First-Year Composition and Transfer: A Quantitative Study

James D. Williams, Rhetoric & Linguistics, Soka University

Minami Hattori, Psychology, University of Notre Dame

Contact: jwilliams@soka.edu

Abstract

The present study investigated the effect of writing pedagogy on transfer by examining the effect of pedagogical orientation (WAC/WID or 'traditional') on content-area grades. Participants were 1,052 undergraduates from 17 schools throughout the United States. Hypothesis was that the WAC/WID orientation would lead to higher transfer levels as measured by participants' higher content-area performance. Composition grades were collected in year one; content-area grades where collected in year two. Propensity scores were calculated to stratify the groups and minimize selection bias of writing-class assignment, thereby allowing quasi-causal inference. An ANOVA was performed on the resulting 2-by-5 stratified data. Results indicated that students who completed the WAC/WID composition classes received significantly higher content grades than those in the 'traditional' writing classes. The results confirmed the hypothesis.

Keywords: transfer, academic performance, composition, pedagogy

First-Year Composition and Transfer: A Quantitative Study

Concerns about the value and intellectual rigor of first-year composition (FYC) are long standing (e.g., Bamberg, 1997; Connors, 1995; Skeffington, 2012), and various studies have reported that FYC does not help students become better writers (e.g., Arum & Roksa, 2011; Fleming, 2002; Zorn, 2013). Although conceptions of 'better writer' vary considerably, a consensus nevertheless has formed that FYC should, at a minimum, provide students with transferable writing skills that help them succeed in content-area courses. As Wardle (2007) noted, the FYC requirement throughout US colleges and universities 'suggests that administrators, policy makers, parents, and students expect the course to prepare students for the writing they will do later—in the university and even beyond it. Implicit in these expectations is the assumption that FYC should and will provide students with knowledge and skills that can transfer to writing tasks in other courses and contexts' (p. 65). The problem, as Wardle pointed out, is that there is little evidence that FYC succeeds in meeting this expectation. Some reports suggest that the failure is systemic, rooted in the discipline's resistance to defining FYC as a service course (e.g., Mahoney, 2011), in which case any lack of transfer could be attributed to FYC pedagogy.

Yancey, Robertson, and Taczak (2015) argued that, like 'better writers,' definitions and concepts of 'transfer' are highly contentious. In the context of FYC, they challenged the commonplace idea that transfer can be understood as the 'application of skills from one situation to another' (p. 7) on the ground that it does not address the rhetorical factors inherent in the adaptive nature of writing transfer—which must meet the varying conditions and contexts of the diverse content-area courses that comprise the undergraduate curriculum.

The psychological literature on transfer, however, is perhaps better characterized as varied rather than contentious, for the more influential perspectives have considerable overlap. Thorndike and Woodworth's (1901) proposal that transfer consists of cognitive improvement in one area that leads to improvement in a related area has much in common with Gagné's (1965) identification of what he referred to as 'lateral' transfer—'a kind of generalization that spreads over a broad set of situations at roughly the same level of complexity' (p. 231) and 'vertical' transfer—skill or knowledge that contributes to a superordinate skill. With Gagné, transfer theory and research began emphasizing context, as is evident in Perkins and Salomon's (1992) 'near' and 'far' transfer. Near transfer occurs between similar contexts, whereas far transfer occurs between contexts that appear remote from each other. Salomon and Perkins (1988) also proposed that transfer occurs via two psychological mechanisms, 'low road' and 'high road.' Low-road transfer occurs when a learned skill set is applied in a context that is perceptually similar to the contextual domain of learning. High-road transfer, on the other hand, occurs when a learned skill set is applied in a context significantly different from the contextual domain of learning. Nelms and Dively (2007) focused on high-road mechanisms when they noted that 'Transfer . . . involves

the application of knowledge acquired in one situation or context to a different situation or context' (p. 215).

Efforts to study the transfer of FYC skills have produced, at best, mixed results. A small number of transfer studies exist, but these have reported conflicting findings. Moreover, they tend to be confounded by questionable methodologies and small sample sizes. Wardle's (2007) study of transfer, for example, included only 7 students, whom she worked with over the course of their undergraduate years. Wardle collected samples of the participants' writing, interviewed them individually and as a group, and had them complete a survey related to their perceptions of whether FYC helped them in content-area courses. She reported that transfer did occur in such areas as how to conduct research and the ability to talk about writing in general. As Wardle indicated, however, observer-expectancy bias and self-enhancing bias were limitations of the study. Moreover, the participants were self-selected, and the lack of randomization was yet another limitation.

Working with nonnative English-speaking students, James (2010) also reported that writing instruction resulted in transfer, but the study was based on students' self-reports, again raising the issue of self-enhancing bias. Zarei and Rahimi (2014), following James, used a combination of self-reports and participant interviews and found that transfer occurred when instruction focused on writing skills that were aligned with specific contexts. Their study, however, involved only 13 participants, which presents a challenge to generalizability.

Nelms and Dively (2007) noted that observing and measuring transfer is difficult because it occurs over time. They concluded that, as a result, transfer of knowledge and proficiency from FYC to content-area courses 'remains largely unexplored' (p. 215). Perkins and Salomon's (1992) theory suggests that effecting the kind of complex transfer involved in applying FYC writing proficiency in content-area courses requires composition pedagogies that involve what they termed 'bridging.' A bridging pedagogy would provide specific connections between FYC writing assignments and activities and those required in the undergraduate curriculum.

Activity Theory (AT) explains the mechanisms involved in bridging by viewing transfer as distributed learning governed by complex interactions with others using cultural tools, particularly language (Kuuti, 1996; Russell, 1995). A particular activity becomes purposeful in a given context when the application of appropriate tools involves the internalization and externalization of cognitive processes (Nadi, 1996). Thus, individual learning involves cognitive changes resulting from repeated social interactions with other people and the cultural tools that characterize the contexts, or activity systems, of those interactions (Hutchins, 1995). The cognitive changes are not limited to internalization of information and/or skills through decontextualized repetition but rather are the result of expanding involvement with the conventions and tools that characterize specific contexts and their activity systems.

