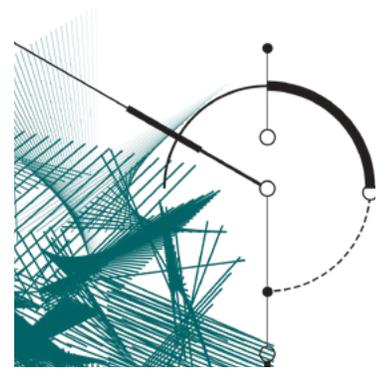


## Abstract

Written communication remains an important learning objective for colleges and universities as more and more students enter the workforce without necessary writing skills and experiences. This importance is increased for public colleges and universities within the state of Texas, as that state has adopted written communication as a core learning objective for its students. The efforts to assess student written communication at one four-year, public university in Texas are highlighted within this study. In particular, differences were examined in student writing performance based upon student race and gender. Using a one-way MANOVA it was determined that no statistically significant relationship existed between student writing performance and either gender or race. It is hoped that the assessment and analysis methodologies presented within this study may serve as models for other researchers seeking to evaluate written communication.



# Examining Differences in Student Writing Proficiency as a Function of Student Race and Gender

## AUTHORS

Jeff Roberts, M.A.  
Sam Houston  
State University

Carroll F. Nardone, Ph.D.  
Sam Houston  
State University

Bill Bridges, Ph.D.  
Sam Houston  
State University

Written communication is widely considered to be an important skill for students to have mastered prior to graduating college and entering the workforce (Allan & Driscoll, 2014; Arum & Roska, 2011; Hart Research Associates, 2013, 2015b; Kelly-Riley, 2015). In a recent survey sponsored by the American Association of Colleges and Universities (AAC&U), 82% of employers indicated that it was important for graduating students to write effectively, and 81% of employers reported that they would be more likely to hire students who took multiple writing-intensive courses in college (Hart Research Associates, 2015b). These results were similar to those from a 2013 employer survey in which 80% of employers noted colleges and universities should place a greater emphasis upon written communication skills (Hart Research Associates, 2013). However, in the face of employer desire for students to have stronger written communication skills only 65% of surveyed students reported that they were well prepared with regard to written communication. Even more troubling, only 27% of the surveyed employers indicated they believed that recent graduates were entering the work force prepared to write effectively (Hart Research Associates, 2015b).

Results like these have led to some higher-education researchers holding negative perceptions regarding student writing proficiency (Arum & Roska, 2011; Secretary of Education's Commission on the Future of Higher Education, 2006). For example, the authors of the Spellings Commission report argued that graduating students lacked fundamental knowledge and skills as they graduated from college (Secretary of Education's Commission on the Future of Higher Education, 2006). These findings served as the basis for the influential and controversial book *Academically Adrift: Limited Learning on College Campuses* (Lederman, 2013; Arum & Roska, 2011). Arum and Roska (2011) further argued that colleges and universities were not doing an adequate job of preparing students with regard to several key skills, which included writing. Using data from the Collegiate Learning Assessment the authors determined that students, in general, made limited gains during

## CORRESPONDENCE

*Email*  
jeff.roberts@shsu.edu

their first two years of college (Arum & Roska, 2011). The scores of minority students lagged behind those of White students, with Black students showing virtually no gain in their scores.

Effective assessment of student writing represents an important tool colleges and universities can use to measure, and ultimately improve, student writing proficiency. However, using third-party, commercial instruments may not provide the meaningful answers institutional leaders are seeking. The measurement of written communication through the evaluation of authentic student artifacts, using locally developed processes, may instead provide institutions with a better perspective of their unique students' writing skills and proficiencies. In turn, these data can help give faculty, staff, and administrators the information they need to identify areas for improvement and to implement curricular and pedagogical changes necessary to increase the writing proficiency of students graduating from their institutions.

