Writing Assessment Across the Curriculum

Dudley W. Reynolds
Dudley W. Reynolds is an Associate Teaching Professor of English and Director of Research in English Language Learning at Carnegie Mellon University, Qatar. He is the author of One on One with Second Language Writers: A Guide for Writing Tutors, Teachers, and Consultants and Assessing Writing, Assessing Learning: A Practical Guide for Evaluating and Reporting on Writing Instruction Programs, both published by the University of Michigan Press. He currently serves on the Board of Directors of TESOL.
The past decade has witnessed significant transformations in the multilingual and multicultural educational contexts of the Arabian Gulf. The number of schools at primary, secondary, and tertiary levels where English is a primary language of instruction has increased dramatically. Coinciding with this growth of courses taught in English has also been a diversification of the curriculum. Gulf students no longer have to study abroad in order to study computer science or aeronautical engineering. This growth in the diversity of degrees offered to individuals for whom English is an additional or second language raises significant issues for the field of language testing and assessment because it means that assessment of language competency can not necessarily be separated out from assessment of content learning.

Once upon a time the term English language testing aptly summarized this field since the majority of instruments, research, and publications related narrowly to English as a second or foreign language. The primary object being assessed at that time was knowledge of the language: its syntactic and phonological patterns, lexis, and generic discourse structures. The field’s focus was on measuring what students knew, frequently for the purpose of placing them into an appropriate level of English language instruction. With many secondary and postsecondary institutions around the world today adopting an approach to literacy instruction known as writing across the curriculum (WAC), however, English language specialists can no longer separate what they do from the assessment carried out by the faculty in degree programs.

This paper offers suggestions for conducting assessment in contexts where literacy and content instruction are coupled and in particular how to deal with the need to establish shared principles and procedures among language, writing, and content area specialists. Following a more detailed explanation of the goals and types of writing across the curriculum programs, five strategies for using assessment to promote student development and learning are discussed.

What is Writing Across the Curriculum?

McLeod and Maimon (2000) established a theoretical orientation for writing across the curriculum programs:

WAC is a pedagogical reform movement that presents an alternative to the delivery of information model of teaching in higher education, to lecture classes and to multiple-choice, true/false testing. In place of this model, WAC presents two ways of using writing in the classroom and the curriculum: writing to learn and learning to write in the disciplines (this latter may also be thought of as writing to communicate) (p.579)

What may be most striking about this definition is that writing is not seen as an end in itself, as something to be taught, learned, or measured for the sake of saying that someone “knows how to write.” Rather, writing is a tool that students use to develop their ability to reason and solve problems; it is also a means of access to professional standing in a particular academic discourse community. The implications for pedagogy—and assessment—of this perspective are clear.

From the pedagogical stance of WAC, writing instruction cannot be limited to an initial college preparatory or first-year program. Students continue to learn through and about writing across their academic careers. In science classes, for example, the act of writing up an experiment forces students to organize their actions into a series of identifiable steps and to distinguish results from implications. In engineering courses, thinking about the audience for an abstract of a design project leads students to pragmatic considerations of a customer’s needs and priorities. Students rarely arrive in these classes, however, knowing how to write lab reports or design documents; they must be taught. Spack (1988) argues that such instruction should be left to the
discipline-area faculty because they are the ones with expertise in the conventions and forms of their academic community. Many faculty members, however, may feel uncomfortable trying to incorporate writing instruction into their courses.

When it comes to assessment, the theoretical orientations of WAC programs suggest that writing should not be evaluated as a stand-alone product, as an indicator that a student does or does not know how to write, as in the commonplace writing proficiency exam. Instead, building from the metaphor of writing as a tool, students’ texts should be evaluated in terms of how well they achieve a purpose and in such a way that they can learn strategies and develop understandings of a particular context that will help them with similar tasks in the future. This does not mean that WAC programs can ignore the political realities of having to be accountable for student literacy; it simply means that they construct the standards for literacy in a different way. At Washington State University, one of the best known WAC programs for example, students in their third year must submit a portfolio containing three course papers for evaluation by discipline-area faculty who have received training in the assessment of writing (Haswell, 2001).

