Elliot, Norbert, and Les Perelman, eds. *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. New York: Hampton Press, 2012. 518 pages. $42.50. ISBN 978-1-61289-087-6.

Reviewed by Chris Blankenship

White's Law: "Assess thyself or assessment shall be done unto thee." This pithy phrase, uttered on the Writing Program Administrator Listserv in 1996, has become one of the most widely referenced maxims among scholars and teachers of writing, no less for its truth than for the deep respect our community holds for its scribe, Edward M. White. For over 40 years, White has been a champion of meaningful writing assessment, working tirelessly to strike a balance between the best pedagogical practices in rhetoric and composition with the realities of a growing culture that desires cost-effective evaluation and accountability. This volume of 27 essays in four sections honors White's contributions to the field of writing assessment; however, as Elliot and Perelman point out, it is not simply a tribute, but "a tribute by emulation" (13), each essay building upon existing scholarship. The authors, academics in rhetoric and composition and educational measurement professionals alike, are united in their admiration of White's work. Although they often disagree with one another, and, at times, with White himself, the editors hope that the book serves as "evidence of a narrowing gap between the two communities" (13).

The first section of the book consists of seven essays that provide the reader with historical and contemporary contexts for writing assessment. On the academic side, John Brereton begins the section with an examination of the late 19th century Harvard entrance exam in writing. He demonstrates that, like so many tests today, reports of extreme rates of failure were likely exaggerated, although these exaggerations led to what is perhaps the first instance of writing assessment driving course development and curricular change as evidenced by the creation of

first-year writing courses in 1884. This history is extended in Margaret Hundleby's description of the development of assessment in technical and professional communication, including its similarities to and divergences from mainstream composition studies. Barry Maid and Barbara D'Angelo also delve into mainstream composition studies by comparing the Council of Writing Program Administrators' Outcomes Statement to the Association of College and Research Libraries' Information Literacy Competency Standards for Higher Education to trace the development of assessment standards regarding digital media and genre studies. They conclude that the best outcomes to apply to assessment practices are those that are flexible and dynamic, demonstrating this principle through an example proposal writing course at Arizona State University.

The remaining chapters in this section are written by educational measurement professionals and provide a broader educational context of writing assessment. Sherry Seale Swain and Paul LeMahieu trace the development of the National Writing Project's analytical writing continuum, providing an in-depth discussion of their current evaluation methods and the program's impact upon teachers. Similarly, Hilary Persky traces key moments in the development of writing assessment by the National Assessment of Education Progress, focusing on the tension between large-scale writing assessment and the need to use meaningful writing tasks. Outside of higher education, Paul Deane considers writing assessment in primary and secondary schools. He outlines the CBAL (Cognitively Based Assessment *of*, *for*, and *as* Learning) model, which draws on activity theory to situate the assessments of reading, writing, and critical thinking in a social context. Deane provides an example of this assessment from one of their pilot programs in 8th grade persuasive writing. Finally, Mary Fowles considers how assessment scholarship has impacted tests for graduate and professional programs, posing a set of critical questions and an explanation of how the designers of the Graduate Record Examination (GRE) went about addressing these concerns.

The second and largest section of the book takes a broad view of contemporary assessment, detailing assessment strategies, complicating existing concepts, and asking critical questions. The most prominent method of direct writing assessment today is the portfolio, and the most famous and longstanding program in portfolio assessment comes from Washington State University. Diane Kelly-Riley provides an update to Haswell's description of the program in *Beyond Outcomes: Assessment and Instruction within a University Writing Program*, focusing on how the expert rater system has inspired greater faculty ownership of the program and increased innovation and effective documentation of their efforts. It is this innovation and improved research through effective documentation that William Condon advises scholars to continue in his cautionary note about portfolio assessment. He demonstrates how other writing assessment methods have fallen into the "trap of efficiency and a reductive kind of cost-effectiveness" (243), resulting in stagnation and misalignment with pedagogical best practices. Assessments and curriculum must be intimately connected, as Irvin Peckham explains in his chapter, providing an overview of the Online Challenge and Semester Assessment protocols used at Louisiana State University for placement, course exemption, and program assessment. These systems insist upon pedagogical alignment, "matching the genres being assessed with the genres being taught" allowing for "greater construct representation" (180).

