

Review Essay: Paul B. Diederich? Which Paul B. Diederich?

by Rich Haswell

Robert L. Hampel's 2014 edited collection of pieces by Paul Diederich, most of them unpublished, casts Diederich in a new light. The articles, reports, and memoranda reveal him and his work in writing assessment as deeply progressive, both in the educational and political sense. They call for a re-interpretation of his factoring of reader judgments (1961), his analytical scale for student essays (1966), and his measuring of student growth in writing (1974). The pieces also depict Diederich as an intricate and sometimes conflicted thinker, who always saw school writing performance and measurement in terms of the psychological, social, and ethical. He still has relevance today, especially for writing assessment specialists wrestling with current issues such as the testing slated for the Common Core State Standards.

Hampel, R. (2014) *Paul Diederich and the progressive American high school*. Charlotte, NC: Information Age Publishing.

It is November 18, 2014, and yesterday the Smarter Balanced Assessment Consortium announced the cut scores they will apply in their testing of school students on the Common Core State Standards (Gewertz, 2014). It's only one step by SBAC in the four-year development of their tests for Common Core, but it raises multiple questions for specialists in writing assessment.

In 22 states next year, SBAC predicts that the cut will deem as "non-proficient" somewhere between one third and one half of middle schoolers and eleventh graders in English language arts. What is the long-term impact upon the psychology of students and teachers? What impact will it have on school or traditional language-arts objectives such as fostering self-esteem and ability to work with others? Will it further erode time-honored teaching practices such as free reading and peer critique? If the final categorization of test-takers has only four levels--roughly college ready, proficient, non-proficient with partial knowledge, non-proficient with minimal knowledge--how badly does it gloss over the complex and uneven language-proficiency profiles of individual students? If "critical thinking" is one of Common Core's language arts targets, what part did it play in the essay evaluation by human raters that supposedly justified these cut-off scores, and what part can it possibly play with the machine diagnosis and machine assessment planned by SBAC? (Deane, et al., 2013; Zhang, 2013). Most perplexing, how did the new political controversy that is jeopardizing the very future of Common Core State Assessment affect SBAC's final cut-score settings, considering they are, in the end, arbitrary decisions that would exempt only eleven percent of graduating seniors from remedial courses at some universities?

Of course for students of writing assessment--pro or con or, like most, somewhere in the middle--SBAC's projection of cut scores, indeed all of Common Core's future testing, looks less like a new controversy than a familiar recrudescence. We have seen all of this before. What we haven't seen are these obdurate issues through the eyes of Paul B. Diederich.

Paul Diederich? A name, if remembered at all by the profession, fixed in a kind of distorting amber. But at this late date Diederich (1906-1997) has come onto a bit of good fortune. A selection of his writings have appeared, most of them previously unpublished (Hampel, 2014). What they show are not solutions to today's assessment woes. Diederich always had solutions and never lacked the feistiness to champion them, but they are solutions to contexts long gone. He had a way of approaching writing assessment problems, however, that proves still viable. The contexts and the solutions may be past, but his perspectives are surprisingly current and his way of dealing with assessment issues surprisingly attractive. Always a teacher first, Diederich can still teach us how to be a writing assessment activist.

This collection is not likely to sidetrack today's accountability barons. *Paul Diederich and the Progressive American High School?* In the eyes of the Common Core State Standards, for instance, progressive education is pretty much out of historical range. Diederich's topics center on students, teachers, instructional objectives and conditions, and the connection of education to moral and social life. There are pieces on educational shortcomings, faculty workshops, free reading, social goals, assignments in philosophy, ethical goals, Latin grammar, seat time, counseling, the 10-14 curriculum, class size, and revamping the Educational Testing Service. One essay, originally published in 1969, even argues that the accountability mission should not kill progressive aims and methods in education: "The Sputnik-inspired demands for superior education did not toll the death knell for progressive education" (Hampel, p. 136). Furthermore, the editor, Robert L. Hampel of the University of Delaware, explains that he has excluded Diederich's work in "technical research tests" (p. xvii), even though Diederich spent 27 years in the research division of ETS. So, the book's vision is not tunneled or fixated on assessment. It is the vision of the social commentator, albeit testing expert, attempting to bridge the chasm between education in context and measurement in theory and practice.

Throughout his long career, the gap between schooling and testing Diederich observed often and hated fully. It was an observation that sometimes put him at odds with his colleagues at ETS. Here is one of his anecdotes, of which this book is delightfully full. Diederich, it should be noted, was hired in 1949 by Henry Chauncey, president of the fledgling organization, for his background both in school assessment and in psychological testing (Elliot, p. 201; Hampel, p. 111). Off and on at meetings attended by other ETS personnel versed in psychological measurement, Diederich would note that problems for them to consider could be found in school objectives. "Whenever I make that suggestion," Diederich writes in a 1970 memo to William Turnbull, his new boss at ETS, "the psychologists in the group act as though I had farted" (Hampel, p. 168).

