



Introduction

BRIAN HUOT

Kent State University

We are pleased to introduce the first issue of the second volume of *The Journal of Writing Assessment (JWA)*. This issue exemplifies the kinds of scholarship in writing assessment that motivated Kathi (Yancey) and I to found the first writing assessment journal more than 10 years ago. All three articles stake out an important set of issues different from each other but important to an understanding of the field. We recognize that writing assessment is a diverse field that draws interested scholars from a network of connected disciplines that have their own journals, conferences, and literatures. Although it is certainly in the best interest for writing assessment scholarship to enjoy the rich and diverse attention from scholars with a diverse array of beliefs and assumptions about language use in and outside of the classroom and especially about its teaching and evaluation, this strength of the scholarship in writing assessment is also one of its glaring weaknesses. Although it is true that there exists a wealth of scholarship addressing such diverse issues as the application of physical science theory to writing assessment and the differences in rater behavior and cognition of large-scale writing assessment raters—an essay on each of these subjects appears in this single issue—the truth is that for the most part, scholars like Edward Wolfe and Bob Broad and Michael Boyd, let alone Edward M. White (the author of the third article in this issue) do not usually publish in the same journals nor are they read by the same audience. Writing assessment’s diversity also promotes its insularity. Those of us who work in writing assessment tend to come from one camp or another (Huot, 2002; Moss, 1996; White, 1993), and we end up representing different, although not unimportant concerns. *JWA*, I believe, is one of the few academic journals where someone can encounter a wide range of scholarship and scholars all devoted to increasing our understanding about writing assessment and ensuring the continued creation of better and better assessments. Of course, this doesn’t mean that this issue or even this journal can end the problems of insularity in writing assessment scholarship. On the other hand, creating a venue like *JWA* that celebrates a range of positions, research questions, methods, viewpoints, and voices can only enrich the understanding of all of us who work in writing assessment and aid in the important and difficult task of representing writing skill and making informed decisions about that skill.

In addition to providing writing assessment scholars and practitioners (not that there is always a difference between the two) with an important, wide-ranging resource for scholarship, *JWA* also wishes to announce the creation of the Stephen P. Witte Award for Excellence in Writing Assessment Scholarship to be given on an annual basis (when merited) to the most important scholarly contribution to writing assessment. The period for the first award will be 2000–2004, if the journal-appointed committee should deem a worthy recipient. After the initial award, the award committee will consider scholarship from a 2-year period. The Stephen P. Witte Award will include a commemorative plaque and a cash award of \$500 supplied by Hampton Press. Stephen P. Witte was a pioneer in writing assessment scholarship who passed away in Spring 2004. Witte's contributions to the field of writing assessment include three co-authored articles in a 4-year period for the then first and fledgling journal, *Assessing Writing*, during the time Kathi Yancey and I founded and edited that journal. Not only did Professor Witte support all of those who work in writing assessment, he was also a strong supporter of my and Yancey's efforts to bring important writing assessment scholarship to the field. This award honors Stephen P. Witte's contributions to writing assessment and to the efforts of junior scholars (at the time) to establish the importance and efficacy of scholarship and scholarly journals dedicated to writing assessment. All nominations, including the name of the scholar, the piece of scholarship and a justification for awarding the prize to your nominee should be sent to me, the editor of *JWA*.

In addition to the pleasure of bringing a wide range of scholarship in this the second issue of *JWA* and announcing the new and only award for writing assessment scholarship that honors the contributions of Stephen P. Witte, I also take great pleasure in introducing the articles for this issue. The first article is "Rhetorical Writing Assessment: The Practice and Theory of Complementarity" by Bob Broad and Michael Boyd. Broad and Boyd's article focuses on the gains made in writing assessment over the last couple of decades away from single-shot impromptu writing scored on a numerical rubric with which raters are trained to agree. Broad and Boyd single out two important changes in some writing assessment procedures: the use of multiple texts (portfolios) to make an assessment decision and the use of multiple readers to make that decision, what they call portfolios and communal writing assessment. This article explores the importance of establishing a firm theoretical and epistemological basis for writing assessment theory and practice. According to Broad and Boyd, unless we establish the theoretical superiority of portfolios and communal assessment, we will ultimately lose these writing assessment options, because they are not supported by psychometric versions of writing assessment or practical, pragmatic approaches interested in producing assessments for the most number of people for the least expense. On the other hand, from a theoretical basis and a consideration of validity as inquiry into the decisions made based on a specific measure, portfolios and communal writing assessment should remain important writing assessment options for assessments that attempt to represent student writing and teachers' reading of that writing in the most sophisticated and nuanced ways.

Edward M. White's article, "The Misuse of Writing Assessment for Political Purposes," examines some major uses for writing assessment and, in his discussion,

underlines how these uses for assessment are not always true to the field's understanding of best practices nor does it represent an understanding or practice of validity theory. Specifically, White looks at state-mandated writing assessments (the subject of two articles in *JWA* 1.1), writing placement exams for entering college students, and mid-level college or exit writing assessments that are often required for student graduation from college. White affirms the consequences for students, teachers, and institutions of all of these writing assessments. "Each of these assessments represents a gate through which students must pass if they are to gain access to the privileges and enhanced salaries of college graduates, and so they carry a particular social weight along with their academic importance. In other words, each of these tests carry significant consequences or high stakes." Although White defends the importance of using writing assessment to teach writing inside the classroom and to document the how well writing programs are serving students, he looks at various practices and trends in these three assessment cites and argues that political expediency and posturing often drive assessments that underrepresent the act of writing and ultimately produce decisions about students that cannot be supported by writing assessment theory and practice. Ironically and importantly, White's arguments about test use and the political issues and pundits that drive this use lead us back to considerations of validity because when we examine current notions of validity theory, we are reminded that validity is about the decisions being made based on a measure and that each use of a test must receive adequate inquiry into their validity.