## Literature Review

To help place this current study within a broader framework it may be of benefit to the reader to briefly examine some of the existing literature on writing assessment and student written communication proficiencies. Anson (2010), Anson and Lyles (2011), and Behizadeh and Engelhard (2011) share many similarities, focusing on the development of "writing across the curriculum" programs throughout the recent history of higher education. The articles by Anson (2010) and Anson and Lyles (2011) were meta-analyses, examining studies pertaining to writing across the curriculum within 14 relevant journals. The authors of both studies used qualitative research techniques (e.g., citation analysis, content analysis, word count) to conduct further analysis of the articles identified from their searches (Anson, 2010; Anson & Lyles, 2011). Both Anson (2010) and Anson and Lyles (2011) examined roughly 20-year periods within their respective studies (1967 to 1986, Anson, 2010; 1986 to 2006, Anson & Lyles, 2011).

It is interesting to note that Anson and Lyles (2011) could only identify a limited number of articles focusing upon the assessment of student writing. The authors stated, "In the context of burgeoning interest in learning outcomes, assessment, and quality enhancement across all of higher education, the potential for further significant exploration of the uses of writing for assessment in other disciplines remains strong" (p. 15). This paucity of research was also recognized by Behizadeh and Engelhard (2011), who argued that the gap between writing theory and writing assessment was widening. Behizadeh and Engelhard (2011) did observe, though, that a new discipline focused on the assessment of student writing, which combined writing, composition, and measurement scholarship, seemed to be emerging within the literature.

Although the body of research on student writing is limited some studies do exist (Allan and Driscoll, 2014; Barnhisel, Stoddard, & Gorman, 2012; Cargill & Kalikoff, 2007; Desmet, Miller, Griffin, Balthazor, & Cummings, 2008; Faulkner, 2013; Good, Osborn, & Birchfield, 2012; Kelly-Riley, 2015). However, only a few of these studies provide significant discussion of assessment processes and student results (Allan & Driscoll, 2014; Desmet et al., 2008; Faulkner, 2013; Good et al., 2012; Kelly-Riley, 2015). An examination of these few studies show some of the interesting research being conducted around student writing.

Good et al. (2012) described how one university used both a locally developed writing rubric and a third-party, commercial assessment product, the Collegiate Assessment of Academic Proficiencies exam, to assess student writing. The use of multiple measures allowed the researchers to determine how well their locally developed instrument correlated with the Collegiate Assessment of Academic Proficiencies exam, to triangulate their assessment results, and to identify areas for improvement (Good et al., 2012).

Allan and Driscoll (2014) examined student written communication at Oakland University, a doctoral-research institution in Detroit, Michigan with roughly 16,000 undergraduate and 3,500 graduate students, scoring student writing artifacts from lower-level English courses with a rubric. They were then able to identify relative points of strength and weakness in student performance, gain perspectives regarding student perceptions of their own abilities, and provide faculty development opportunities (Allan & Driscoll, 2014). Finally,

**The measurement of written communication through the evaluation of authentic student artifacts, using locally developed processes, may instead provide institutions with a better perspective of their unique students' writing skills and proficiencies.**

Allan and Driscoll (2014) concluded that written reflections could be used alongside other in-class assignments to triangulate assessment and provide a better picture of student learning. Similarly, Desmet et al. (2008) examined students taking freshman composition courses at the University of Georgia. In particular, the authors looked to determine whether “revision improve(d) the quality of writing products” (p. 22). Their study showed that students did improve in a pre-to-post assessment.

Faulkner (2013) also conducted a university-level study of student writing, examining students at Cedarville University. Faulkner strongly advocated for both greater writing instruction and remediation across the curriculum and for implementing Writing in the Disciplines or Writing Across the Curriculum programs, arguing that one-semester remedial English programs cannot meaningfully improve student writing. Alarming, the results of Faulkner’s study demonstrated that student writing scores actually went down from the freshman to senior years at that particular university (2013).