Brent (2011) suggested that, although AT provides insight into how FYC pedagogy may be shaped to facilitate transfer to the 'larger world of activity systems in general' (p. 400), the functional characteristics of the activity systems in composition classes often differ significantly from the activity systems in content-area courses, thus complicating transfer of knowledge and proficiency. For example, an FYC pedagogy that focuses on developing abilities in self-expression, theme-based writing, or literary analysis would be unlikely to transfer to success in a history class that requires a paper analyzing the effect of the Industrial Revolution on England's population growth.

When writing across the curriculum (WAC) and writing in the disciplines (WID) pedagogies emerged in the late 1970s and early 1980s, they aimed, in part, to actualize bridging by contextualizing writing, giving students opportunities to practice the rhetorical conventions, standards of proof, and ways of knowing that characterize academic writing in various disciplines (Neff & Whithaus, 2009; Williams, 2014). WAC/WID programs have become common nationwide over the last 40 years. Although these programs differ in various ways, a frequently expressed goal is that WAC/WID will emphasize writing for academic purposes and will introduce students to the activity systems in content-area courses. As originally conceived, WAC/WID programs recognized that a writer must have knowledge not only of the tools appropriate to a given context but also knowledge of the various conventions, rhetorical and otherwise, associated with using those tools in specific contexts. On this account, these programs have

the potential to serve as an alternative to FYC pedagogies that emphasize general writing skills for general audiences, and they may offer a principled approach to assessing FYC transfer.

The bulk of existing research on WAC/WID pedagogy has focused on efficacy, however, not transfer, perhaps because administrators and content-area faculty commonly expect WAC/WID programs to demonstrate their effectiveness in helping students write proficiently. Although a large body of assessment literature exists on writing efficacy (e.g., Adler-Kassner & O'Neill, 2010; Elliot, 2005; Huot & O'Neill, 2009, O'Neill, Moore, & Huot, 2009), it suffers from many of the same problems evident in FYC transfer research. The tendency has been to reframe the research question in terms of 'stakeholder satisfaction' and non-quantifiable constructs (e.g., Ochsner & Fowler, 2004; Yancey & Huot, 1997). As Kuh et al. (2006) noted, 'satisfaction' is merely one factor in assessing academic performance—indeed, a factor of relatively low significance compared to acquisition of knowledge and competencies that are applicable in specific contexts.

Rose and Theilheimer (2002), for example, sought to assess the efficacy of their school's WAC/WID program by soliciting 'two students' points of view about WAC, based on data gleaned from interviews with them' (p. 18). They based their methodology on 'the WAC/WID assessment literature [, which] repeatedly recommends that evaluators turn to stakeholders as they set their research agenda' (p. 18). The results showed that the students, as well as the faculty involved, were 'satisfied' with the WAC/WID program. The authors also noted that 'through experience, or by design, the instructors we interviewed crafted writing assignments that drew initially on students' personal experiences' (p. 27), which cannot be aligned with either academic writing in general or with writing in the specific contexts of content-area courses because personal-experience writing is private, whereas academic writing is public.

In recognition that stakeholder satisfaction is not a valid measure of performance, several studies have sought to examine learning outcomes. Robinson and Burton (2009), for example, asked 256 undergraduates enrolled in upper-division writing-intensive courses to complete a questionnaire that aimed to engage the participants in self-reflection and goal setting related to their writing practices and their writing-intensive courses. The researchers found that a majority reported setting goals and engaging in reflection on various components of their writing experiences. The lack of any controls for self-reporting bias compromised the study, however, as did the failure to consider whether students in the writing-intensive courses performed better than cohorts in sections that were not designated as writing intensive.

Seeking to avoid such problems, some studies have included analyses of student writing in addition to surveys and self-reports to assess FYC efficacy and transfer. Herman et al., (2011), for example, selected 20 WAC/WID courses (designated as WRITD, or writing in the disciplines) at Gustavus Adolphus College[1] and used a multifaceted approach involving student and faculty surveys, review of faculty syllabi and assignments, and faculty evaluation of 113 first-draft/final draft student papers using a 3-point rubric ('high, ' 'medium, ' 'low '). The researchers found that 73% of the participants felt that their writing had improved as a result of the WAC/WID course. Faculty readers scored 85.7% of the student papers as demonstrating adequacy or mastery (60.7% and 25%, respectively) of academic conventions.

Although the Herman et al. (2011) results are suggestive, they reveal very little about either efficacy or transfer owing to methodological shortcomings that compromised the findings. The 3-point rubric used for assessing student writing, for example, was not based on specific writing criteria and therefore relied on highly subjective assessments. In addition, the methods section provided no discussion of how the readers were socialized to the rubric, raising questions of assessment reliability.

In sum, the available assessment literature tends to exhibit a number of methodological problems that prevent the results from providing much insight into the question of either transfer or FYC efficacy. On this account, it seems reasonable to conclude, as Olds, Leydens, and Miller (1999) did two decades ago,

---

[1] Gustavus Adolphus College requires students to take 2 writing intensive (WRITI) courses and 1 WRITD course for graduation. The distinction between the two courses on the school website is unclear.

that FYC assessment research tends to be anecdotal and/or idiosyncratic, largely owing to reliance on 'stakeholders' satisfaction' (p. 120) measured via questionnaire surveys.

A potential alternative to assessing transfer through questionnaires and interviews is to use students' academic achievement in content-area classes that require writing, using FYC pedagogy as the independent variable. Activity Theory predicts that FYC instruction that provides genre-specific training targeting specific audiences will result in higher overall grades in content-area classes that ask students to write papers than FYC instruction that provides general skills for general audiences, which for the purposes of this study will be referred to as 'traditional' FYC. The rationale is that genre-specific writing will be more closely aligned with the activity systems of content-area classes. If the theory is sound and transfer occurs, it can be measured in any given content-area class that requires writing: Students who experience FYC pedagogy that focuses on genre-specific writing for specific audiences will generally receive higher grades on content-area papers than students who experience FYC pedagogy that focuses on general writing skills for general audiences. All other things being equal, the higher grades on papers will affect overall grades in the content-area courses. In this respect, such an alternative to assessing transfer is congruent with York, Gibson, and Rankin's (2015) meta-analysis of academic performance, which indicated that 'achievement should be a direct result of attaining learning objectives and acquiring desired skills and competencies' (p. 6).