Assessment within the classroom itself is also an important part of this field, although, as Chris Anson points out in his chapter, there is a relative dearth of research on the effects of teacher response on student development. He calls for more student voices in the research as well as a greater breadth of information gathered when conducting the research. This response to student writing may not always come from teachers, though. The essay from Jill Burstein, the only educational measurement professional in this section of the book, provides a positive evaluation of E-rater, an automated essay scoring (AES) system used in the ETS Criterion classroom software. She argues that the system is meant

to support people, not replace them, and that AES can be useful, helpful, and dependable. She also engages Anne Herrington and Charles Moran, whose essay directly challenges the notion that AES can be anything but harmful in its current form. Also using E-rater and a timed essay written by one of the authors, Herrington and Moran point out numerous misidentified errors and poor advice given by E-rater. They also criticize the implicit ideology built into the system that there is one "standard English," and problems E-rater has in helping writers whose first language doesn't fit this standard.

The final three chapters in this section work to complicate our notions of what makes good assessment. Jon A. Leydens and Barbara M. Olds outline an assessment project in a first-year humanities course that both failed and succeeded. They discuss what they learned in this project, asserting that "there is a clear difference between poor assessment and disappointing results" (248). Bob Broad's approach to assessment moves beyond rubrics and holistic assessment through Dynamic Criteria Mapping (fully explained in his 2003 book *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*). After providing a brief overview of DCM, he then points out all that it has in common with White's holistic scoring, despite the on-face disagreements. Lee Odell's essay also takes up this reconceptualization of assessment by looking for a standard language that will let us find "elements of similarity among different types of composition without oversimplifying or ignoring significant differences" (273). His "given-new contract" works toward this goal, and much of the chapter is devoted to demonstrating its utility in several pieces of real student writing.

The third section of six essays shifts focus from strategy and philosophy to examining the impacts of writing assessment, moving from broad considerations of evaluation and public perception to particular consequences dealing with race and native language. Peter Elbow begins this section with a discussion of "good enough" evaluation, striking a balance between the need for assessment and the potential harm from untrustworthy results. He

points out three common traps into which teachers fall in classroom evaluation and then extends his analysis to placement tests, program evaluations, and large-scale national writing tests. Doug Baldwin, a program director at ETS, pursues a similar topic in his essay about the challenges associated with large-scale writing tests, such as promoting validity and reliability, administering assessments in international settings, and the relatively new problem of "canned" essay responses. This concept of fairness and transparency is echoed in Daniel J. Royer and Roger Giles' chapter, which explains the advantages of making the private practices of a program, such as placement and evaluation, public for all stakeholders.

The second prevailing thread in this section is that of the intersections of race and language with writing assessment practices. Asao B. Inoue and Mya Poe's essay revisits the study by White and Thomas (1981) examining California's English Placement Test (EPT) scores by racial categories. They find that racial formations still play a role in EPT scores and caution readers about the ways that tests can produce "educational environments that could be unequal, either in terms of access, opportunities, or possibilities" (353). Gita DasBender finds a similarly complex situation in using Directed Self Placement with generation 1.5 students, a term she uses interchangeably with "multilingual students" to acknowledge the diversity of these language learners. Her research uncovers that the best way to serve the needs of these students is to seek out a more nuanced understanding of their experiences with language and literacy. Liz Hamp-Lyons also writes about multilingual students but in the context of her work designing a school-based assessment (SBA) of spoken English in Hong Kong. Drawing on her experience with writing assessment, Hamp-Lyons departs from White's holistic scoring model in the SBA, preferring a multi-trait scoring model. She argues quite strongly for her approach, remarking that "holisticism is not teacherly" (395) because it is focused more on the reader than the writer. While giving credit to White, she concludes, "you cannot build a sturdy house with only one brick" (395).