Finally, Kelly-Riley (2015) represents one of the more interesting studies examining student writing. Like the others, Kelly-Riley (2015) examined student writing; however, Kelly-Riley did so using a validation framework, attempting to validate the findings from a previous study of student writing (Hasswell, 2000). Kelly-Riley (2015) examined work from 30 students, from multiple points across their academic careers. Eight different domains related to student writing success were examined using a holistic rubric. The author determined that students made statistical gains across multiple domains and showed statistical improvement over time (Kelly-Riley, 2015).

What is currently missing from this literature are studies examining student writing as a function of race and gender. Race and gender can both represent at-risk factors in higher education (Gray, 2013). The influence of race (Aud, Fox, KewalRamani, 2010; Corona et al., 2017; Harper, 2012; Kim, 2011; Lucas & Paret, 2005; Strayhorn, 2010) and gender (Corona et al., 2017; Kim, 2011; Strayhorn, 2010; Voyer & Voyer, 2014) upon student success is prevalent within educational literature. However, these studies typically focus on general student success. Research examining student writing proficiency as a function of race or gender are almost nonexistent within the literature. In fact, higher-education institutions, in general, are not examining their data in this way. Acting on behalf of the AAC&U, Hart Research Associates conducted a survey with which they determined that 70% of institutional leaders reported tracking learning outcomes achievement data; however, only 16% of the responding institutions reported disaggregating data by race (Hart Research Associates, 2015a). Ultimately, student success in higher education is increasingly becoming a social justice issue (Gray, 2013); therefore, it is key for higher-education professionals to better understand how these factors may influence student writing performance.

### **Statement of the Problem**

In the face of the challenges and concerns posed by government agencies, researchers (Arum & Roska, 2011; Secretary of Education’s Commission on the Future of Higher Education, 2006), business leaders (Hart Research Associates, 2015b), and institutions must find ways to accurately assess, and help improve, student writing. These issues are particularly important for public colleges and universities in Texas, for the Texas Higher Education Coordinating Board has identified student written communication as one of the core learning objectives adopted for all public institutions within the state (Texas Higher Education Coordinating Board, 2015). However, all institutions seeking to assess student written communication, whether by state mandate or faculty choice, face similar challenges. The importance of assessing student written communication through the lens of race and gender is magnified given the importance of equity in higher education (Gray, 2013; Montenegro & Jankowski, 2017). However, there remains a significant gap in the literature in this area that needs to be addressed.

### **Purpose and Significance of the Study**

Given the demonstrated importance of written communication for undergraduate students, and the criticisms of colleges and universities to adequately prepare students to write effectively, faculty and staff need to develop ways to assess student written communication.

**What is currently missing from this literature are studies examining student writing as a function of race and gender.**

The authors of this study seek to highlight the efforts of one four-year public university in southeast Texas to use a locally developed writing rubric to effectively assess the writing proficiency of students as they approached graduation. In particular, the authors attempted to determine what, if any, differences might exist in student writing scores as a function of student race and student gender. Not only does this study join the growing body of literature related to assessing student writing (Allan and Driscoll, 2014; Barnhisel et al., 2012; Cargill & Kalikoff, 2007; Desmet et al., 2008; Faulkner, 2013; Good et al., 2012; Kelly-Riley, 2015) this study also seeks to examine the important issue of equity in student achievement that many are raising in higher education (Gray, 2013; Hart Research Associates 2015a; Montenegro & Jankowski, 2017). Finally, the assessment methodologies and analysis techniques presented within this study may also serve as an example to other institutions seeking to evaluate student writing.

## Research Questions

The following research questions were addressed in this study: (a) What was the difference in the student performance on an end-of-experience student writing assessment as a function of student race (i.e., White, Black, Hispanic, Other)?; and (b) What was the difference in the student performance on an end-of-experience student writing assessment as a function of student gender (i.e., male, female)?

