Curriculum-Based Measurement for Written Expression with Postsecondary Students with Intellectual and Developmental Disabilities

Jeremy W. Ford, Ph.D., NCSP[1]
Boise State University

Erica R. Kaldenberg, Ph.D.[2]
The University of Iowa

The purpose of this study was to replicate, and extend, previous research using curriculum-based measurement for written expression (CBM-WE) with 22 postsecondary students with intellectual and developmental disabilities from a large Midwestern university. Students were administered three CBM-WE passages and the Woodcock-Johnson Tests of Academic Achievement: Third Edition (WJIII) Broad Written Language cluster. CBM-WE passages were scored using a variety of metrics, some previously unexamined with this population. Results suggest the relation between previously examined metrics and the WJIII may be stronger than initially found. Metrics not previously examined demonstrated similar promise as when used with non-elementary age students.

Keywords: curriculum-based measurement, intellectual disabilities, developmental disabilities, postsecondary education, written expression

Since the late 20th century, federal legislation such as the Individuals with Disabilities Education Act has emphasized postsecondary education opportunities for all students, including those with intellectual and other developmental disabilities (ID/DD; Papay & Bambara, 2011). Yet, educators and researchers alike have long known that the needs of such students require more intense support compared to what their peers with less significant disabilities require (Newman, Cameto, Garza, & Levine, 2005). It is difficult to define the heterogeneous population of students with ID/DD increasingly participating in postsecondary education (Papay & Bambara, 2011). However, Plotner and Marshall (2014) state it is appropriate to note such students are those “who do not meet the academic program admission requirements used by traditional degree-seeking students” (p. 50-51). Thus, it is important to recognize this population does not include students with mild to moderate learning disabilities, those with difficulty sustaining attention, or those with higher functioning autism who – should they seek traditional postsecondary education – could choose to access support from their institution’s Disability Services Office. Given the unique needs of students with ID/DD who seek education beyond high school exceed the capacity of Disability Services Offices to provide, postsecondary education programs for this population vary considerably in their structure, options, and mission.

DOI: 10.13021/jipe.2019.2473

Postsecondary Education Program Models for Students with ID/DD

Grigal, Dwyre, & Davis (2006) identify three models for providing postsecondary education opportunities for students with ID/DD: (a) Substantially separate models, (b) Mixed/hybrid models, and (c) Inclusive individual support models. Substantially separate models involve students with ID/DD being physically located on a college or university campus but, typically, such students only participate in classes with other students with disabilities. In substantially separate models, students with ID/DD are unable to enroll in typical college classes. In mixed/hybrid models, students with ID/DD engage in academic coursework, participate in social events, and are involved in service activities along with students without disabilities. While students with ID/DD attending mixed/hybrid models take courses not available to students attending the college or university (e.g., courses in independent living or job exploration), they may also enroll in typical college courses as well. Inclusive individual support models provide students with ID/DD individualized services to facilitate their participation in typical college activities (including coursework, social events, and service activities). In order to implement such a model, a significant degree of support from academic coaches, mentors, and advisers is required.

Though the three models of postsecondary education for students with ID/DD described above function differently in many ways, similarities also exist. For example, regardless of the model, postsecondary education programs for students with ID/DD tend to focus on five domains: academic, community resources, vocational/employment, recreational, and independent living (Grigal, Hart, & Weir, 2013; Papay & Bambara, 2011). These domains represent areas of focus of postsecondary programs for students with ID/DD given such students need for continued – often intensive – related instruction and support beyond high school, not typical of their peers without ID/DD. In particular, and of interest to the current study, for students with ID/DD who seek out postsecondary education it can be expected many will do so with unsatisfactory academic skills, including in writing. As an academic skill, writing is crucial for students' achievement in grades K-12 schools, and also their success in postsecondary education and employment (Graham, Capizzi, Harris, Hebert, & Morphy, 2014).