Cumulative grades are widely used as legitimate measures of performance throughout education (York, Gibson, & Rankin, 2015). Teachers commonly use multiple grade points to assess students' individual course performance. For decades, high school GPA has been touted as an accurate indicator of college success, which a recent large-scale study appeared to confirm (Hiss & Franks, 2014). Nevertheless, a potential obstacle to using grades as a dependent variable is that they are merely proxies of performance rather than absolute measures. As Allen (2005) noted, individual grades have questionable validity owing to factors such as variations in how and why teachers assign them, subjectivity, and, as Arum and Roksa (2011) reported, the influence of professors' yearning for high student evaluations (Stark & Freishtat, 2014). Using grades to assess entire programs is also problematic because the courses within those programs may differ significantly.

Using grades to assess FYC transfer, however, may be feasible owing to the fact that within any given program there are certain shared characteristics, irrespective of content, associated with effective writing within a given context or activity system, such as applying discipline-specific rhetorical conventions, use of evidence, and standards of proof. Moreover, in a sufficiently large dataset, subjective assignment would be random error, or noise in the data, and not a systematic error affecting the hypothesis.

Like all academic performance measures, grades may be imperfect, but various studies have reported their predictive validity. Geiser (2004), for example, found the squared correlation between high school GPA and college first-year GPA to be .15, compared to only .13 for SAT scores. Likewise, Zekarias, Aba-Milki, and Mikre (2015) reported that grade-12 GPA, along with university entrance exam score, accounted for 43.5% (R^2 = 0. 435, F = 129. 2, $p < 0.05$) of academic achievement. Zahner, Ramsaran, and Steedle (2014), in a sample of 4,500 sophomores and juniors, found that high school GPA was the best predictor of college GPA. Schuler, Funke, and Baron-Boldt (1990), in a study of more than 29,000 students, found that 'the mean corrected validity of final school grades for the prediction of university examinations is ϱ = 0.456; for the prediction of vocational training success it is ϱ= 0.408, matching the validities of the best psychological predictors in personnel selection' (p. 89).

The Present Study
WAC/WID programs tend to focus on genre-specific writing for specific audiences, whereas 'traditional' FYC pedagogies emphasize general writing skills for general audiences. Activity Theory therefore suggests that FYC transfer can be investigated in a principled way by examining writing performance in content-area classes that require genre-specific writing for specific audiences. On this account, it was hypothesized that students who completed a FYC class with a WAC/WID orientation would receive higher grades in content-area courses, thus demonstrating a higher level of transfer, than those who completed a 'traditional' writing course.

The present study was reviewed and approved by the principle investigator's university IRB. Participants consisted of two undergraduate groups: (1) those who had completed a FYC course with a WAC/WID orientation and (2) those who had completed a FYC course with a 'traditional' orientation. FYC data were collected nationwide from colleges and universities over a 2-year period. At the end of Year 1, data were collected from participants who had completed their FYC requirement. At the end of Year 2, data were collected from a range of participating content-area faculty for undergraduate courses—such as American History, 18th Century British Literature, Introduction to Biology, Women's Studies, Ethnic Studies, International Relations, Introduction to Psychology, Introduction to Sociology, Anthropology, American Literature, etc.

Methods
School Selection
Data were collected over two years (2013–2014) at three levels:

1. university (doctoral-granting institutions)
2. colleges (MA-granting institutions)
3. community colleges (AA granting institutions)

School selection began with a review of websites to identify schools that have in place a WAC/WID program. The review showed that: (1) some schools' offered a 'traditional' FYC program exclusively; (2) some offered WAC/WID exclusively; and (3) some offered a two-tiered program consisting mostly of 'traditional' writing courses but with a small number of WAC/WID courses. (In these two-tiered programs, students had the ability to choose which class they would take.) The result was an initial pool of 425 colleges and universities distributed throughout in the West, Midwest, South, and East. The writing program administrator (WPA) at each school was then solicited via email to participate in the study. Sixty-eight WPAs responded (16%).

The next step involved providing the respondents with the following criteria for the study:

- WAC/WID writing courses had to focus on genre-specific writing for more than one specific audience (e.g., humanities, social/behavioral sciences, life science).
- WPAs had to collect and provide student grade rosters for a minimum of 4 classes from the previous year (year 1).
- WPAs had to be willing to contact undergraduate content-area faculty and solicit their cooperation in the study. They then had to provide cooperating content-area faculty with the names of FYC students who had completed the FYC requirement during the previous year. Cooperating content-area faculty had to match names to those students who had completed a content-area course and provide the corresponding course grade (year 2). The content-area faculty were not to be informed of the pedagogical orientation of students' FYC so as to ensure they were blind to the condition and thus not subject to bias.
- Cooperating content-area faculty had to require at least one paper of 5 pages or more as part of the course requirements. Finally, they had to provide the titles of their courses and departmental affiliation.

Nineteen percent of responding programs limited their WAC/WID effort to writing-intensive courses that involved asking content-area faculty to increase the amount of writing students produce in these courses. Because these programs did not offer any WAC/WID-oriented composition classes, they were excluded from the study.

Participants
A total of 17 schools met all the study criteria, resulting in 1052 student participants (451 males, 601 females) after adjustments. The participants were undergraduates nationwide who had completed the FYC requirement and who subsequently had enrolled in content-area courses. The FYC courses and the content-area courses were not linked, and the content-area courses were not designated as writing intensive. To increase the generalizability of the findings, data were collected at three higher-education levels in all four regions of the United States: university (doctoral-granting institutions), colleges (BA-

and MA-granting institutions, and community colleges (AA-granting institutions) located in the West, Midwest, South, and East.

Composition Class Grades. The participating WPAs provided students' grades in first-year composition classes. Letter grades were converted to numeric values, using the following nominal conversion: A = 4, A- = 3.75, B+ = 3.5, B = 3, B- = 2.75, C+ = 2.5, C = 2, C- = 1.75, D+ = 1.5, D = 1, D- = .75, F = 0.