## Method

### Participants

Student writing artifacts were selected from 4000-level, writing-enhanced courses at a four-year, public university in southeast Texas during the spring 2013 semester. A stratified, random sampling process was used in order to select authentic student writing artifacts. Several steps were taken to identify this sample pool and collect the writing artifacts. As the purpose of the original study was to examine the writing proficiencies of upper-division students, all students not classified as being juniors or seniors were excluded from the initial sample pool. This potential sample pool was then divided into separate stratum, by academic college. Students were randomly selected from within these stratum, with the total number of students selected being based upon the percentage of junior- and senior-level majors within each college for the spring 2013 semester. To identify the total number of artifacts selected from each course within the various stratum the total number of declared majors within the sample population for each college was divided by the total number of courses within that stratum. This methodology resulted in a sample pool that was representative of both the size and diversity of the studied university.

The instructors of record for each of the 203 writing-enhanced courses within the sample were then emailed requesting the selected student artifacts. All received artifacts were redacted of student and faculty identifying information in preparation for scoring, and were assigned a unique tracking code. Ultimately, 430 student artifacts from 153 writing-enhanced courses were received, of which 395 were chosen for scoring. A total of 27 submitted artifacts were unusable for the writing assessment (e.g., short-answer tests, papers written in a foreign language, illegible handwritten student work). Additionally, eight artifacts were used as anchor papers to norm faculty raters and were not included within the data for analysis.

### Instrumentation

To obtain the writing scores used for data analysis in this research article the sampled student writing artifacts were scored using a locally developed writing rubric. Kuh et al. (2015) argued that “rubrics encourage the use of authentic student work for assessment” (p. 39). The rubric was separated into four different domains of student writing (i.e., Ideas/Critical Thinking/Synthesis, Style, Organization, Conventions). All artifacts were scored independently by two raters, with each rater scoring the artifact for each domain using a four-point scale. These individual domain scores were then averaged to provide an overall score for each student artifact.

**The authors of this study seek to highlight the efforts of one four-year public university in southeast Texas to use a locally developed writing rubric to effectively assess the writing proficiency of students as they approached graduation. In particular, the authors attempted to determine what, if any, differences might exist in student writing scores as a function of student race and student gender.**

## Score Reliability

With any rubric-based assessments one important measure of reliability is the consistence of the scores (Banta & Palomba, 2015; Millett, Payne, Dwyer, Stickler, & Alexiou, 2008). Therefore, several steps were taken to ensure the consistency of the scoring process. An interdisciplinary group of faculty raters evaluated student artifacts over a two-day period using a locally developed rubric. At the beginning of the scoring session the group of raters were normed to the rubric using anchor papers. The entire group of raters scored identical papers and were then led through a discussion of their scores by a facilitator in order to bring everyone into agreement regarding how to appropriately apply the rubric. Twelve of these faculty members served as either a first or second rater for each artifact, with the first rater's score not being known by the second rater. When a discrepancy of two or more points was present between the average total scores for the first two raters one of two different faculty members served as a third rater. The score from the third rater was then used in place of the score that was furthest out of agreement.

Intraclass correlation coefficients (ICCs) were calculated to determine the level of interrater agreement for each of the four writing domains (i.e., Ideas/Critical Thinking/Synthesis, Style, Organization, and Conventions), the total overall score, and the overall average (Fleiss, 2003; Shrout & Fleiss, 1979). Because every rater did not evaluate every student writing artifact, a one-way random ICC was calculated. According to Cicchetti (1994), ICC agreement values below .40 demonstrate poor agreement, values from .40–.59 demonstrate fair agreement, values from .60–.74 demonstrate good agreement, and values above .75 demonstrate excellent agreement. The ICC agreement values for three of the four writing domains (i.e., Ideas/Critical Thinking/Synthesis, Style, Organization) were above a .60, indicating good agreement, while the ICC agreement value for conventions was .58, indicating fair agreement. The ICC agreement values for the total overall score and the overall average were both .80, indicating excellent agreement for the total scores (see Table 1 for a full breakdown of the ICC agreement values for this study).