Writing and Postsecondary Education

Despite the importance given to writing, very little data are collected on the writing ability of students – with or without disabilities – beyond elementary school (Applebee & Langer, 2011; Kiuhara, Graham, & Hawken, 2009). In particular, data for postsecondary students, again those with and without disabilities, is especially lacking and what is available is concerning. For example, in 2005 the American Association of Colleges and Universities reported that only 11% of college seniors were able to write proficiently. In addition, the National Commission on Writing (NCW, 2005) reported that 30% of college graduates must undergo remedial writing training on the job, further indicating many students do not leave college with the level of writing skills employers expect. Given this data suggests typically developing students finish their postsecondary education absent the writing skills needed for adequate performance, concern for students with ID/DD is prudent.

DOI: 10.13021/jipe.2019.2473

Fortunately, a growing research base has provided evidence that postsecondary students with ID/DD can improve their writing skills via explicit instruction (cf. Woods-Groves et al., 2014; Woods-Groves et al., 2015). However, to coincide with the implementation of evidence-based writing instruction, instructors must rely on methods to monitor student progress. While evaluation standards for postsecondary education programs for students with ID/DD are broad (Grigal, Hart, & Weir, 2012), the inclusion of the aforementioned academic domain (i.e., basic and new skill instruction) across models, requires the use of measurement tools with appropriate technical adequacy (i.e., reliability and evidence of validity) for instructional decision-making. Curriculum-based measurement (CBM; Hosp, Hosp, & Howell, 2016) is a measurement technology with empirical support for such a purpose.

Curriculum-Based Measurement

Originally developed at the University of Minnesota in the late 1970s and 1980s, CBM consists of multiple tools with appropriate technical adequacy for assisting in screening and progress decision-making across academic areas (i.e., reading, mathematics, and written expression). Fuchs (2004) described three stages of CBM research. Stage 1 CBM research is designed to provide evidence of a metric being appropriate for screening decisions. The purpose of such research is to examine the technical features of the static score (e.g., words read correctly, correct digits, correct writing sequences). In Stage 2 CBM research the technical features of slope are examined to provide evidence of a metric being appropriate for progress decisions. That is, Stage 2 CBM research involves repeated measurement of students' progress using CBM. This is done in order to determine if improving CBM scores correspond to increasing academic skill. Moreover, Stage 3 CBM research investigates the instructional utility of CBM. That is, this stage of research examines whether educators can use results in order to inform their instructional decision-making and facilitate students' academic progress.

With regard to CBM for written expression (CBM-WE), McMaster and Espin (2007) synthesized the literature on reliability and evidence of validity for students in elementary and secondary school. Postsecondary students with ID/DD were not included in this review by McMaster and Espin (2007) because (a) their interest was students in grades K-12 and (b) even if they had wanted to extend their synthesis beyond high school, no research using CBM-WE with this population had been conducted at that time. However, given these students' age and educational level it is likely their writing skills are no more than comparable to those of secondary students. Thus, the results of McMaster and Espin (2007) for this population are of more interest than those for elementary age students for considering the use of CBM-WE with postsecondary students with ID/DD. Such results included alternate-form reliability of mostly greater than $r = .70$. Further, coefficients for evidence of criterion-related validity were found to range from $r = .02$ to .99, with such a range likely due to the selected criterions and the degree to which they were directly related to writing (Campbell, Espin, & McMaster, 2013).

A second reason for such a range is related to the differences in metrics examined (McMaster & Espin, 2007). Traditional metrics for measuring students' performance include: total word written (TWW), number of words spelled correctly (WSC), number of

DOI: 10.13021/jipe.2019.2473

correct writing sequences (CWS), and the number of CWS minus the number of incorrect word sequences (CIWS). TWW is calculated by counting the number of words a student writes without considering errors in spelling or grammar. WSC is calculated by counting the number of words a student spells correctly regardless of context. A CWS is "two adjacent, correctly spelled words that are acceptable within the context of the [written] phase to a native speaker of the English language" (Videen, et al. 1982, p. 7 as cited in Hosp et al., 2016).

