Assessing Writing 23 (2015) 60–74

Contents lists available at ScienceDirect

Assessing Writing

Toward a validational framework using student course papers from common undergraduate curricular requirements as viable outcomes evidence

Diane Kelly-Riley*

University of Idaho, Moscow, ID 83844, United States

ARTICLE INFO

Article history:
Received 21 June 2014
Received in revised form 1 October 2014
Accepted 3 October 2014

Keywords:
Outcomes
Assessment
Local context
Longitudinal studies of writing
Validity
Accountability

ABSTRACT

Examining gains in undergraduate writing abilities, Haswell (2000) applied a multi-dimension construct of writing to impromptu writing exams composed at the first- and third years. This project replicates Haswell's original study to impromptu writing exams composed at the same points, and extends that methodology to course papers written for common undergraduate curricular contexts—first-year composition, general education requirements, and advanced undergraduate writing in the disciplines requirements—to consider the use of such assessment scores as plausible and appropriate evidence for outcomes assessment purposes within a validational framework (articulated by Kane, 2006, 2013). This study considers the feasibility of reporting such localized assessment information as an alternative to represent progress for undergraduate writing ability, and reports preliminary evidence suggesting positive effects of distributed writing requirements across undergraduate curriculums on student writing performance.

* Correspondence to: Department of English, University of Idaho, 875 Perimeter Drive MS 1102, Moscow, ID 83844-1102, United States. Tel.: +1 208 885 5704.
E-mail address: dianek@uidaho.edu

http://dx.doi.org/10.1016/j.asw.2014.10.001
1075-2935/

1. Introduction

In *Academically Adrift: Limited Learning on College Campuses*, Arum and Roksa (2011) observe the "dearth of systematic longitudinal research. . .on the state of undergraduate learning in higher education," (p. 2), and use the results of their study to controversially assert that undergraduates attending colleges and universities in the United States make minimal gains in critical thinking, analytic abilities and writing in their first two years of school. The authors criticize higher education's drift away from the education of undergraduates for the allure of institutional prestige and external funding, and allege that "organizational inertia, the assumption that students are meeting the academic goals espoused in mission statements, and a lack of external pressure to demonstrate learning have all contributed to a failure systematically to measure and evaluate students' gains in higher education" (p. 17).

Their study "track[s] students through a large and representative sample of higher education institutions with objective measures of their learning as well as of their coursework, social background, and the experience of life on today's college campus" (p. 19) by analyzing standardized test scores from the *Collegiate Learning Assessment* (CLA) administered to students at twenty-four, four-year colleges and universities around the United States. The CLA purports to "assess the quality of undergraduate education by directly measuring student learning outcomes through performance tasks" (CAE). Arum and Roksa report two major findings: "many students are only minimally improving their skills in critical thinking, complex reasoning, and writing during their journeys through higher education" (p. 35), and "students are also likely to leave higher education as unequal, or more so, than whey they entered" (p. 37).

Arum and Roksa's alarming findings were widely published, but scholars decried the study's methodological approaches and the hyperbolic tenor of the findings. Astin (2011), Haswell (2012), and others criticized the methodological shakiness and the cherry-picking interpretive approaches of the data. Haswell observed that the authors neglected to reference extensive scholarship related to student learning implying that the authors selectively tuned out key research to amplify their points, and noted that the authors relied on a phantom measure to assess undergraduate learning that other researchers couldn't access since the CLA is a for-profit testing product. Others criticized the use of the CLA as the centerpiece assessment instrument and the subsequent interpretation of the data (Jaschik, 2013; Possin, 2013). Possin, in particular, observed the convoluted nature of the construct used by the CLA. The interpretation of this type of standardized data across institution types implies that all colleges and universities are created equally, and that student populations can (and should) attain similar and uniform levels of undergraduate performance. Finally, and perhaps most curiously, Arum and Roksa's findings actually documented positive gains in students' performance on the CLA across all sub-groups, a point minimized in their data table and their subsequent analysis (see Arum and Roksa, Table A2.1)[1].

Arum and Roksa employ a common narrative frequently used in accountability arguments. Adler-Kassner and Harrington (2010) assert that "accountability is widely used in discussions about what should be happening in school. . .In these 'reform' efforts, writing is narrowly conceived, sometimes as grammatical correctness or, more recently, as the reproduction of particular interpretations or modes, and that [such arguments are] not an appropriate guiding concepts for assessments designed to improve teaching and learning" (pp. 73–76). Such narratives are easy to employ since alternative research has not been conducted to counteract claims about limitations in undergraduate writing performance. In 1989, White observed that "there is no replicated design in existence for demonstrating that any writing instructional program in fact improves writing" (p. 198). This study takes the first step to fill this long-standing void by replicating research conducted by Haswell (2000) that documented gains between entry- and junior-level impromptu writing samples composed for institutional assessment purposes.

[1] In Table A2.1, Arum and Roksa report multiple levels of data. First, the table reports gains in student learning on the CLA across all areas and all sub-groups (my italics for emphasis). Tests of significance on these comparisons are not reported. In the same table, comparisons are made between a comparison category and subcategories, and tests of significance are reported here as well as how much the subcategories differ from the comparison category. Combining all of this data in one table minimizes the actual gains students made across all areas made between first and second years.