Writing across the curriculum programs, therefore, typically involve collaborations between discipline-area faculty and writing faculty at multiple levels. At the curricular level, WAC programs may organize workshops and training seminars for faculty on issues such as working with second language writers; they may involve discipline-area faculty in assessment schemes such as the one at Washington State; or they may seek to designate specific courses in each academic major that require intensive writing. At the course level, collaboration may come in the form of team-teaching, where a portion of the classes are turned over to a literacy instructor; it may consist of a dedicated writing tutor who knows the course and its assignments well, or it may be simply a set of materials or workshops developed by a writing center. As part of this collaboration, the writing faculty can offer expertise in writing processes such as invention, planning, and revising, and more importantly for our purposes, strategies for teaching and assessing students writing in a second language.

Collaboration between faculty with a primary concern for literacy and faculty for whom reading and writing are more or less subconscious processes can be difficult. There may be a tendency for each to say “You worry about your area, and I will worry about mine.” There may also be a tendency for discipline-area faculty to notice language only when it deviates from what they expect, which leads them to see students’ writing as evidence of what the students cannot do rather than what they can. Finally, there may be a tendency for discipline-area faculty to perceive writing assessment as too time consuming, which in turn prompts them to limit the amount of writing in their courses. As suggested in the rest of this paper, however, strategic consideration of assessment goals and processes can help overcome some of these difficulties.

**Strategies for Collaborating on Writing Assessment Across the Curriculum**

The five strategies that follow apply to most large scale test design projects (see for example Bachman & Palmer, 1996). In the parlance of language testing, the strategies include generating a description of the construct to be tested, the use of multiple measures to achieve greater reliability, development of item specifications and explicit measurement procedures, and plans for feedback and impact. These principles for good assessment also provide a useful framework for collaborative discussions between writing and discipline-area faculty at either the curricular or course level. In essence, the principles of good testing can reinforce the principles of good teaching.
Discuss Purpose

Many curricular writing initiatives begin with a discipline-area faculty member seeking help from a writing teacher to improve students’ writing. They often assume that a diagnostic test would reveal what the students do not know so that the writing teacher could provide remediation. This may sound like a legitimate request, but it also reflects a very incomplete view of what it takes to learn to write and also the options for addressing student needs. If the writing teacher rebuffs the need for a test, however, they risk being accused of inadequately responding to the problem. One strategy, therefore, is to agree to work on a test but ask for a discussion of its goals.

Discussions of test purpose force both faculty and writing teacher jointly to consider certain questions. Firstly, are they assessing abilities that they can reasonably expect students to have gained prior to starting a course? If engineering faculty members are concerned about the quality of abstracts, it can be asked when and where students had chances to learn abstract writing. Secondly, is the purpose of the test simply to provide information on individual students, or will the results collectively be used to inform instruction—wherever it exists in the curriculum? Thirdly, will the results of the test be used to inform the students themselves with the expectation that they will then engage in independent learning, or does the instructor have the means to invest sufficient class time for instruction in writing? The answers to this last question may determine what further types of collaboration will be possible. Will self-directed learning materials need to be developed? Is it possible to require students to attend workshops outside of class? Does a department need to develop a communication course specifically for its majors? In sum, discussions of test purpose force a reality check, and they shift the focus from why students cannot write to options for addressing student and programmatic needs.

Use Varied Assessment Instruments

Assuming that faculty and writing teacher can establish a shared understanding of why an assessment of student writing could be useful, the next step is to develop one or more instruments. As noted earlier, what writing teachers bring to collaboration with discipline-area faculty is an understanding of writing as a process, knowledge of issues faced by different groups of writers, and awareness that writing is used for specific purposes in specific contexts. In short, they are more likely to have a developed theory of writing as a complex system. Discipline-area instructors, on the other hand, contribute reliable knowledge about purposes, contexts, and likely audiences, but they may not have ever thought about the complexities involved in the processes of written communication. The discussion about how to assess can help with this.