Researchers (e.g., Malecki & Jewell, 2003; Mercer, Martinez, Faust, & Mitchell, 2012) have grouped TWW, WSC, and CWS together as production-dependent indices while CIWS has been identified as an "accurate-production indicator." In addition, such researchers have also identified "production-independent indices" such as the percentage WSC (%WSC) and the percentage of CWS (%CWS). Much research has supported the technical adequacy of production-independent indices as the strongest indicators of writing performance for older students (Amato & Watkins, 2009; Espin, De La Paz, Scierka & Roelofs, 2005; Parker, Tindal, & Hasbrouck, 1991; Tindal & Parker, 1989; Watkinson & Lee, 1992). This is likely due to these more complex metrics better reflecting the more complex nature of older students' writing. However, production-independent metrics also fail to meet the requirements of CBM as it is possible for percentage measures to remain consistent (or vary greatly) over time even as the amount a student writes increases, decreases, or does not change (Espin, Weissenburger, & Benson, 2004). As such, production-independent metrics are unlikely to be useful to monitor students' progress over time. However, it is possible such metrics could be useful in making screening decisions. In order to determine if this is the case for postsecondary students with ID/DD, Stage 1 CBM research is necessary. In addition, research has also suggested correct punctuation marks (CPM) to be a better indicator for older students as well (Amato & Watkins, 2009; Diercks-Gransee, Weissenburger, Johnson, & Christensen, 2009; Gansle, Noell, VanDerHeyden, Naquin & Slider, 2002).

CBM and Postsecondary Students with ID/DD

While the evidence for using CBM-WE with secondary students suggests its use with postsecondary students with ID/DD to be appropriate, this should not be an automatic conclusion. Almost no research has examined the use of CBM at the postsecondary level. Two recent studies that have done so by Hosp, Hensley, Huddle and Ford (2014) and Hosp, Ford, Huddle, and Hensley (2018). These studies examined the technical adequacy of static CBM scores and thus are both examples of Stage 1 CBM research. Hosp et al. (2014) examined the use of CBM for postsecondary students with ID/DD for reading, mathematics, and written expression, finding results to be largely congruent with those related to students in K-12. Hosp et al. (2018) attempted to replicate the findings of their original study with regard to reading and mathematics, but written expression was not included (due to lack of interest of program leadership and staff at the time). There are multiple purposes to our study, as outlined below, however a primary purpose is to attempt to replicate the findings of Hosp et al. (2014) related to written expression. The findings of Hosp et al. (2014) examined traditional CBM metrics (TWW, WSC, CWS, and CWIS) for measuring student writing performance. Results found the four CBM-WE metrics to be moderately correlated with WJIII Broad Written Language Cluster (r = .53

DOI: 10.13021/jipe.2019.2473

to .67), with CIWS producing the strongest correlation (*r* = .67) followed by CWS (*r* = .65), consistent with research related to CBM-WE and secondary students (McMaster & Espin, 2007). However, the production-independent metrics of %WSC and %CWS were not examined due to issues with distribution of the study's sample.

Purpose

Our first, and primary, purpose for conducting our study was to attempt to replicate the findings of Hosp et al. (2014) regarding the relation between CBM-WE and postsecondary students with ID/DD, thus extending the Stage 1 CBM research in this area. Replication is an important means by which to improve the research base of a discipline, and the validity of findings (Cook, 2014). Further, replication allows for generating defensible knowledge. While Riley-Tillman and Burns (2009) discuss generating defensible knowledge about the effectiveness of interventions at the individual level, we believe the concept is pertinent to the use of tools for instructional decision-making at the group level as well. Thus, concluding the use of CBM-WE with postsecondary students with ID/DD as appropriate would be immature based solely on the results of Hosp et al. (2014).

In addition to our primary purpose, there are two additional purposes for our study. Our second purpose was to examine potential differences when using a single writing passage (as did Hosp et al., 2014) versus the median of three writing passages (as is standard practice in using CBM) in order to make screening decisions. Our third purpose was to examine if metrics shown to be more effective for measuring older students' writing skill level would be, in fact, more effective for measuring postsecondary students with ID/DD writing skills compared to traditional metrics. Given these purposes, the following research questions guided our study: (1) Using a different sample, can the results of Hosp et al. (2014) regarding evidence of criterion-related validity of the static score of CBM-WE metrics be replicated? (2) Does the relation between CBM-WE metrics and the WJIII vary when using performance on one CBM-WE passage compared to using the median performance from three passages? (3) Using median performance, is there a statistical difference in how different CBM-WE metrics predict performance? In particular, are the CBM-WE metrics shown to be better predictors for older students (i.e., %WSC, %CWS, CPM) better predictors of performance than traditional CBM-WE metrics (TWW, WSC, CWS, CIWS)?