Edward Wolfe's article provides an interesting look into the different ways raters process their readings and evaluations of student writing within a large-scale, psychometric writing assessment, depending on their ability to read student writing effectively in a particular assessment context. Successful raters who can agree with each other at a high rate read and think about their reading in different ways than those raters who are less able to agree on rankings for the same papers. Successful raters tend to base decisions more holistically about the entire paper, reserving judgment until the paper has been completely read, whereas less successful raters render judgments more atomistically and then attempt to combine these partial judgments into an overall score for the paper. Of course, Wolfe's findings confirm the advice and direction that the inventors (Godshalk, Swineford & Coffman, 1966) and early promoters (Myers, 1980) of holistic scoring have been saying for years—that the most reliable judgment is a single "holistic judgment" based on a clearly defined and articulated set of scoring guidelines. In addition to confirming the early advice of holistic scoring users, Wolfe's research outlines the cognitive demands of being able to reserve judgment for an overall understanding of a student's text. Wolfe's other findings, like the importance of learning to think within the language of the scoring guideline, also supports conventional holistic scoring about the importance of using rubrics to train raters to agree. In addition to documenting how successful raters score student writing in large-scale assessments, Wolfe also documents that successful raters' reading behavior and cognitive processes are an acceptable way of making meaning from text and providing a judgment about that text. The overall picture of rater behavior provided in this article

can be used to revise scoring procedures that encourage a greater agreement among raters and support defensible practices for reading and rating student writing.

Anthony Edgington's review of Lad Tobin's book leads us into the same territory explored in Edward Wolfe's article, but within a different context and from a different perspective. Instead of looking at the cognitive processes for reading and rating student writing in large-scale psychometric writing assessments, *Reading Student Writing: Confessions, Meditations, and Rants* looks at classroom contexts and relates personal accounts of interactions with students and their texts, urging us to examine our underlying postures and psychological issues in order to truly understand the nature and value of our responses to student writing. Edgington's reading of Tobin's book causes him to reflect on some of the underlying motives he might have had about the way he had designed and conducted his study, which looked at the talk-aloud protocols of teachers reading their students' writing in a normal instructional context—in a class they were teaching and in the same place they normally used to read and respond to student writing. Edgington's appreciative reading of Tobin's book and his own willingness to use Tobin's message to reflect on his decision making in his recent empirical study illustrates the value of a varied and robust writing assessment literature in which lessons learned in classroom contexts can also inform decision making in large-scale assessments or empirical research studying the way teachers respond to their students.

All of the articles in this issue are ultimately about the validity of specific writing assessment practices and programs. Whether we consider Broad and Boyd's ideas about complementarity in portfolios and communal writing assessment, White's misuse of writing assessment for political purposes, or Wolfe's delineation and documentation about the way raters read and score student writing in large-scale writing assessments, we are ultimately talking about the degree to which a measure can support certain kinds of decisions. In addition, the segment of annotated bibliography we include in this issue is the second half of the entries on validity and reliability. To come back to my earlier point about the compartmentalization and insularity of much writing assessment scholarship, I think a review of the contents of this issue of *JWA* is a good illustration of how when we consider a broad range of theoretical and research perspectives can we expect to create a writing assessment scholarship representative of all the nuances and complexities that actually support defensible and valid decision making in writing assessment.

Before getting out of your (the reader's) way, I would like to make an announcement and acknowledge a couple of debts for this issue. After being a faculty member in the English Department at the University of Louisville, I joined the faculty in the English Department at Kent State University for the Fall 2004 semester. Although the University of Louisville supported my efforts as a journal editor for 5 years or so, it had been unable to support either *Assessing Writing* during the last 3 years of my tenure as editor or in the time I remained at Louisville and began editing *JWA*. Fortunately, Kent State is supporting *JWA*, and for this reason, all manuscripts and correspondence for *JWA* can be directed toward my Kent State address listed on the journal's inside page under information for submission. I would like to thank the University of Louisville for the support it rendered in the past and my new academic home Kent State University for supporting the current

publication of *JWA*. I would also like to thank Barbara Bernstein for all of her help and support with *JWA*—as well, I thank her for supporting The Stephen P. Witte Award for Excellence in Writing Assessment Scholarship. My thanks also go out to Michael Neal, *JWA*'s managing editor. My last and most important thanks is for you our readers, that whole range of scholars who make this journal and a bright future for writing assessment a real possibility.

References

- Godshalk, F. I. et al. (1966). *The measurement of writing ability*. Princeton, NJ: Educational Testing Service. CEEB RM No. 6.
- Huot, B. (2002). *(Re)Articulating writing assessment for teaching and learning*. Logan: Utah State University Press.
- Moss, P. (1996). Enlarging the dialogue in educational measurement: Voices from interpretive research traditions. *Educational Researcher*, 25(1), 20-28.
- Myers, M. (1980). *A procedure for writing assessment and holistic scoring*. Urbana, IL: National Council of Teachers of English.
- Smith, W. L. (1993). Assessing the reliability and adequacy of using holistic scoring of essays as a college composition placement technique. In M. Williamson & B. Huot (Eds.), *Validating holistic scoring for writing assessment* (pp. 142-205). Cresskill, NJ: Hampton Press.