Out of experience, many people’s idea of a writing test may be the traditional placement essay, a timed response to an intentionally generic prompt. As an indication of what it means to know how to write, placement essays focus attention on the final product and its conformity to usage standards, paradigmatic organizational patterns, and overall persuasiveness. They generally fail to reveal much about the breadth or versatility of students’ abilities or their command of distinct writing processes such as invention and audience analysis (Hamp-Lyons, 2005; Raimes, 1990). The assumption is simply that these are necessary to produce the finished product. One way to help faculty, and students, develop a more detailed theory of writing is to employ instruments that assess different writing abilities from those emphasized by timed writing samples. Consider, for example:

**Behavioral surveys.** Students may be asked to: rate the frequency with which they imagine someone reading their writing; ask themselves questions while writing; use visual schematics to plan, move sections of text while writing, and so on.

**Conference records.** If opportunities exist for students to talk one-on-one or in small groups with a writing tutor, then it may be possible to record or video-tape these sessions. These conference records can then be coded using a checklist with multiple purposes: to deal with issues that
arise, such as the writer’s understanding of the task; to develop language-focused discussions on problematic areas such as verb tenses; or to focus on strategy use, such as discussing ideas with another person. The frequency with which different topics appear in the conference records can then be used to describe the group as a whole, compare more and less successful samples of writing, or possibly examine the effectiveness of the conferencing method.

**Objective tests.** Although writing clearly depends on an ability to perform, many readers have expectations for texts that incorporate what might be called discrete-point knowledge. Readers expect adherence to conventions for spelling, grammar, mechanics, and style sheets. They also expect precise use of technical vocabulary. Multiple-choice or other controlled-response tests signal that good writers can consciously draw on such knowledge.

**Portfolios.** Portfolios can consist of collections of finished products produced throughout a course or they can include multiple drafts of one or more papers. Either way, they shift the focus from inferences that can be drawn about a piece of writing to a broader view of the writer’s development over time.

These suggestions focus on the abilities, behaviors, and knowledge of the students as writers. If our goal is to broaden understandings not only of what students learn but also how, then it is also useful to assess the instruction they are receiving. Schools typically employ course evaluation forms for this purpose, but the information provided by these forms is generally much more useful for accountability than as a tool for course improvement. More formative options include:

- asking students at the end of each class to write one or two things they learned that day as a means to assess the saliency of topics;
- coding a recording of a class or writing conference for teacher actions (e.g. explanation, questioning, correction) and evidence of student uptake (e.g. reformulations, follow-up questions)—similar to the use of conference records described above—as a means for comparing teacher emphasis with student interest;
- creating course-specific evaluation forms that target actual course activities (e.g., writing conferences, teacher explanations, peer evaluation).

There may be instructors and students who assume that writing is learned simply by doing and that the main function of a writing class is to provide practice rather than instruction. Broadening what and how we assess identifies behaviors that students can be taught to use or change; it also clarifies the options that a teacher has.

**Develop Rubrics**

Building on the discussions about why and how to assess writing, the next step is for writing teachers and their faculty colleagues to consider what they most value in the finished product. Many faculty members grade papers by writing comments as they read in response to particular sections of a paper or by inserting a general, summative comment at the end (Ferris, 1997). The comments may focus on needed revisions, express agreement or disagreement with the writer’s position, or provide an explanation that would be useful for the writer. Because they are created in response to an individual text, however, they do not really require principled analysis of what are essential elements for a successful response to the assignment. They may also focus much more on producing a better version of the text than a better writer (Sommers, 1982).

Grading rubrics provides a commonly cited way of encouraging students to abstract from a particular text in order to broaden their understanding of writing in general and deepen their understanding of the specific genre represented by the text (Coombe, Folse, & Hubley, 2007; Weigle, 2002). When discipline-area faculty members are engaged in developing a rubric for an assignment, they are forced to make similar abstractions. They must first decide on the traits
that are important for the assignment. Rubric traits should represent the instructor’s theory for how effective writing is recognized in his or her field. Again, perhaps what is most important is for the instructor to realize that there are options. For example, many rubrics commonly used by language teachers include traits related to content, organization, language control, and sometimes vocabulary. These traits apply to texts as a whole and emphasize the writer’s role in constructing the text. A rubric for a biology lab report, on the other hand, might label traits according to the sections of the report, such as abstract, method, and results. By communicating that individual sections can be assessed according to multiple criteria, this system stresses the need for knowledge of the discourse community’s norms. Finally, traits may be worded in terms of a reader’s likely perceptions, such as persuasiveness, credibility, usefulness, ease of access. This rubric emphasizes the communicative nature of writing and the importance of understanding the reader’s reasons for reading.