Method

Participants and Setting

Participants were postsecondary students with ID/DD enrolled in their second year in a 2-year postsecondary education program designed to facilitate young adults' independence and community integration with a focus on academics, life skills/social skills, and career development. The program is located at a Midwestern research university and housed in that university's College of Education. Students live on campus in dormitories and room with others in the program. Typical college students attending the traditional university live on the same floor. Program coursework includes instruction targeting career and independent life skills (e.g., job search skills, internship opportunities) and integration in

DOI: 10.13021/jipe.2019.2473

undergraduate courses with support. Academic coursework includes instruction in reading (e.g., increasing oral reading rate), mathematics (e.g., applied computation such as budgeting and tip calculation), and writing (e.g., editing skills). In addition to their coursework, students also participate in social events and service activities offered by the university (e.g., attending sporting events, joining campus clubs).

Participants in this study were 22 students enrolled in the postsecondary education program. They were 50.0% female (n = 11) and 95.6% white (n = 20). The mean age of the participants was 19.5 years. All students had completed high school, earning a diploma or a certificate. With regard to their previously identified (i.e., prior to enrolling in the program) primary disability category, 12 students (54.5%) were diagnosed with an intellectual disability, three students (13.6%) with an intellectual disability (along with a secondary diagnosis of autism), two students (9.1%) with autism, four students (18.2%) with a learning disability, and one student (4.5%) with attention deficit / hyperactivity disorder. Three students (13.6%) did not have an official diagnosis. While a range of official disability categories are represented in our sample, all students in the study meet the aforementioned definition of a student with ID/DD provided by Plotner and Marshall (2014) in that they did not meet the admission requirements of the university associated with the program.

Moreover, it is imperative to highlight issues related to the use of labels with regard to individuals with disabilities as we believe they are especially relevant to describing our study's sample, and the population we seek to generalize our findings to. That is, while there are pros and cons to the use of labeling (Heward, Alber-Morgan, & Konrad, 2017), the underlying assumption that individual differences within a specific disability category uniquely vary compared to those within other disability categories has long been questioned (Ysseldyke & Marston, 1999). Indeed, recent scholarship (Burns et al., 2016; Fletcher & Miciak, 2017) has supported oft-repeated claims (e.g., E. Deno, 1970; Germann, 2010) that accurately distinguishing students with one disability from another – or indeed from distinguishing whether one has a disability or not – is far from an exact science.

Given that the primary disability category of students in our sample was not determined by program staff, we are unable to speak to the rationale one, as opposed to another category, was determined to be accurate. However, we again state all students meet the definition of possessing characteristics described by Plotner and Marshall (2014) as those of one with an ID/DD and, therefore, enrolled in the postsecondary education program of interest due to not meeting the admission requirements of a typical two- or four-year college or university.

Instruments

Curriculum-based measurement for written expression (CBM-WE)

We used traditional procedures for administering and scoring CBM-WE (cf. Hosp et al., 2016). Thus, students were read an open ended story starter, given 1 minute to think about and plan their writing, and reminded of the story starter halfway through their "think

DOI: 10.13021/jipe.2019.2473

time" to be thinking about the story starter. At the end of the 1 min, students were given 3 min to write a story. Prompts used in the study included: (a) "If I could fly, I would go …", (b) "The best vacation I ever had was …", and (c) "The dog jumped over the fence and …".