Content-Area Class Grades. The WPAs also provided copies of the students' content-area grades, which they had received from cooperating content-area faculty. The content-area grades came from multiple sections of 23 different undergraduate classes, such as history, sociology, psychology, American literature, philosophy, environmental studies, and women's studies. Letter grades for the courses were again converted to numeric values, using the same conversion as for the composition class grades.

Writing Class Type. Writing classes were coded as follows: 1 = WAC/WID, 0 = Traditional. As a check on type designation, the WPAs were asked to submit copies of syllabi for the selected sections. They also were asked to obtain and submit copies of syllabi from the participating content-area faculty to check that the courses met the specification for writing in the study criteria.

Owing to the logistics involved, WPAs could not collect and send copies for all classes. They did, however, provide what they considered to be representative samples for composition as well as content-area classes. A total of 123 syllabi were received (75 for writing classes; 48 for content-area classes). These were then reviewed to validate study parameters. The sample syllabi were also reviewed to check the distinction between WAC/WID assignments (i.e., gaining specific knowledge and skills that are applicable in specific contexts) and 'traditional' assignments (gaining general knowledge and skills that is applicable in general contexts). One hundred percent of the Traditional syllabi corresponded with the designation. Fourteen percent of the WAC/WID syllabi, however, did not indicate a WAC/WID orientation but rather a Traditional orientation. These sections were dropped from the study. The content-area syllabi showed that 100% met the study criteria.

Coding the Data

The two sets of grades were matched by student names, and the data were entered into an SPSS data file. Upon entry, names were replaced with numeric codes to de-identify the student participants so as to ensure and maintain confidentiality, and the original data were purged in keeping with US HHS provision 45 C.F.R. §164.514(a)(b). The codes also identified the regional source of the data (West = 1, Midwest = 2, South = 3, East = 4) and a nominal section number (1, 2, 3, etc.) for each writing class at each school. Gender identity (based on given names) was retained and entered into the data set (0 = male; 1 = female).

Participants were evenly enrolled in one of the two types of writing-class orientation (WAC/WID = 525; Traditional = 527). In 81 cases, an FYC student subsequently enrolled in more than one participating content-area class and thus had multiple content-area grades. These were dropped from the final data set to simplify the analysis. Missing values (some composition grades and/or content-area grades and some names that did not clearly indicate gender) were also deleted from the data set.

The final data set consisted of the following variables:

- ClassType (WAC, Traditional)
- CompGrade (composition grade)
- ContGrade (content-area grade)
- gender
- SchoolType (community college, college, university)

The data were then analyzed using SPSS, with p set at .05.

Results

Analysis 1

Sample sizes, means, and standard deviations for content-area grades were calculated for the entire data set. The results are shown in Table 1.

The effect of a FYC orientation on transfer cannot be investigated in a non-randomized study owing to the possibility that students chose their writing classes on the basis of pedagogical orientation rather than, for example, the class schedule or the individual teacher. In such a case, the direct estimation of treatment effect can be biased (Rubin, 1991). Covariates that are considered to be related to writing class selection (IV) and content-area grades (DV) should be statistically controlled. Analysis of covariance, however, is not appropriate because the categorical variables, such as gender, are likely to violate the assumption of linearity (Tabachnick & Fidell, 2007). Thus, propensity score analysis is more appropriate (Rosenbaum & Rubin, 1983, 1984; Thoemmes & Kim, 2011).

A binary logistic regression using three covariates (gender, geographical region, and writing class grades) was conducted to compute a propensity score for each student. Students were then divided into five quintiles based on these scores. For example, the fifth quintile included students with conditional probability of choosing to be enrolled in a WAC/WID class that was higher than the 80th percentile. A quintile stratification is a common choice because five strata remove approximately 90% of the selection bias due to measured covariates when estimating a treatment effect (Cochran, 1968).

Respondents in the first group had propensity scores ranging from 0 to .2999343 (N = 213; 20.2%); those in the second group had propensity scores ranging from .299934300001 to .4991136 (N = 261; 24.8%); those in the third group had propensity scores ranging from .499113600001 to .5335585 (N = 125; 11.9%); those in the fourth group had propensity scores ranging from .533558500001 to .7060559 (N = 256; 24.3%); and those in the fifth group had propensity scores ranging from .706055900001 to 1 (N = 197; 18.7%). Including categorical covariates resulted in quintiles with unequal sample sizes, but this does not pose a statistical problem (Thoemmes & Kim, 2011).

Propensity score analysis is valid when substantial overlap exists on propensity scores across groups (Imai, King, & Stuart, 2008). The overlap of range, mean, and standard deviation of propensity scores was therefore evaluated for both groups (WAC/WID vs. Traditional) across quintiles. The results showed that the range of propensity scores between groups overlapped substantially between groups across all quintiles (see Table 2), indicating no need for trimming any data.

Statistical Assumptions

Propensity score stratification has two underlying statistical assumptions (Thoemmes & Kim, 2011). The first is that the difference in propensity scores between the groups is approximately one standard deviation or less. Analysis showed that the WAC/WID group had higher propensity scores (M = .5986, SD = .1854) than the Traditional group (M = .3999, SD = .2062, Cohen's d =1.01). The second assumption is homogeneity of variance across groups. In this study, the ratio of variance was 1.24, which is between the acceptable range of .5 and 2. Thus, the two underlying assumptions were sufficiently met: The quintiles effectively stratified the two groups. These results therefore indicate an absence of selection bias, allowing us to conclude that the groups could be treated as randomized samples.

Adjusted Analysis of Variance

A 2- (WAC/WID vs. Traditional) by-5 (quintiles based on propensity score) two-way analysis of variance (ANOVA) was then conducted to estimate the WAC/WID main effect, controlling for stratification and its interaction. That is, quintiles were entered as a fixed factor. The WAC/WID main effect was significant, F(1,1042) = 76.24, p <. 001, partial η^2 = .24. Students in WAC/WID classes received significantly higher content-area grades than those in Traditional writing classes across the quintiles that matched on gender, geographical region, and writing grades. The trend is shown in Figure 1. This result indicated that the study hypothesis was correct and that transfer occurred.