There is no best choice for rubric traits, but considering the options leads to discussion of whether the purpose of writing in a course is to facilitate student thinking about course topics or to prepare them for workplace tasks. It also further underscores the message that assessment should serve to improve the writer more than the individual piece of writing.

**Choose a Measurement System**

Whereas traits represent a theory of how texts are constructed or used, the rubric’s system for measuring performance constitutes a stance on the possibilities for cognitive development. Bringing a rubric’s algorithm for converting qualitative judgments into quantitative representations out into the open as a topic for discussion forces consideration of how we want students to change as a result of instruction and course experiences.

The first question to ask is whether the grade should reflect an overall impression of the product or can be determined by attending to individual traits (however those have been defined). In the language of testers this is referred to as holistic versus trait scoring (Hamp-Lyons, 1991). Although this may seem difficult to answer, it may be helpful to consider whether the pedagogy has discussed writing in terms of overall effects and purposes or in terms of component parts (e.g., steps in a process, requirements for different sections, organizational strategies). It may also be useful to consider whether you can make claims such as the text is strong in this area and weak in another or vice versa. Many discussions of writing assessment take this discussion one step further. It is asked if different areas or traits should be weighted differently; for example, is appropriate bibliographic form as important as a consistent organizational strategy? Furthermore, it is questioned whether or not objective correlates for different performance levels can be determined; for example, what signals high, average, or low persuasiveness? Ultimately decisions about how to compute the grade rest on how fine-tuned we feel our judgments of a text can be.

This leads to the next question: what should testers base their judgments on? How do they determine their expectations for performance? For language testers, the measurement choice is between norm-referenced, comparing the individual’s performance to that of a population sample, and criterion-referenced, comparing the individual’s performance to a theoretical model of performance (McNamara, 2000). When discussing the choices for measuring writing with people outside the field of educational psychology, it may be more helpful to ask whether we should compare students to themselves at an earlier time, to other students who have completed the assignment, or to criteria that we have established based on our expert knowledge. Whether we have multiple drafts or work from other students may answer this question; nevertheless, asking the question often points out whether we are relying on explicit or implicit norms for development.
Make Assessment Cyclical

The final strategy for working with discipline-area colleagues is to return to the initial question regarding the purpose of assessing writing, but this time with a much more practical consideration of how the results will be used and their likely impact. If the discussion started because a colleague was worried about his or her students’ writing, will the instruments and measurement system we have created address that need? Too often we may assume that assigning a writing task and then grading sufficiently absolves us of all responsibility to our students. Clearly, we cannot force students to take advantage of our guidance, but we can make it easier for them.

Rubrics are a first step in this direction because they channel our feedback and comments into categories of information that are easier for students to take in and apply the next time around. The next step is for writing teachers to discuss with discipline-area colleagues what the next time around will be. If the course structure permits, they may jointly want to create a series of assignments that build off of each other. They might suggest writing sections of a longer document before assembling the whole; another possibility would be similar assignments with varied audiences. If it is not possible to make connections between the assignments, then we should find ways to make students aware of similar academic and professional writing tasks that they may encounter at some time in the future similar to the socioliterate approach advocated by Johns (1997). Finally, we may want to discuss whether students are likely to see writing as important for their futures. Will they be motivated to improve their writing, and if not, how can we change this? Only when we begin to talk about the possibilities for assessment to lead to change do our efforts become meaningful.

Faculty Across the Curriculum

Involving discipline-area faculty and writing teachers in discussions about writing assessment can overcome concerns about the burden of grading and the natural tendency to focus on what students are not doing rather than what they can do. At the same time that these discussions provide practical methods for generating useful feedback, they build vocabulary and conceptual awareness. One subtext throughout this list of strategies, however, has been the need to consider carefully our audience—the faculty “across the curriculum”—what they are likely to know, at varying levels, about writing, the development of writing abilities, and the pedagogic options. For WAC to have persuasive appeal, it must be grounded in concrete examples and oriented towards results. It must eschew the vocabulary of educational measurement and assume the mantle of promoting student learning.

References