**The first purpose was to examine what differences might exist in student writing scores as a function of student race and student gender.**

Table 1

*Breakdown of ICC Agreement by Category Area*

Domain Area	Intraclass Correlation for Average Measures
Ideas/Critical Thinking/Synthesis	.69
Style	.65
Organization	.64
Conventions	.58
Overall Artifact Average	.80

## Results

Prior to conducting statistical procedures to address differences in student performance on an end-of-experience writing assessment as a function of student race and of student gender the normality of the dependent variables were first ascertained. The standardized skewness coefficients (i.e., the skewness value divided by its standard error) and the standardized kurtosis coefficients (i.e., the kurtosis value divided by its standard error) were all within the boundaries of normality,  $\pm 3$  (Onwuegbuzie & Daniel, 2001) for both research questions. However, the assumption for the Box's Test of Equality of Covariance was violated for both research questions. Finally, the Levene's Test of Equality of Error Variances revealed that the assumptions were met for both research questions. As the majority of the assumptions were met for both research questions, the use of a parametric, one-way Multivariate Analysis of Variance (MANOVA) was justified for this study (Field, 2009). The MANOVA procedures did not reveal a statistically significant difference

in student writing performance as a function of race (i.e., White, Black, Hispanic, Other), Wilks'  $\Lambda = .97$ ,  $p = .56$ , or as a function of gender (i.e., male, female), Wilks'  $\Lambda = .99$ ,  $p = .65$  (see Table 2 for the descriptive statistics for these analyses).

Table 2

*Descriptive Statistics for Student Writing Scores by Student Race and Gender*

Student Demographic Characteristic	Ideas, Critical Thinking, Synthesis, Style, Organization, Conventions								Overall Student Average	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Race										
White ( $n = 259$ )	2.75	0.74	2.72	0.69	2.67	0.72	2.64	0.75	2.69	0.64
Black ( $n = 51$ )	2.61	0.81	2.56	0.72	2.55	0.66	2.41	0.73	2.53	0.65
Hispanic ( $n = 56$ )	2.57	0.66	2.62	0.60	2.53	0.64	2.46	0.70	2.54	0.56
Other ( $n = 28$ )	2.43	0.68	2.46	0.71	2.55	0.61	2.45	0.55	2.48	0.53
Gender										
Male ( $n = 143$ )	2.64	0.76	2.63	0.72	2.60	0.68	2.52	0.77	2.59	0.65
Female ( $n = 251$ )	2.71	0.72	2.69	0.67	2.64	0.71	2.60	0.71	2.66	0.61

## Discussion

**At first glance, the lack of statistically significant results within this study would seem disheartening. However, from an institutional perspective, these results are very important as they highlight the actual performance of that institution's students with regards to written communication.**

The purpose of this study was twofold. The first purpose was to examine what differences might exist in student writing scores as a function of student race and student gender. In doing so, this study joins a growing body of literature on the assessment of student writing (Allan and Driscoll, 2014; Barnhisel, et al., 2012; Cargill & Kalikoff, 2007; Desmet, et al., 2008; Faulkner, 2013; Good, et al., 2012; Kelly-Riley, 2015). Additionally, this study represents an example of how an institution is disaggregating student performance data (Hart Research Associates, 2015a), and is helping to answer questions regarding the equity of student learning (Montenegro & Jankowski, 2017). Finally, this study provides a model to other institutions for assessing student writing performance and analyzing those results. At first glance, the lack of statistically significant results within this study would seem disheartening. However, from an institutional perspective, these results are very important as they highlight the actual performance of that institution's students with regards to written communication.