And so it is that writing assessment scholars interested in contextual measurement and its relationship to curricula should find this collection enlightening, even supportive. Whatever Diederich found to write about, and that was a lot, never far from his mind was language performance. He had a major impact on writing assessment for several generations. But we must keep in mind that his professional love was ever high school life. True, before he joined the new-born ETS in 1949, he endured some years at the University of Chicago, where he supervised basic writing teachers; devised exemption tests in language, reading, and composition; and worked on the GED examination for the Armed Forces. But the beginning and the rest of his professional life were focused on secondary education.

Early in my career I decided that I was not good enough to teach in elementary schools; I was just about up to teaching in secondary schools; but I would be damned if I would lower myself to teach in a college. I receded from that lofty position when I accepted employment at the University of Chicago, but I never felt quite at home there. Now I am back where I belong--in secondary schools. *The Lord's Vineyard*. (pp. 112-113)

In the same internal report and in the same colorful, direct, and sometimes acerbic language, Diederich asserts that his lack of interest in post-secondary education stems from the belief that college thinking "about educational policies and procedures is about 50 years behind that of our public secondary schools--although they may be more expert in particular tasks, such as the teaching of Hamlet" (p. 112). Hampel's focus on Diederich's views of high-school teaching, curriculum, and evaluation makes good sense if we are after a balanced view of his work and legacy.

To be frank, in the field of college writing assessment Diederich's legacy could use a bit of rebalancing. It's beginning to tip. His idea to use paraprofessionals, "college-educated housewives," to help evaluate student essays receives attention nowadays only as a historical embarrassment. The field has also looked dourly on his 1961 study (with John French and Sydell Carlton) that statistically rendered down rates and commentary on student essays into the five factors of ideas, mechanics, organization, wording, and flavor. The weighted rubric that he derived from this factoring to measure essay performance, often called simply the Diederich scale and popular with researchers for more than two decades, sees little use today. Also long under the bridge are his techniques for documenting progress in writing over school years in *Measuring Growth in English* (1974). For instance, his local teacher scoring has largely been replaced by standardized and sometimes mechanized scoring, replaced long before Common Core assessment consortiums started their work. All in all, it is understandable that the field has tended to stack up the legacy of Diederich on the pass, conservative, regressive, or benighted side of the beam.

A reading of the pieces in this collection will help right this interpretation. Consider Diederich's paraprofessionals. In a 1959 essay reprinted in this collection, Diederich outlines the motives for his plan to hire "lay readers," motives that could only be called progressive then and would still be understood as such today. His purpose was to make English classes teachable--to reduce class size to no more than 25, to free one day of the week so teachers could attend to individual students in conference ("to catch those who have strayed from the flock--especially those who have strayed ahead," p. 160), to liberate another day of the week for "free reading," to increase the amount of student writing to an essay every two weeks ("no one has yet learned to read or write by attending lectures," p. 157), and to allow class time for peer critique. It is reasonable to call sexist Diederich's choice of unemployed women with BAs for the "technicians" that would allow this pedagogy. But the pedagogy itself was far in advance of its times--and in advance, one is tempted to hazard, of the language-arts classroom envisioned by the Common Core testing.

Or consider the "Diederich scale," his analytic measure for documenting growth in school writing. There are eight criteria, each on a five-point scale, weighted and then summed--wide open to the charge of empirical positivism. The empiricism can't be rationalized away, nor would Diederich want it to be. With students, Diederich strongly supported the use of "genuine scientific data on the status and growth in important aspects of development" (p. 118). But what are the "important aspects" and what kind of positivism should be adduced to measure them? Diederich answers this question in an important 1970 report written for ETS directors, published here for the first time. "Developing Educational Awareness within ETS" was the title he gave this "inferentially long memo" (p. 172). In terms of measuring student growth, he states, ETS of the 1960s was mired in the stage of L. L. Thurstone. When the ETS was at all interested in the "personal and social characteristics" of school students, it valued only those traits that are presumed not to change, as in the Myers-Briggs types or as in what ETS was still calling "aptitude." Diederich wants to apply a different kind of positivism, if it can be termed as such. He is after measurements of personal qualities that change and that form the "supposed outcomes of education." Among other personality traits, he lists initiative, perseverance, honesty, orderliness, good judgment, ability to work with others, pride, and cheerfulness (p. 165). Such qualities manifest in student writing, and the Diederich scale recognizes them with criteria such as "Ideas," "Organization," and "Flavor." As Hampel points out, Diederich was an advocate of broad construct representation in assessment. He would certainly have applauded the *Framework for Success in Postsecondary Writing*, for instance, with its insistence on the experiences of reading, writing, and critical analysis as well as on habits of mind. He probably would have applauded Common Core language arts standards such as weighing the credibility and accuracy of multiple sources or developing an essay organization and style appropriate to the audience.