Woodcock-Johnson Tests of Achievement: Third Edition

We used the Woodcock-Johnson Tests of Achievement: Third Edition (WJII; Woodcock, McGrew, & Mather, 2001) as the criterion measure in our study. Our primary reason for choosing to do so is students in the program where the study took place are administered the instrument upon enrollment. The WJIII is a standardized, norm-referenced battery of achievement tests for use with individuals age two to 90. The WJIII clusters (i.e., Broad Reading, Broad Math, and Broad Written Language) have been observed to have strong reliability, generally at .90 or greater (Riverside Publishing, 2011). Specific tests are administered to measure student's skill level in each cluster. For the Broad Written Language cluster these tests are Spelling, Writing Fluency, and Writing Samples. Reliability for Broad Written Language is reported to be strong (r = .94) with reliability for the tests that comprise it ranging from .87 to .90 (Riverside Publishing, 2011).
Despite Writing Samples being administered, as it is included in the Broad Written Language cluster, the test was not individually included in our analysis as it was not a part of the Hosp et al. (2014) study.

Procedures

All CBM measures were administered by the second author during participants' regularly scheduled "special topics" course using standardized procedures. Special topics courses are designed to provide academic skill, or adaptive behavior, instruction to students in the program. Each student completed three CBM writing samples. Two passages were collected during one class period and the third passage was collected two days later during the same class period. All students were able to sufficiently respond to prompts and no accommodations were provided. Students completed passages in the same order due to the need for group administration. WJIII data were independently collected by appropriately trained staff from the postsecondary program as part of regular administrative procedures. CBM-WE data were collected in the fall of 2013, with WJIII data collected the previous year in the fall of 2012 as part of admissions procedures for the program. Each CBM-WE writing probe was de-identified and scored independently by the first and second author. The authors met and reconciled the scores across multiple sessions. Initial inter-rater agreement was as follows: CWS = 90.1%, TWW = 99.0%, WSC = 96.9%, and CIWS = 80.8%. Differences between CWS and CIWS are due to initial differences in scoring between Hosp et al. early in the scoring process. All differences were reconciled with CBM scoring guidelines (Hosp et al., 2016), resulting in 100% agreement prior to data being analyzed. Despite Writing Samples being administered, as it is included in the Broad Written Language cluster, the test was not individually included in our analysis as it was not a part of the Hosp et al. (2014) study.

Data Analysis

DOI: 10.13021/jipe.2019.2473

In order to examine technical adequacy, we first calculated descriptive statistics for each CBM metric. Next, we calculated bivariate correlations between each CBM metric for each passage, as well as the median, and each content-appropriate cluster/test from the WJIII. Significance was noted at traditional *p*-values of < .1, < .05, and < .01 and further noted with a Bonferroni correction ($p < .0007$) to account for multiple comparisons. Obtained correlations were further compared to the findings of Hosp et al. (2014), presented in the discussion section, using criteria established by Marston (1989) to interpret the strength of relations in CBM research (i.e., strong relations, $r = \geq .70$; moderate relations, $r = .50$ to .69, and weak relations, $r = \leq .50$. Last, following Fisher's *r* to *z* transformation, the bivariate correlations obtain from each median CBM metric were compared using Meng's *z* in order to determine which were statistically significantly better predictors (Meng, Rosenthal, & Rubin, 1992). For example, Meng's *z* test allowed us to examine the correlations between students' median TWW score and median CWS to Broad Written Language and compare these results to determine which metric has a better predictor (i.e., had a stronger correlation) to performance on the WJII cluster.

Results

Table 1 includes descriptive statistics for each CBM metric across passages and the median as well as descriptive statistics for the WJIII. CBM metrics across passages and the median were judged for deviations of skewness and kurtosis with values above 1.0 considered questionable and above 2.0 problematic (Tabachnick & Fidell, 2013). Congruent with the findings of Hosp et al. (2014), the traditional CBM-WE metrics (TWW, WSC, CWS, and CIWS) all demonstrated low levels of skewness and kurtosis. Unlike Hosp et al. (2014), the degree of skewness and kurtosis for accuracy metrics (i.e., % of CWS and % of WSC) were largely not observed to be problematic, though they remain questionable. Given the nature of our research questions these metrics, along with CPM, which also showed questionable kurtosis, were included for additional analyses. Note, further results will be discussed in terms of "traditional metrics" (i.e., TWW, WSC, CWS, and CIWS), "less common metrics" (i.e., %WSC, %CWS, and CPM), and "differences in prediction" (i.e., Meng's *z* test) by cluster/test. Bivariate correlations for metrics and Broad Written Language can be found in Table 2, while bivariate correlations for metrics and Spelling and Writing Fluency can be found in Tables 3 and 4. Results using Meng's *z* test to examine differences in prediction can be found in Table 5.