If a college or university is doing an adequate job of preparing its students to write effectively it would be natural to expect that all students, regardless of race or gender, would perform equitably upon an authentic writing assessment. Therefore, the lack of statistically significant results observed within this study could be interpreted by decision makers from that university to mean that they are preparing students equally well with regard to written communication. That said, equity does not necessarily mean quality. More information is needed to determine whether the level of student performance observed within this study was sufficient for end-of-experience students. A possible explanation for these results may also be that weaker students, regardless of race or gender, did not persist to the junior- or senior-year to be measured, thus limiting the differences observed by race or gender. It should be noted that while statistically significant differences in student scores by race and gender were not observed, White students scored higher than all other races and females scored higher than males across all four rubric domains. Further study is needed to better understand and interpret these results.

More work is also needed in order to determine whether the findings of this study are the result of some outside factors. The lack of statistical significance in the results of this study may not be representative of actual student performance, but instead may reflect error within the assessment process itself. For example, the locally developed rubric used within this study may not be sensitive enough to pick up the differences between the various student groups. There may also be flaws with the rubric itself which may be impacting the collected results. Finally, the sample size used for this study may also not have been sufficient to identify any differences by race or by gender.

Several steps can be taken in order to address these possible concerns. The first logical course of action would be to increase the size of the sample being used for analysis. This would allow the researchers to determine whether the results were the result of an insufficient sample size or were actually representative of student performance. It might also allow for separate statistical analysis on the racial groups included within the Other category (e.g., Asian or Pacific Islander, American Indian, International). Further replication of this study is needed to replicate and validate the results identified here (cf. Kelly-Riley, 2015).

Furthermore, as this initial study used a one-way MANOVA no attempt was made to examine the interactions between race and gender upon student written communication proficiencies. Follow-up studies are needed, with larger samples, to better understand how student performance can be affected by student membership within multiple groups. Additional variables, like socio-economic status and first-generation status, could also be included within such an analysis to better understand the nuances of student writing.

Efforts could also be made to help further validity of the rubric used to score student writing artifacts. For example, the same rubric could be used to also score writing artifacts from beginning students, the scores from which could be compared to those of end-of-experience students in order to determine whether the rubric was sensitive enough to pick up potential differences between the two groups. Also, cross-institutional scoring and comparison could offer opportunities for rubric validation. Already scored, redacted, and coded student artifacts could be traded between, and scored by, peer institutions in order to determine how student artifacts from one institution scored using the instrument from the other. Scores could then be compared using statistical analysis in order to determine how well the scores from the two rubrics correlated. This would both provide evidence for the validity of both institutions' assessment instruments, and would possibly give insight into how an institution's students were doing in comparison to peers.

As a parting warning, readers are cautioned to not overgeneralize the findings presented within this study. The examined population was limited to junior- and senior-level students attending one public, four-year Texas university, in 2013. The results from the analysis may therefore not be generalizable beyond the time, setting, and population involved within this study. Finally, although several steps were taken to try to ensure the validity and reliability of the methodologies used in this study, faculty, staff, and administrators may experience different results if they attempt to replicate the methodologies at their own institutions.

## Conclusion

The data presented within this study represent only the first effort by one institution to evaluate the written communication proficiencies of its students, and the specific assessment methodologies highlighted here are not the only ways to evaluate student writing. Despite the promises of some groups to provide the magic bullet for evaluating written communication (e.g., the CLA+; Council for Aid to Education, 2015), it is impossible for any single test, measure, or rubric to provide all the information needed by institutions to improve student writing. Institutional improvement does not occur over night but instead takes the time and intentionality of faculty, staff, and administrators.

Student writing remains of great importance (Allan & Driscoll, 2014; Arum & Roska, 2011; Hart Research Associates, 2013, 2015b; Kelly-Riley, 2015), and those within higher education need to better prepare students to write effectively. In order to make the changes

**Therefore, the lack of statistically significant results observed within this study could be interpreted by decision makers from that university to mean that they are preparing students equally well with regard to written communication. That said, equity does not necessarily mean quality.**

that are necessary to improve student written communication, faculty, staff, and administrators must have the necessary data to make those changes. This study provides an overview of one institution's attempts to use authentic assessments to gather this needed data. In doing so, readers may be inspired to engage in their own local assessments of student writing.