Which leads us back to Diederich, French, and Carlton's 1961 factoring of reader judgment, where the criteria of Ideas, Organization, and Flavor surfaced. This study has been painted as seriously reductive, by Bob Broad (2003), Maja Wilson (2006)

and others, including myself (1998). For instance, Broad charges that it “set out to simplify and standardize the evaluation of writing (p. 5). In writing assessment studies, its method of factor analysis (ironically, derived from Thurstone) stands as a kind of evil grandfather, heir to a progeny of rubrics that have cabined, cribbed, and confined the rife and exuberant communicative urges of student writers and their readers. But which Diederich conducted this study? Let’s factor in his work as a teacher and evaluation staff member (1932-1940) for the Eight-Year Study of select Ohio schools, a famous experiment to improve pre-collegiate education through local teacher initiatives. In one of the most important of the unpublished reports in this collection (Hampel, pp. 4-38), Diederich reviews the Eight-Year Study two years afterward. He begins by excoriating grades: “Don’t we know enough to say that the whole system of marks in courses is bad, and to hell with it? Needed are “records of growth that really mean something (p. 5). He recommends a student folder containing “evidences of growth that are interpretable and individualized, such as “tests, papers, questionnaires, anecdotes, excuses, complaints, summaries of reading (p. 28). This doesn’t much look like simplification or standardization. He might have approved the Common Core standards, but he would have excoriated Smarter Balanced’s measurement reduction of student language performance to levels of 1, 2, 3, and 4.

The year he moved to ETS Diederich wrote an internal report called “Toward a Comprehensive Evaluation System. For students of assessment, it is the most revealing piece in this collection. “The time has come, he prophesies, not very accurately, “for a major revision in our methods of academic bookkeeping. Marks alone, supplemented by a few tests, are not enough (p. 114). He explains that his contracted purpose in joining ETS in 1949 was to create a better bookkeeping system for public schools. He called his improvement the “Profile Index. It looks much like his Eight-Year Study recommendation with the addition of a detailed, practical plan by which “folders bulging with unsorted and unexamined evidence (p. 125) can be sorted, examined, and then used to improve learning, teaching, and curriculum. The Profile Index is the opposite of a depersonalized rubric, in part because it is so robust. Hampel notes that three years later, Diederich drafted a Profile Index of 97 items in 6 categories (p. 127, fn. 19). Such complexity forces the school counselor to consider each student “as a human being rather than as a consumer of the particular wares that he, as a teacher, has to peddle (p. 118). At the least, the Diederich Profile Index helps us re-interpret the Diederich scale that he devised a decade later. Like the Diederich, French, Carlton 1961 factoring, the 1966 scale was not intended as a rubric for stamping a paper with a unitary score but rather as a “common vocabulary to help teachers resolve differences in evaluation, explicitly in contrast to the “blur of “general impression scoring (Diederich, *Measuring Growth*, p. 55). If the factoring and the scale has since been used by others to mark papers, teachers, or curricula as pass or fail, that application runs counter to Diederich’s thorough-going educational progressivism. I might add that Diederich’s Profile Index, now unearthed by this collection, helps explain why there is a direct historical line from Diederich’s scale to later “profile assessments in ESL instruction that depart from general-impression methods (e.g., Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey, 1981; Hamp-Lyons, 1986).

Paul Diederich and the Progressive American High School suggests that in the field of college writing assessment Diederich has been given something of a bum rap. But that, I hope, is the least of the book’s messages. As I have suggested, for me the greatest impact is to flesh out Diederich as an imposing and complex figure in language evaluation. He was that rare salmagundi of pragmatist, idealist, and inventor. One of Melville’s isolatoes (born and raised Catholic in Kansas!), he resisted many of his era’s educational trends yet seemed always to operate with his finger on the pulse. In the pre-WWII Eight-Year Study, he supported the instructional objective of “critical thinking but saw the instructional means, more teaching of Freud and Marxism, as having “produced altogether too many disorganized, irresponsible, unhappy, ineffective people (Hampel, pp. 32-33). He supported faculty workshops but concocted a list of thirty ways workshops avoid facing the issue, for instance, “Retreat from the problem into endless discussion of ways to study it (p. 41). He assisted I. A. Richards in translating Plato’s Republic into Basic English but always believed that the learning of Latin and Greek would help students develop an ample and savory vocabulary in their native tongue. He wrote general objectives for schools yet felt they should objectify “virtues such as “The ability to find aesthetic delight in ideas (p. 85) and later in his career declared, “I no longer write objectives (p. 77).