To further test the hypothesis, the unadjusted effect of WAC/WID on content-area grades was tested for comparison using an independent sample t-test. When not matched, those in the WAC/WID classes received significantly higher content-area grades (M = 3.23, SD = .56) than those in the Traditional

writing classes (*M* = 2.81, *SD* = .58), *t*(1051) = 11.70, *p* < .001. These results further confirmed the study hypothesis; moreover, the mean difference of .42 on a 4-point scale seems substantively important. The *t*-test also indicated a significant difference in the grades students received in their writing classes (WAC: *n* = 525, *M* = 2.7248, *SD* = .66894; Traditional: *n* = 527, *M* = 3.3036, *SD* = .64033, *F* = .458, *t* = -14.337, *p* < .001). Students in the WAC/WID writing classes received significantly lower grades than their Traditional counterparts. This finding suggests two possible explanations: Either the WAC/WID courses overall were more challenging than the Traditional or the WAC/WID students overall were less proficient writers to begin with than their counterparts. The latter seems unlikely, given that Traditional classes tend to be very similar to the writing instruction undergraduates receive in high school insofar as such instruction generally focuses on general skills for general audiences, but the issue nevertheless warrants further investigation.

Analysis 2

Analysis 1 did not consider two potential confounding factors: (1) the participants came from different schools, and (2) the data had a nested structure. These factors suggest a possible content-grades dependency that could violate the statistical assumptions of ANOVA. The violation of independence may lead to alpha-inflation, producing false-positive results (Barcikowski, 1981). Therefore, the data were analyzed further using multilevel modeling to consider the nested structure of the data (Bryk & Raudenbush, 1992). Specifically, multilevel models with 1052 students (level-1) nested within 17 schools (level-2) were examined with three sequential models with three variables: writing class type as a predictor (Traditional = 0, WAC/WID = 1), gender as a covariate (male = 0, female = 1), and content-area grades as an outcome variable. A maximum likelihood estimation was used because it allows estimating more parameters than the variant methods (Tabachnick & Fidell, 2007).

Model 1 was a two-level empty model without any predictor. This model served as a baseline for the size of contextual variations in students' content-area grades in all subsequent models. The intra-class correlation (ICC), a proportion of group-level variance in total variance, was .003.

Model 2 was a random intercepts model that allowed the mean content-area grades to vary across schools. Gender was included in the model as a covariate. A scaled chi-square difference test (Satorra & Bentler, 2001) indicated significant improvement of Model 2 over Model 1, $\Delta\chi^2$ (2) = 132.54, *p* < .001. Similarly, a smaller AIC for Model 2 (1809.45) than for Model 1 (1937.99) and a smaller BIC for Model 2 (1834.24) than for Model 1 (1952.86) indicate the superiority of Model 2. That is, mean content-area grades varied across schools for each gender.

The nonstandardized regression weight on writing class type from Model 2 was .41, SE = .04, CI_{95} [.34, .48], *p* < .001. These results indicated that participants of the same gender and at the same school who were enrolled in a WAC/WID class had content-area grades that were .41 *points higher* than their counterparts in a traditional writing class. The results replicated those from Analysis 1 and further confirmed the study hypothesis.

Model 3 was a random intercepts and random slopes model, allowing the slope between the predictors and content-area grades to vary across schools. However, the scaled chi-square difference test, as well as AIC and BIC, indicated no substantive improvement of Model 3 over Model 2. In fact, no random slope was statistically significant at α = .05, indicating that the overall effects of WAC/WID on content-area grades did not vary across schools. Thus, Model 3 was not interpreted. The lack of variation across schools indicated that the effect of WAC/WID orientation, as measured by content-area grades, was significant regardless of where WAC/WID courses were offered.

Discussion

Many variables contribute to students' academic success as undergraduates, but various studies have emphasized the importance of writing ability (e.g., Alliance for Excellent Education, 2007; Arum & Roksa, 2011; Defazio, Jones, Tennant, & Hook, 2010; Levine & Dean, 2012). Although courses and instruction vary across the undergraduate curriculum, the grades students receive on the papers they write in their content-area courses contribute to their grades in those courses (Defazio, et al., 2010; Zhu, 2004). The findings of this study indicate that FYC transfer does occur, but they also offer an alternative to the interviews and satisfaction surveys that have characterized similar investigations for years.

The results do not indicate that the Traditional group failed to experience FYC transfer. They may have. However, the effect size in Analysis 1 (Choen's d = 1.01) is large, leading us to conclude that the degree of transfer for the WAC/WID students was significantly greater than for those in the Traditional group and that they were significantly better prepared for the challenges of situated, evidence-based writing in specific contexts. If this interpretation is correct, it raises questions about 'traditional' pedagogy's efficacy in promoting undergraduate success.

In addition, the results indicate that the benefits of a pedagogy that focuses on specific skills for specific contexts are independent of school type and location. These covariables did not moderate the results. The WAC/WID participants performed better than the Traditional participants at all schools and all locations. This finding is congruent with Activity Theory and should be encouraging to schools that are considering whether to invest in developing and implementing a WAC/WID program. It also should serve WPAs whose university administrators require evidence that WAC/WID programs are worth the considerable resources necessary to develop and maintain them.

Our interpretation of the results is that the WAC/WID pedagogy provided skills and knowledge that were transferable because they were relevant to the writing demands in participants' subsequent content-area classes. Those demands activated the cognitive representations developed during participants' WAC/WID classes owing to rhetorical and structural similarities between the tasks. Our examination of syllabi showed that WAC/WID assignments for environmental studies papers, for example, shared many characteristics with writing assignments in environmental studies classes. Whereas one might ask students to analyze the effects of population growth on the environment of a selected area, the other might ask students to analyze the effect of acidification on the planet's oceans. The topics implied the audience (insiders) and associated rhetorical features, and the assignments further implied that the papers were to be analytical, not argumentative, not journalistic, and not expressive. Finally, both assignments specified use of Chicago formatting and documentation conventions.