## References

- Allan, E. G., & Driscoll, D. L. (2014). The three-fold benefit of reflective writing: Improving program assessment, student learning, and faculty professional development. *Assessing Writing, 21*, 37–55. doi:10.1016/j.asw.2014.03.001
- Aud, S., Fox, M., & KewalRamani, A. (2010). *Status and trends in the education of racial and ethnic groups* (NCES 2010–015). Washington, DC: U.S. Department of Education, National Center for Education Statistics, U.S. Government Printing Office. Retrieved from <http://nces.ed.gov/pubs2010/2010015.pdf>
- Anson, C. M. (2010). The intradisciplinary influence of composition and WAC, 1967-1986. *The WAC Journal, 21*, 5–19.
- Anson, C. M., & Lyles, K. (2011). The intradisciplinary influence of composition and WAC, part two: 1986-2006. *The WAC Journal, 22*, 7–19.
- Arum, R., & Roska, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.
- Banta, T. W., & Palomba, C. A. (2015). *Assessment essentials: Planning, implementing, and improving assessment in higher education* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Barnhisel, G., Stoddard, E., & Gorman, J. (2012). *Incorporating process-based writing pedagogy into first-year learning communities: Strategies and outcomes*. *The Journal of General Education, 61*, 461–487. doi:10.1353/jge.2012.0041
- Behizadeh, N., & Engelhard, G., Jr. (2011). Historical view of the influences of measurement and writing theories on the practices of writing assessment in the United States. *Writing Assessment, 16*, 190–211. doi:10.1353/jge.2012.0041
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Cargill, K., & Kalikoff, B. (2007). Linked psychology and writing courses across the curriculum. *The Journal of General Education, 56*, 83–92. doi:10.1353/jge.2007.0017
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284–290. doi:10.1037/1040-3590.6.4.284
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field settings*. Chicago, IL: Rand McNally.
- Corona, R., Rodriguez, V. M, McDonald, S. E., Velazquez, E., Rodriguez, A., & Fuentes, V. E. (2017). Associations between cultural stressors, cultural values, and Latina/o college students' mental health. *Journal of Youth & Adolescence, 46*(1), 63–77. doi:10.1007/s10964-016-0600-5
- Council for Aid to Education. (2015). *CLA+ overview*. Retrieved from <http://cae.org/participating-institutions/cla-overview/>
- Desmet, C., Miller, D. C., Griffin, J., Balthazor, R., & Cummings, R. E. (2008). Reflection, revision, and assessment in first-year composition eportfolios. *The Journal of General Education, 1*, 15–30.
- Faulkner, M. (2013). Remediating remediation: From basic writing to writing across the curriculum. *The CEA Forum, 42*(2), 45–60.
- Fleis, J. L. (2003). *Statistical methods for rates and proportions* (3rd ed.). New York, NY: Wiley. doi:10.1002/0471445428
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.
- Gray, S. S. (2013). Framing “at risk” students: struggles at the boundaries of access to higher education. *Children & Youth Services Review, 35*, 1245–1251. doi:10.1016/j.childyouth.2013.04.011
- Good, J. M., Osborne, K., & Birchfield, K. (2012). Placing data in the hands of discipline-specific decision makers: Campus-wide writing program assessment. *Assessing Writing, 17*, 140-149. doi:10.1016/j.asw.2012.02.003
- Harper, S. R. (2012). *Black male student success in higher education: A report from the national Black male college achievement study*. Philadelphia, PA: University of Pennsylvania, Center for the Study of Race and Equity in Education.

