

Introduction to a Special Issue on a Theory of Ethics for Writing Assessment

by Diane Kelly-Riley and Carl Whithaus, Editors

Editors' introduction to the Special Issue on a Theory of Ethics for Writing Assessment.

Last March, we met with the co-authors of this Special Issue on the Theory of Ethics for Writing Assessment at the CCCC meeting in Tampa, Florida. Our discussion was wide ranging, but it centered on concerns about how fairness had been underplayed in the research literature on writing assessment. Drawing on educational measurement discourse, reliability and validity had occupied primary locations within the field of writing assessment. In the 1990's educational measurement scholars had explored the implications of including consideration of consequences in testing situations. Lee Cronbach (1989) observed, "tests that impinge on the rights and life chances of individuals are inherently disputable" (p. 6) and that adverse social conditions, in and of themselves, call the validity of a test use into question. At that time, educational measurement scholars debated the merits of including the consideration of consequences for validity. Scholars including Pamela Moss, Samuel Messick, Robert Linn, and Lorie Shepard advocated that consequences be explored so the impact of tests on students could be considered, and that such a focus was a worthy investigation within the framework of validity research. Others scholars, such as Michael Kane, William Mehrens, and James Popham, argued that while important, consequences should be considered in some form, but not under the auspices of validity inquiry. These debates have continued to evolve within the field of educational measurement, and are reflected in their most current form with the 2014 publication of the sixth edition of *The Standards for Educational and Psychological Testing*, which elevates the consideration of fairness as primary concern in testing situations. In the new *Standards*, fairness is defined as the

use of test score interpretations for intended use(s) for individuals for all relevant subgroups. A test that is fair minimizes the construct-irrelevant variance associated with individual characteristics and testing contexts that otherwise would compromise the validity of scores for some individuals. (p. 219)

The conversation of consequences shifted to articulations of fairness, and while this has been an important shift, the authors of this special issue argue that consideration of fairness is not enough. The next phase must take up larger questions of involving the ethics of assessments

In writing studies, we've observed similar evolutions in how we wrestle with the testing situations those in our field often oversee. The important, vital turn towards paying more attention to validity--and to use validity--marked by the publication of Brian Huot's *(Re)Articulating Writing Assessment for Teaching and Learning* in 2002 and the theoretical and pragmatic development of those ideas in works such as Bob Broad's *What We Really Value* (2003) and his *Organic Writing Assessment* (2009) did not seem to fully address the underlying tensions emerging in public debates about large-scale standardized testing, currently embodied in the Common Core State Standards and their assessments and the myriad of locally developed writing assessment enacted in post-secondary settings. In fact, as a research field, writing assessment seemed to have skirted the issues of fairness and ethics in favor of in-depth discussions about the need to make sure test designers were developing valid as well as reliable assessments.

At the current moment, adding--or returning to--humanities-based concerns about fairness and ethics seems essential to us and the authors of this special issue. Writing assessment, perhaps even more so than writing studies as a larger field, balances between the detailed, idiosyncratic, individualized approach of the humanities and the generalizable, policy-driven and policy-setting research agendas found in the social sciences. For writing assessment in the age of the Common Core State Standards, the impulse to draw on rigorous approaches to fairness informed by ethical theories from within philosophy, the law, and other humanities fields offers an alternative to continuing down the well-trodden path of scholarly discourse shaped more by those working in educational and psychological measurement than those working in writing studies.

While the discussion at CCCCs in Tampa returned again and again to ethics and fairness, concerns about the effects of tests are not entirely new to educational measurement nor the subfield of writing assessment. Samuel Messick, one of the leading assessment researchers of the 20th century, argued in the 1980s that implementing an assessment program without first assessing its effects on students, teachers, and systems of education is like releasing a new drug on the market without first understanding either its properties or its effects on people. He wrote,

Using test scores that "work" in practice without some understanding of what they mean is like using a drug that works without knowing its properties and reactions. You may get some immediate relief, to be sure, but you had better ascertain and monitor the side effects. And although evaluation of side effects--or more generally, of the social consequences of the testing--contributes to score meaning, it is a weak substitute for score meaning in the rational justification of test use. (Messick, 1989, p. 15)

His argument highlights an important difference between the field of medicine and the field of educational measurement: one is

governed by a well-established code of ethics, the other is not. This is not to suggest that those who work in the field of educational measurement are not ethical, but rather that this field is not governed by a well-defined and broadly accepted theory of ethics: We have no equivalent to the Hippocratic Oath. Addressing this lack is part of the impulse behind this special issue: What would a defined theory of ethics and set of ethical professional practices mean for the field of writing assessment? What implications would the development of an ethics for writing assessment, particularly one informed by legal concepts and theories of social justice, have for the design and implementation of both large-scale and classroom writing assessments? Those are key questions, and ones the authors in this special issue work towards.