Broad Written Language

Traditional metrics

For TWW, the first and third passages were found to have a weak relation ($r = .442$ and .322 respectively) with Broad Written Language and the second passage was found to have a strong relation ($r = .702$). TWW for the median passage was found to have a moderate relation ($r = .524$) with the cluster. For WSC, individual passages were found to have a weak to strong relation ($r = .398$ to .720), while the median passage was found to have a moderate relation ($r = .620$). The relation between CWS and Broad Written Language was found to be mostly strong ($r = .537$ to .745), as was the relation with the median passage ($r = .749$). The relation between individual passages for CIWS and the

DOI: 10.13021/jipe.2019.2473

cluster was also mostly strong (r = .534 to .766), and the median passage was again found to have a strong relation (r = .769) as well.

Less common metrics

For %WSC and %CWS, individual passages were found to have a mostly strong relation (r = .534 to .766, .627 to .774 respectively) with Broad Written Language, while the median passages were found to a have a strong relation (r = .854 and .810 respectively). For CPM, individual passages were found to have a mostly moderate relation (r = .493 to .572), while the median passage was found to have a moderate relation (r = .515) as well.

Differences in prediction

Most other metrics did a better job than TWW for predicting performance on the Broad Written Language cluster of the WJIII. For example, WSC, CWS, and %WSC were all statistically better predictors of Broad Written Language than TWW (z = 3.81, 2.86, and 2.58, respectively, $p < .01$). CIWS and %CWS were also both statistically better predictors of the cluster than TWW (z = 1.99 for both, $p < .05$).

In addition, CWS and %WSC were observed to be statistically better predictors of performance on Broad Written Language than WSC (z = 2.14 and 2.13, respectively, $p < .05$). Further, CPM was observed to be a statistically better predictor of performance on Broad Written Language than CIWS, %WSC, and %CWS (z = 2.07, 2.53, and 2.17, respectively, $p < .05$).

Spelling

Traditional metrics

For TWW, individual passages were found to have a weak to moderate relation (r = .307 to .662) with Spelling, while the median passage was found to have a moderate relation (r = .502). For WSC, individual passages also were found to have a weak to moderate relation (r = .387 to .406) and the median passage was found to have a moderate relation (r = .582) as well. A moderate relation was found between CWS and Spelling for individual passages (r = .526 to .694) and the median passage (r = .690). For CIWS, the relation for the first and third passages was moderate (r = .617 and .520 respectively) while the relation for the second and median passages was found to be strong (r = .701 and .702 respectively).

Less common metrics

A moderate relation (r = .615) was found for %WSC and Spelling for the first individual passage, while strong relations were found for the second, third, and median passages (r = .764, .836, and .896 respectively). For %CWS, a strong relation was found with Spelling for all individual passages (r = .701 to .751) and the median passage (r = .783). The

DOI: 10.13021/jipe.2019.2473

relation between CPM and spelling was mostly found to be weak (r = .372, first passage; r = .519, second passage; r = .408, third passage; and r = .395, median passage).

Differences in prediction

Similar to Broad Written Language, most metrics did a better job of predicting performance on the Spelling test compared to TWW. For example, WSC and %WSC were both statistically better predictors of Spelling compared to TWW (z = 3.10 and 3.33, respectively, p < .01 for both). CWS was also a statistically better predictor of Spelling compared to TWW (z = 2.27, p < .01). %CWS was a statistically better predictor of Spelling compared to TWW (z = 1.85, p < .10) as well.