- Hart Research Associates. (2013). *It takes more than a major: Employer priorities for college learning and student success*. Washington, DC: Association of American Colleges and Universities. Retrieved from <http://www.aacu.org/leap/presidentstrust/compact/2013SurveySummary>
- Hart Research Associates. (2015a). *Bringing equity and quality learning together: Institutional priorities for tracking and advancing underserved students' success. Key findings from a survey and in-depth interviews among administrators at AAC&U member institutions*. Washington, DC: Association of American Colleges and Universities. Retrieved from <http://www.aacu.org/publications/bringing-equity-and-quality-learning-together>
- Hart Research Associates. (2015b). *Falling short? College learning and career success*. Washington, DC: Association of American Colleges and Universities. Retrieved from <http://www.aacu.org/leap/public-opinion-research/2015-survey-falling-short>
- Hasswell, R. H. (2000). Documenting improvement in college writing: A longitudinal approach. *Written Communication*, 17, 307–352. doi:10.1177/0741088300017003001
- Johnson, B., & Christensen, L. (2012). *Educational research: Quantitative, qualitative, and mixed approaches* (4th ed.). Los Angeles, CA: Sage.
- Kelly-Riley, D. (2015). Toward a validation framework using student course papers from common undergraduate curricular requirements as viable outcomes evidence. *Assessing Writing*, 23, 60–74. doi: 10.1016/j.asw.2014.10.001
- Kim, Y. M. (2011). *Minorities in higher education: Twenty-fourth Status Report* (2011 Supplement). Washington, DC: American Council on Education. Retrieved from [http://diversity.ucsc.edu/resources/images/ace\\_report.pdf](http://diversity.ucsc.edu/resources/images/ace_report.pdf)
- Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T., Hutchings, P., Kinzie, J. (2015). *Using evidence of student learning to improve higher education*. San Francisco, CA: Jossey-Bass.
- Leaderman, D. (2013, May). Less academically adrift? *Inside Higher Education*. Retrieved from <https://www.insidehighered.com/news/2013/05/20/studies-challenge-findings-academically-adrift>
- Lucas, S. R., & Paret, M. (2005). Law, race, and education in the United States. *The Annual Review of Law and Social Science*, 1, 203–231. doi:10.1146/annurev.ls.1.110105.100001
- Millett, C. M., Payne, D. G., Dwyer, C. A., Stickler, L. M., & Alexiou, J. J. (2008). *A culture of evidence: An evidence-centered approach to accountability for student learning outcomes*. Princeton, NJ: Educational Testing Service. Retrieved from: [https://www.ets.org/Media/Education\\_Topics/pdf/COEIII\\_report.pdf](https://www.ets.org/Media/Education_Topics/pdf/COEIII_report.pdf)
- Montenegro, E., & Jankowski, N. A. (2017). *Equity and assessment: Moving towards culturally responsive assessment*. (Occasional Paper #29). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
- Onwuegbuzie, A. J. (2003). Expanding the framework of internal and external validity in quantitative research. *Research in Schools*, 10(1), 71–89.
- Onwuegbuzie, A. J., & Daniel, L. G. (2001). Uses and misuses of the correlation coefficient. *Research in the Schools*, 9(1), 73–90.
- Secretary of Education's Commission on the Future of Higher Education. (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, DC: U.S. Department of Education. Retrieved from <http://www.ed.gov/>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychology Bulletin*, 86, 420–428. doi:10.1037/0033-2909.86.2.420
- Smith, M. L., & Glass, G. V. (1987). *Research and evaluation in education and the social sciences*. Englewood Cliffs, NJ: Prentice Hall.
- Strayhorn, T. L. (2010). When race and gender collide: Social and cultural capital's influence on the academic achievement of African American and Latino males. *The Review of Higher Education*, 33(3), 307–332. doi:10.1353/rhe.0.0147
- Texas Higher Education Coordinating Board. (2015). *Elements of the Texas Core Curriculum*. Retrieved from <http://www.theccb.state.tx.us/index.cfm?objectid=427FDE26-AF5D-F1A1-E6FDB62091E2A507>
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 114, 1174–1204. doi:10.1037/a0036620