On this account, we conclude—congruent with cognitive theory—that robust transfer occurs when FYC instruction increases the richness of the cognitive interconnections that will lead to transfer. More specifically, effective, transferable instruction will focus on providing students writing assignments that are aligned with the kinds of writing assigned throughout the undergraduate curriculum. We also conclude that robust transfer occurs not at the far or high level but rather at the near or low level, where the required tasks share conceptual and rhetorical domains. If this conclusion is correct, then what Yancey, Robertson, and Taczak (2015) identified as the typical concern of compositionists—'high-road transfer . . . when two occasions are 'paradoxical'' (p. 16)—further explains (along with questionable methodology) why so much of the FYC transfer research has produced largely insubstantial results.

With regard to generalizability, two factors are important for transfer research. First, the sample size is several orders of magnitude larger than previous studies. Second, both the FYC and content-area faculty were blind, an important factor in controlling for bias. We therefore conclude that the results are highly generalizable.

Among the programs that met the study criteria, the finding that the WAC/WID students received lower grades in their composition classes than their counterparts in the traditional classes is interesting. The possibility that students in these classes were less proficient than their Traditional counterparts seems, on its face, difficult to support, given the diversity of school types and locations. It therefore seems more reasonable to conclude that the WAC/WID courses were more challenging, regardless of region and school type. The FYC sample syllabi supported this conclusion. Those WAC/WID syllabi that included writing assignments required students to produce evidenced-based papers dealing with content-area issues. The Traditional syllabi, on the other hand, commonly asked students to write expressive responses to readings and social issues—assignments that, on the surface, at least—appeared to replicate the types of assignments that are frequently provided at the secondary level (see Williams, 2014). That is, the WAC/WID syllabi focused on situated writing for specific audiences, whereas the Traditional syllabi focused on general writing for general audiences.

Worth noting is that a significant percentage of WAC/WID programs and courses reviewed for this study were not aligned with FYC transfer and therefore did not meet the study criteria. Among the responding WPAs, 31% indicated that their WAC/WID programs serve in merely a support capacity for writing-intensive efforts; they do not actually offer any WAC/WID writing courses. Instead, the WPA and staff provided consultation to content-area faculty on such topics as grading criteria and effective assignments while offering FYC courses with a 'traditional' orientation. Another 26% implemented a curriculum that asked students to write expressive papers on topics related journalistically to content areas rather than to write papers similar to those that students would actually write in those areas. In both cases, the pedagogical focus appeared to be on analyzing texts rather than on writing instruction.

Limitations

Although the sample size in this study is significantly greater than all previous investigations of FYC transfer, a larger sample was desirable. Sample size was limited by three factors: (1) the lack of sufficient investigator resources; (2) the large-scale absence of bridging writing assignments in WAC/WID programs; and (3) the significant time commitment required of WPAs and cooperating content-area faculty. Given the lack of quantitative investigations of FYC transfer—and thus its importance to the field of composition studies as well as to undergraduate education—our inability to obtain a higher response rate was disappointing. Moreover, the number of WAC/WID programs that did not focus on bridging assignments raises concerns about the current status of WAC/WID and its viability as a genuine pedagogical orientation. Future research would need to anticipate these challenges and plan accordingly.

Another limitation is that propensity score analysis works best when the data set for the groups under investigation includes a high number of covariables. Given the nature of the study and the data-collection procedures, it was not possible to obtain numerous covariables. Moreover, designing a follow-up study to do so would increase the demands on participating faculty. We would point out, however, that this limitation is mitigated by the fact that propensity score analysis was used here to assess randomization, but this factor is nevertheless a concern.

Finally, only a few of the content-area syllabi (<10%) indicated the weight given to student papers when determining course grades. In our experience, larger classes often assign less weight to writing assignments than smaller classes, and knowing the class sizes and the weight given to student papers when determining course grades would have provided additional data for interpreting the results, so any future studies should take this issue into account. Even so, our view is that the weight given to writing assignments in content-area classes is a random variable. Therefore, the fact that some content area grades may have been influenced by a low writing weight assignment suggests a higher level of transfer than the aggregate data indicate.

Conclusion

The results of this study provide evidence of measurable FYC transfer among students when their writing courses focused on meeting specific criteria for specific contexts, thereby bridging activity systems. This finding is important not only because of the existing research gap with regard to FYC transfer but also because it serves to confirm the value of writing for academic purposes. Replicating the study will potentially have significant implications for writing pedagogy and academic performance.

The results also raise questions regarding existing FYC curricula and learning outcomes. The relatively small percentage of WAC/WID programs that focused on rhetorical conventions, ways of knowing, and standards of proof across the curriculum suggests that in many instances there is little difference between existing WAC/WID courses and 'traditional' ones. If future research confirms this finding, it would call into doubt the substantial investment in resources necessary to develop and maintain WAC/WID programs.

References

Adler-Kassner, L., & O'Neill, P. (2010). *Reframing writing assessment to improve teaching and learning*. Logan, UT: Utah State University Press.

Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 78,* 218–223.

Alliance for Excellent Education. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement.* Washington, DC: Author.

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses.* Chicago, IL: University of Chicago Press.

Bamberg, B. (1997). Alternative models of first-year composition: Possibilities and problems. *WPA, 21,* 7–18.

Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics, 6,* 267–285.

Brent, D. (2011). Transfer, Transformation, and Rhetorical Knowledge: Insights from Transfer Theory. *Journal of Business and Technical Communication, 25,* 396–420.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage.

Cochran, W. G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics, 24*, 295–313.

Connors, R. (1996). The new abolitionism: Toward a historical background. In J. Petraglia (Ed.), *Reconceiving writing, rethinking writing instruction* (pp. 3–16). Mahwah, NJ: Lawrence Erlbaum Associates.

Defazio, J., Jones, J., Tennant, F., & Hook, S. A. (2010). Academic literacy: the importance and impact of writing across the curriculum—a case study. *Journal of the Scholarship of Teaching and Learning,* 10, 34–47.

Elliott, N. (2005). *On a scale: A social history of writing assessment in America*. New York: Peter Lang.

Fleming, D. (2002). The end of composition-rhetoric. In J. D. Williams (Ed.), *Visions and revisions: Continuity and change in rhetoric and composition* (pp. 109–130). Carbondale, IL: Southern Illinois University Press.