The final section of this volume looks towards the future of writing assessment. As Elliot and Perelman observe in the introduction, "if writing assessment is to have a valid use, the role of classroom practice must inform the assessment design" (409). As a result, this section is the only one in the book with all chapters written by teachers of writing and rhetoric. Rich Haswell begins this section by discussing the use of "rhetorical numbering" as an anticipatory counter argument against external writing assessment being imposed upon a program. He provides five brief stories of WPAs who have used the strategy to great effect. While he does acknowledge that there are potential pitfalls in this philosophy, he concludes, "when the currency is numbers, it is not a good time to be numerophobic" (423). Rhetorical numbering can help to protect programs against mass-market writing assessments, which, as Les Perelman warns in the next chapter, are bullshit. Using Harry Frakfurt's 2004 definition of bullshit as being simply unconcerned with the truth, Perelman points out several distinct ways in which these organizations have "whole-heartedly embraced bullshit" (427), such as the use of automated essay scoring, requiring human essay scorers to ignore factual inaccuracies in essays, and even in the writing of their own organizational research reports. Even if writing programs manage to avoid these mass-market tests, Peggy O'Neil points out that local writing assessments can also have negative or not clearly positive results as well. Using two WAC/WID programs as examples, O'Neil examines the various ways that assessments can frame writing programs. She concludes that "assessment frameworks need to be flexible, allowing for additions and renovations," although, even then, it can be difficult to "meet the demands of both disciplinary standards and institutional realities" (453). Reframing the language around writing assessment is one possible solution, Cindy Moore argues. Using examples from the feminist movement's successful work, Moore makes a strong case for changing the language surrounding writing assessment to empower our efforts to make assessments meaningful and sustainable. In the final chapter of this volume, Kathleen Blake

Yancey describes the nascent fourth wave of writing assessment as being characterized by greater involvement of the federal government in higher education assessment but also by the drive of a "*self-created exigence* independent of any specific local need" (477). Using the two examples of Portnet and the Inter/National Coalition for Electronic Portfolio Research (INCEPR), Yancey demonstrates the promise of future writing assessment research in pointing the way "toward new ways of identifying critical questions, of organizing methods to inquire into them, of producing new knowledge, and of designing new ways of sharing both their knowledge and progress" (489).

As any book honoring an exemplary scholar like Ed White should be, this robust volume is notable in the breadth of its coverage, the depth of its insights, and the rigor of its methods. The authors include an amazing array of widely respected writing assessment scholars, exemplary teachers, and experienced educational measurement professionals, all of whom contribute admirably to the text. While Elliot and Perelman's assertion of "tension between the composition community and the educational measurement community" (407) is a bit of an understatement at times (see Herrington and Moran's direct attack on AES and Burstein's aggressive defense of it), I believe that the editors have accomplished their goal of demonstrating a narrowing gap between these two groups. The cross-references between chapters is generally friendly and respectful, and the engagement of different positions is admirable, although, as White himself points out in the afterword, there is often a relative lack of cross-citations from the scholarly journals of the other community in the individual chapters.

Holistically speaking, I highly recommend *Writing Assessment in the 21st Century* to any teacher of writing at any level. The content is always quite accessible for novices in writing assessment practices while still being useful to those who have more experience in the field. Due to its breadth, this book is also an excellent reference guide for additional resources in writing assessment theory and practice. Although much of the content

focuses on practices in higher education, the strategies, techniques, and lessons in this volume can be applied at any level to design, implement, and sustain systems of writing assessment that can provide meaningful data for program development and student learning.

Works Cited

Broad, Bob. *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*. Logan, UT: Utah State University Press, 2003. Print.

Frakfurt, Harry. *On Bullshit*. Princeton, NJ: Princeton University Press, 2004. Print.

Haswell, Rich. (Ed). *Beyond Outcomes: Assessment and Instruction within a University Writing Program*. Westport, CT: Ablex, 2001. Print.

White, Edward M., and L.L Thomas. "Racial Minorities and Writing Skills Assessment in the California State University and Colleges." *College English* 43.3 (1981): 276-283. Print.