However, those are not the only questions these authors wrestle with. In fact, their work is richly contextualized within the history and field of writing assessment. They move us towards fairer and more ethical writing assessment by exploring the following questions:

- Why is it so difficult for the testing industry and others to address problems with current large-scale writing assessment models?
- What key ideas should guide a theory of ethics for the field of writing assessment?
- How might a theory of ethics change how we design classroom writing assessments?
- How might a theory of ethics enable us to better attend to the needs of the diversity of students in our classrooms?
- How might a theory of ethics enable us to better advocate for changes in the design of large-scale writing assessments?

Appropriately enough, they do not answer these questions in a color-by-numbers based approach. Rather, these questions inform each of the articles in slightly different ways. They reach towards ethical and fair approaches to writing assessment for students, teachers, and educational policy makers by considering nuanced and contextualized issues.

Their work opens with David Slomp's "Ethical Considerations and Writing Assessment," which lays out the context of large-scale writing assessment in North American settings and details the ways in which large-scale standardized testing have become so prevalent in the landscape of primary, secondary, and post-secondary education. His article begins to lay the groundwork, identifying the conditions that ask us to think about how a test can be fair yet not ethical at the same time. In the next article, "A Theory of Ethics for Writing Assessment," Norbert Elliot meticulously details a theory based on fairness and the structure of opportunity "created through maximum construct representation under conditions of constraints...to the extent which benefits are realized for the least advantaged" to articulate a vision of a theory of ethics enacted in writing assessment. Elliot illustrates his theory through an extended example, and positions the field of writing studies to take up further inquiry. Slomp's and Elliot's pieces serve as framing articles. They address not only the stakes of developing fairer writing assessments but sketch out pragmatic and theoretical ways to consider ethics in the development of writing assessments.

The next set of articles respond to essential considerations raised in Slomp's detailed examples and in Elliot's articulated ethical theory. Slomp's second piece in the special issue, "An Integrated Design and Appraisal Framework for Ethical Writing Assessment" explores how the integrative framework of ethics allows for more complete understanding of measurement concepts in action. Slomp provides an extended example using a case study of the Alberta English 30-1 diploma exam (equivalent to a US graduation test) as well as an assessment design project at the middle school level. Mya Poe and John Cogan explore ethics from the standpoint of legal frameworks, and see that fairness of writing assessment practices can be examined through "burden-shifting heuristics" like those used of the Office for Civil Rights to consider disparate or unequal impact of exams. They suggest methodological approaches that can be used by other writing studies practitioners and researchers. In "This is Not Only a Test: Exploring Structured Ethical Blindness in the Testing Industry," Bob Broad explores the question of how a test can meet technical specifications and thus be considered "fair," but still not be ethical. Broad examines Pearson's and ETS's articulation of fairness through their company documents, and asserts the need for some sort of governing entity to be an advocate for test takers in the process. In "Decolonizing Validity," Ellen Cushman responds to shortcomings in current conceptualizations of validity both as a Composition and Rhetoric scholar and as a member of the Cherokee tribe. Finally, the authors of this special issue wrap up this volume with a forum, which provides a leaping off point for other researchers. They explore the implications of shifting demographic trends in North America; changing educational and political landscapes and the implications upon our research agendas; and ways in which the humanities-based approaches can facilitate our collective inquiry.