Gagné, R. M. (1965). *The conditions of learning*. New York: Holt, Reinhart, and Winston.

Herman, J., Goodwin, D., Dobler, C., Fremo, R., Banks, A., Rao, S., & Tunheim, K. (2011). *Writing Across the Curriculum (WAC) Assessment Report*. Adolphus College. Retrieved from https://gustavus.edu/wac/assessment

Hiss, W. C., & Franks, V. W. (2014). Defining promise: Optional standardized testing polices in American college and university admissions. National Association for College Admission Counseling. Retrieved from http://www.nacacnet.org/research/research-data/nacac-research/Documents/DefiningPromise.pdf

Huot, B., & O'Neill, P. (2009). *Assessing writing: A critical sourcebook*. Boston: Bedford.

Hutchins, E. (1995). *Cognition in the wild.* Boston: MIT Press.

Imai, K., King, G., & Stuart, E. (2008). Misunderstandings among experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society, Series A, 171*, 481–502.

James, M. A. (2010). An investigation of learning transfer in English for general academic purposes writing instruction. *Journal of Second Language Writing, 19,* 183–206.

Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). What matters in student success: A review of the literature. Commissioned report for the National Symposium on Postsecondary Student Success: Spearheading a Dialog on Student Success. Washington, DC: National Postsecondary Education Cooperative.

Kuutti, K. (1996). Activity theory as a potential framework for human-computer interaction research. In B. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (17–44). Cambridge, MA: MIT press.

Levine, A., & Dean, D. R. (2012). *Generation on a tightrope*. San Francisco: Jossey-Base.

Mahoney, K. (2011). You can't get there from here: Higher education, labor activism, and challenges of neoliberal globalization. In S. Kahn and J. H. Lee (Eds.), *Activism and rhetoric: Theories and contexts for political engagement* (pp. 147–158) New York: Routledge.

Nardi, B.A. (1996). *Context and consciousness: Activity theory and human-computer interaction*. Cambridge, MA: MIT Press.

Neff, J. M., & Whithaus, C. (2009). Writing across distances and disciplines: *Research and pedagogy in distributed learning*. New York: Erlbaum.

Nelms, G., & Dively, R. L. (2007). Perceived roadblocks to transferring knowledge from first-year composition to writing-intensive major courses: A pilot study. *Writing Program Administration, 31,* 214–240.

O'Neill, P., Moore, C., & Huot, B. (2009). *A guide to college writing assessment.* Logan, UT: Utah State University Press.

Ochsner, R., & Fowler, J. (2004). Playing the devil's advocate: Evaluating the literature of the WAC/WID movement. *Review of Educational Research, 74,* 117–140.

Olds, B. M., Leydens, J. A., & Miller, R. L. (2009). A flexible model for assessing WAC/WID programs. *Language and Learning Across the Disciplines, 3,* 123–129.

Perkins, D. N., & Salomon, G. (1992). The science and art of transfer. In A. L. Costa, J. Bellanca, & R. Fogarty (Eds.), *If minds matter: A forward to the future* (Vol. 1) (pp. 201–210). Palatine, IL: Skylight Publishing.

Robinson, T. A., & Burton, V. T. (2009). The writer's personal profile: Student self-assessment and goal of setting at start of term. *Across the Disciplines: A Journal of Language, Learning, and Academic Writing, 6,* np. Retrieved from http://wac.colostate.edu/atd/assessment/robinson_burton.cfm

Rose, L., & Theilheimer, R. (2004). You write what you know: Writing, learning, and student construction of knowledge. *The WAC/WID Journal, 13,* 17–29.

Rosenbaum, P., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects**.** *Biometrica, 70,* 41–55.

Rosenbaum, P., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*, 516–524.

Rubin, R. B. (1991). Practical implications of modes of statistical inference for causal effects and the critical role of the assignment mechanism. *Biometrics, 47*, 1213–1234.

Russell, D. (1995). Activity theory and its implications for writing instruction. In J. Perraglia (Ed.), *Reconceiving writing, rethinking writing instruction* (pp. 51–77). Mahwah, NJ: Lawrence Erlbaum Associates.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66,* 507–514.

Skeffington J. K. (2012). Enhancing transfer from first-year composition: A pedagogy of shorter essays. *Journal of Teaching Writing, 27,* 27–45.

Salomon, G., & Perkins, D. N. (1988). Teaching for transfer. *Educational Leadership, 46,* 22–32.

Schuler, H., Funke, U., & Baron-Boldt, J. (1990). *Applied Psychology,* 39, 89–103.

Stark, P. B., & Freishtat, R. (2014). An evaluation of course evaluations. *ScienceOpen,* 1–26. DOI: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.)*. Boston, MA: Allyn & Bacon/Pearson Education.

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research, 46*, 90–118.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, 8, 247–261 DOI: 10.1037/h0074898.

Wardle, E. (2007). Understanding 'transfer' from FYC: Preliminary results of a longitudinal study. *WPA, 31,* 65–85.

Williams, J. D. (2014). *Preparing to teach writing: Research, theory, and practice,* 4th edition. New York: Routledge.

Yancey, K., Robertson, L., & Taczak, K. (2015). *Writing across contexts: Transfer, composition, and sites of writing*. Logan, UT: Utah State University Press.

Yancey, K., & Huot, B. (Eds.) (1997). *Assessing writing across the curriculum: Diverse approaches and practices.* New York, Ablex.

York, T., Gibson, C., & Rankin, S. (2015). Defining and measuring academic success. Practical Assessment Research & Evaluation, 20, 1–20.

Zahner, D., Ramsaran, L. M., & Steedle, J. T. (2014). Comparing alternatives in the prediction of college success. Council for Aid to Education. Retrieved from http://cae.org/images/ uploads/pdf/

Zarei, G. R., & Rahimi, A. (2014). Learning transfer in English for general academic purposes writing. DOI: 10.1177/2158244013518925.

Zekarias, Z., Aba-Milki, N., & Mikre, Fisseha. (2015). Predictors of academic achievement for first year students: The case of Wolaita-Sodo University, Ethiopia. *European Scientific Journal, 11,* 160–173.