When James Conant’s *The American High School Today* came out in 1959, supported and publicized by ETS and a bellwether for the accountability creep US students have bodily felt ever since, Diederich approved of the book’s push for higher academic standards and more vocational courses but lambasted its call for more classroom sit time: “Flesh and blood rebel at the monstrous, inhuman schedule of cramped inactivity listening to voices talking nonsense, endlessly repeating themselves, scolding, prompting, belaboring the obvious and scampering over the obscure (p. 104). He worked for an organization that helped develop and validate the Advanced Placement exams, yet he deemed the “subject approach to assessment “a blind alley (p. 122). He supported Skinnerian learning machines and devised some himself and wrote some self-correcting programmed vocabulary workbooks that earned ETS \$40,000 a year in royalties, but he says by his own observation that when students use computer-assisted programmed instruction on a regular basis, “By Christmas at least the better students are so bored that they could scream (p. 174).

With Diederich there are plenty more of these Whitmanesque self-contradictions: “Do I contradict myself? Very well, I contradict myself, I am large, I contain multitudes. Maybe they sound more like Yeatsian contraries. Hampel wisely warns the reader in his introduction that “The boxes in which historians place individuals like Diederich may be too small to accommodate the scope of their world views (p. xv). Hampel also wisely glosses Diederich’s pieces with introductions and footnotes replete with historical contexts and details. They testify to Hampel’s unrivaled knowledge of American education and to his exemplary search across the

nation for repositories of Diederich's letters, articles, reports, and memoranda.

So the book left me with three wishes. I would like to have seen this volume supplied with a chronology of Diederich's wandering-scholar life. I would like to see a second volume of Diederich's unpublished writings, this time centering on his years with ETS. And concerning the people who have made and are making accountability ventures like Common Core and its testing apparatus a fact of life in our schools, I wish a critical number of them had been more imbued with Diederich's contradictions and contraries. I am speaking of writing assessment experts, not just of funders, entrepreneurs, politicians, educational policy makers, state overseers, and school officers.

Wishes aside, I appreciate this book as it stands for the positive message it sends to a certain kind of writing-assessment specialist. I mean the kind who would like to fight for evaluative approaches to language performance that are not simplistic, decontextualized, abstract, boring, illiberal, or repressive. Such specialists might do well to remember Paul B. Diederich, whose tactics always took into account complexity, context, data, inventiveness, liberality, and love of students and their teachers. His message is stand your ground and speak your mind despite the opposition or the odds. The vexed and ill-mapped space between formal testing and existential teaching, he says, can be crossed. It doesn't matter, finally, whether we are fighting for or against the standards or the testing of the Common Core initiative, or for or against any other writing assessment program, on the books or on the horizon. The progressive education that Diederich tried to keep alive and the progressive politics that struggles for life today are two different but meshed ideals. Neither are dead. And neither are divorced from language assessment.

References

- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Logan, UT: Utah State University Press.
- Council of Writing Program Administrators, National Council of Teachers of English, & the National Writing Project. (2014). *Framework for Success in Postsecondary Writing*. Retrieved from <http://wpacouncil.org/framework>
- Diederich, P. B. (1966). How to measure growth in writing ability. *English Journal*, 55(4), 435-449.
- Diederich, P. B. (1974). *Measuring growth in English*. Urbana, IL: National Council of Teachers of English.
- Diederich, P. B., French, J. W., & Carlton, S. T. (1961). *Factors in judgments of writing ability*. ETS Research Bulletin, No. 61.15. Princeton, NJ: Educational Testing Service.
- Elliot, N. (2014). *Henry Chauncey: An American life*. History of Schools and Schooling, Vol. 54. New York: Peter Lang.
- Gewertz, C. (2014). Cutoff scores set for Common-Core tests. *Education Week*, Nov. 21. Retrieved from <http://www.edweek.org/ew/articles/2014/11/17/13sbac.h34.html>
- Hampel, R. L. (Ed.). (2014). *Paul Diederich and the progressive American high school*. Readings in Educational Thought, Vol. 5. Charlotte, NC: Information Age Publishing.
- Hamp-Lyons, L. (1986). *Testing second language writing in academic settings*. Dissertation. Edinburgh: University of Edinburgh.
- Haswell, R. H. (1998). Rubrics, prototypes, and exemplars: Categorization theory and systems of writing placement. *Assessing Writing*, 5(2), 231-268.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- Wilson, M. (2006). *Rethinking rubrics in writing assessment*. Portsmouth, NH: Heinemann.
- Zhang, M. (2013). Contrasting automated and human scoring of essays. *R & D Connections* (March). Retrieved from http://www.ets.org/Media/Research/pdf/RD_Connections_21.pdf