We envision that academic researchers and scholars can take up these questions; they can also be taken up by students, parents, administrators, politicians, and policy-makers. Engaging with these issues will continue the work that researchers in writing assessment have developed over the last decade and a half to make sure that content validity has equal consideration with inter-rater and test reliability. Taking up these questions will also show how many large scale assessment--particularly those developed by PARCC and Smarter Balance to support the Common Core State Standards--have continued to privilege forms of reliability. A focus on content, or even use validity, would leave the field trapped in ways that neither researchers nor the public want. Intervening in discussions about writing assessment where we argue about the role of validity in relation to reliability still leaves only two terms in the equation--validity and reliability. The focus on fairness and the ethics of writing assessment introduced in this Special Issue might help change the fundamental elements of the equation. That is a broad and ambitious task. But given the stakes associated with writing assessments in 2016, it is a necessary undertaking. We are delighted with the articles that follow because they take up

that task with such verve.[\[1\]](#)

As ever, we wish to thank the multitude of supporters of *JWA*. The continued financial support of the Department of English and the College of Arts, Letters, and Social Sciences at the University of Idaho enables us to publish independent scholarship that is free and accessible to all. We are grateful for this support, and know that this freely available access was a primary reason the authors of this special issue sought to publish their work in *JWA*.

We could not publish this excellent scholarship without the dedication and hard work of our extended editorial team: Jessica Nastal-Dema, Prairie State College, Associate Editor; Tialitha Macklin, Washington State University, Assistant Editor; and David Bedsole and Bruce Bowles, Jr., both of Florida State University, co-editors of the *JWA* Reading List.

For this Special Issue on the theory of ethics for writing assessment, we gratefully acknowledge the expertise and work of multiple reviewers:

Kathy Charmaz, Sonoma State University

Chris Gallagher, Northeastern University

Richard Haswell, Texas A & M Corpus Christi

Asao Inoue, University of Washington Tacoma

Patricia Lynne, Framingham University

Dan Melzer, University of California Davis

Robert Mislevy, ETS

Duane Roen, Arizona State University

Victor Villanueva, Washington State University

Finally, we look forward to future research that takes up the myriad of issues laid out in this special issue. We invite you to share your research on the theoretical implications of ethical considerations in writing assessment with us for future issues of the *Journal of Writing Assessment*.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014) *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.
- Broad, B. (2003) *What we really value: Beyond rubrics in teaching and assessing writing*. Logan, Utah: Utah State University Press.
- Broad, B., Adler-Kassner, L., Alford, B., Detweiler, J., Estrem, H., Harrington, S., McBride, M., Stalions, E., Weeden, S. (2009) *Organic writing assessment: Dynamic Criteria Mapping in action*. Logan, Utah: Utah State University Press.
- Cronbach, L.J. (1989). Construct validation after thirty years. In R. E. Linn (Ed.), *Intelligence. Measurement, theory and public policy* (pp. 147-171). Urbana, IL: University of Illinois Press.
- Huot, B. (2002). *(Re)articulating writing assessment for teaching and learning*. Logan, Utah: Utah State University Press.
- Kane, M.T. (2001) Current concerns in validity theory. *Journal of Educational Measurement*, 38(4), 319-342.
- Linn, R. L. (1997) Evaluating the validity of assessments. The consequences of use. *Educational Measurement. Issues and Practice*, 16(2),14-18.

Mehrens, W. A. (1998) Consequences of assessment. What is the evidence? *Education Policy Analysis*, 6(13).

Messick, S. (1989) Validity. In R. Linn (Ed.), *Educational Measurement, Third Edition*. (pp.13-105). New York, NY: American Council on Education.

Moss, P. A. (1998) The role of consequences in validity theory. *Educational Measurement. Issues and Practice*, 6-12.

Popham, W. J. (1997) Consequential validity. Right concern—wrong concept. *Educational Measurement. Issues and Practice*, 16(2) 9-13.

Shepard, L. A. (1997) The centrality of test use and consequences for test validity. *Educational Measurement. Issues and Practice*, 16(2) 5-8.

[1] The editors and contributors talk more about issues relating to ethics in writing assessments in the archived NCTE On-Air Web seminar "[No Test Is Neutral: Writing Assessments, Equity, Ethics, and Social Justice](#)." For an interview with the authors, see "[The Ethics of Writing Assessments: Moving from Exclusion to Opportunity](#)" in *Council Chronicle* (2016), 25(3), 6-9.

Copyright © 2021 - *The Journal of Writing Assessment* - All Rights Reserved.