Zhu, W. (2004). Faculty views on the importance of writing, the nature of academic writing, and reaching and responding to writing in the disciplines. *Journal of Second Language Writing, 13,* 29–48.

Zorn, J. (2013). English composition as fraud and failure. *National Academic of Sciences Report.* Retrieved from www.nas.org/articles/english_composition_as_fraud_and_failure.

Table 1

Sample sizes, means and standard deviations of content grades of students in 17 schools.

School	Region	Type	n	Mean	SD
1	West	PhD Granting Institution	43	3.18	.73
2	West	PhD Granting Institution	41	2.91	.71
3	West	Community College	44	3.10	.60
4	West	BA/MA Granting Institution	51	3.20	.65
5	Midwest	BA/MA Granting Institution	91	2.96	.65
6	Midwest	PhD Granting Institution	49	2.96	.58
7	Midwest	Community College	42	3.03	.65
8	Midwest	BA/MA Granting Institution	45	2.93	.53
9	Midwest	BA/MA Granting Institution	91	3.04	.59
10	Midwest	BA/MA Granting Institution	38	2.88	.70
11	South	BA/MA Granting Institution	125	2.94	.66
12	South	PhD Granting Institution	82	3.09	.58
13	East	BA/MA Granting Institution	48	3.08	.40
14	East	BA/MA Granting Institution	126	3.05	.57
15	East	BA/MA Granting Institution	47	3.06	.56
16	East	PhD Granting Institution	40	3.09	.55
17	East	Community College	49	2.87	.49

Note. SD = standard deviation. Content grades are on 4-point scale.

Table 2

Sample size and minimum, maximum, and mean propensity score for all groups across quintiles.

Quintile	Group	N	Min	Max	M	SD
1	Traditional	191	.1600	.2999	.1981	.0399
	WAC/WID	22	.1645	.2999	.2052	.0478
2	Traditional	148	.3070	.4991	.3698	.0752
	WAC/WID	113	.3070	.4991	.4262	.0822
3	Traditional	51	.5081	.5253	.5184	.0068
	WAC/WID	74	.5081	.5253	.5175	.0074
4	Traditional	86	.5357	.7061	.5737	.0578
	WAC/WID	170	.5357	.7061	.5946	.0665
5	Traditional	51	.7134	.9691	.8313	.0668
	WAC/WID	146	.7134	.9670	.8370	.0792

Note. N = sample size. Min = minimum; Max = maximum; *M* = mean; *SD* = standard deviation, of propensity score.

Figure 1

Estimated content grades for WAC/WID or Traditional groups across quintiles.

Estimated Marginal Means of Content Grades
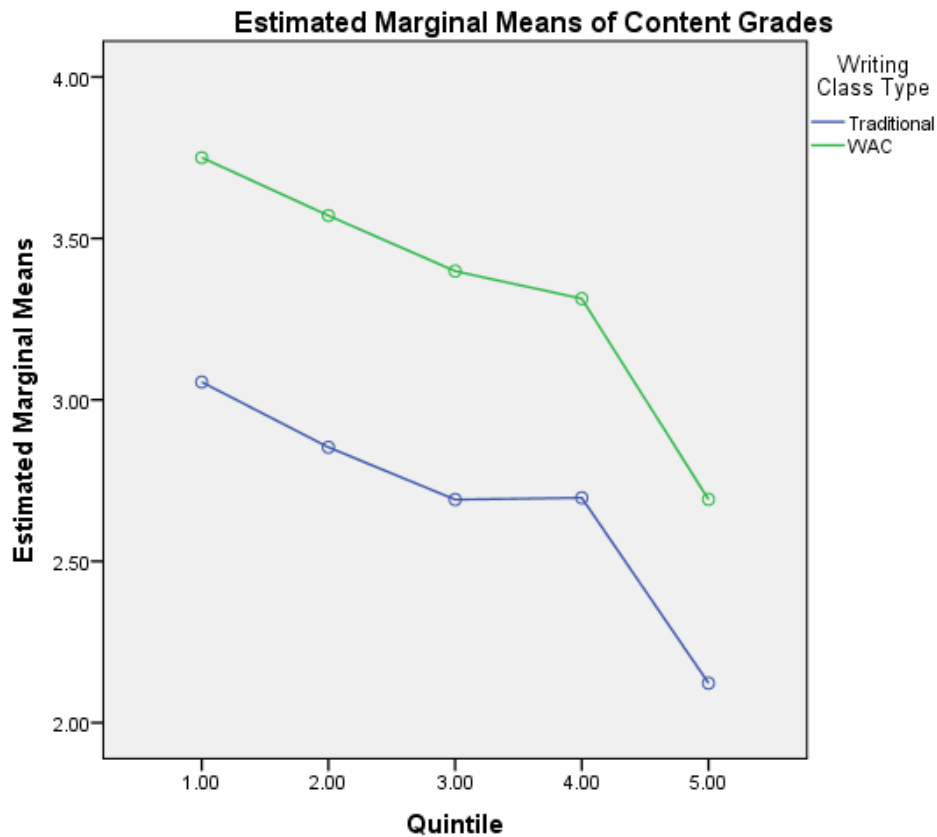
Note. Quintiles are based on propensity score using gender, geographical region, and writing grades. The fifth quintile has the highest conditional probability of choosing to be enrolled in a WAC/WID class.

Discovering OIL: The Role of Online International Learning and International Field Trips in Enhancing Student Engagement and Performance

Aaron Taylor, Human Resource Management, Coventry University London Campus
Contact: aaron.taylor@culc.coventry.ac.uk

Abstract
Online International Learning (OIL) is an innovative teaching paradigm that facilitates intercultural competence via meaningful online discussions between higher education practitioners and students in distant locations (de Wit 2013). OIL has been elucidated as a collaborative form of pedagogy that enhances 'virtual mobility', collaborative learning and the student experience (ibid).

Similarly, international field trips allow students the opportunity to enhance their cultural awareness by active learning and immersion in new, dynamic and exciting learning environments (Jakubowski 2003). Piggott (2012) argues higher education students revel in experiencing real situations that can often bring what is taught in the classroom 'to life'.