In addition, %WSC and CWS were observed to be statistically better predictors of Spelling compared to WSC (z = 3.05, p < .01 and z = 1.68, p < .10, respectively). Further, %WSC was observed to be a statistically better predictor compared to CWS and CIWS (z = 2.61, p < .01 and z = 2.58, p < .01, respectively). %WSC was also observed to be a statistically better predictor of Spelling compared to %CWS (z = 2.12, p < .05). Last, CPM was observed to be a statistically better predictor of Spelling than CWS (z = 1.80, p < .10), CIWS (z = 2.23, p < .05), %WSC (z = 3.75, p < .01), and %CWS (z = 2.58, p < .01).

Writing Fluency

Traditional metrics

For TWW, WSC, and CWS individual passages were found to have a weak to strong relation (r = .230 to .750, .264 to .740, and .425 to .740 respectively) with Writing Fluency, while the median passage was found to have a strong relation (r = .752, .761, and .791 respectively). For CIWS, individual passages were found to have a moderate relation (.545 to .690), while the median passage was found to have a strong relation (r = .710).

Less common metrics

For %WSC, %CWS, and CPM individual passages were found to have mostly moderate relations with Writing Fluency (r = .435 to .642, .500 to .649, and .550 to .671 respectively). The relation between median passages for %WSC, %CWS, and CPM and Writing Fluency was moderate (r = .583, .638, and .560 respectively).

Differences in prediction

Few differences in prediction were observed for CBM metrics and the Writing Fluency test of the WJIII as well. CWS was observed to be a better predicator of Written Fluency compared to %WSC and % CWS (z = 1.99 and 2.09, p < .05 for both). In addition, TWW, WSC, and CWS were all found to be better predictors of Written Fluency compared to CBM (z = 1.69, 1.88, and 1.71, respectively, p < .10 for all).

DOI: 10.13021/jipe.2019.2473

Discussion

We conducted our study given the importance of replication for building scientific knowledge (cf. Francis, 2012; Jasny, Chin, Chong, & Vignieri, 2011) and the increase of attention given to the topic recently in special education (e.g., Cook, 2014). While the number of students with ID/DD attending postsecondary schools is increasing, and the research base including such students is growing, definitive empirically-based conclusions are limited. One area of emerging research is the use of CBM with this population. The initial research (Hosp et al., 2014) examining the use of CBM with postsecondary students with ID/DD explored evidence of criterion-related validity for CBM reading, mathematics and written expression. Their results largely found that these CBM tools have the potential to function for such students in a similar manner as they do for students in kindergarten through high school. The goal of replication has also been described by Hosp et al. (2018) which was limited to CBM for reading and mathematics. At the time of Hosp et al. (2018) WJIII written language tests scores for students were not available, however, such data continued to be collected by program staff, and the writing skills of students in the program was anecdotally observed by staff to vary greatly across cohorts. Given the investigation of Hosp et al. (2018) to replicate the findings of Hosp et al. (2014) did not include examination of CBM-WE, a need to do so was present. In discussing our results, we compare the relation between CBM metrics and cluster/test performance on the WJIII for written language to those observed by Hosp et al. (2014). Marston's (1989) guidelines for interpreting for the strength of relations in CBM research are used to frame this comparison.

Broad Written Language

The traditional metrics previously investigated by Hosp et al. (2014) were all observed to have a moderate relation with the Broad Written Language cluster of the WJIII. In our study, however, we observed the strength of the relation between CBM metrics and the WJIII varied across individual and median passage performance. For example, while the relation for TWW and WSC on the median passage was observed to have a moderate relation with Broad Written Language, individual passages were observed to have either weak or strong relations. We observed a similar pattern for CWS in our study with the first individual and median passages observed to have a moderate relation to Broad Written Language, but strong and weak relations observed for the second and third individual passages respectively. Further, while Hosp et al. (2014) observed a moderate relation for CWS and CIWS to Broad Written Language, in our study a mostly strong relation was observed across individual passages, and the median passage for both metrics was observed to have a strong relation as well.

Hosp et al. (2014) did not examine the metrics of %WSC and %CWS due to problematic levels of skewness and kurtosis. However, in our replication study these indices were found to be questionable and included in analyses for exploration. Given the questionable nature of our data for %WSC and %CWS, interpretation should be cautious. It is noteworthy though that mostly strong relations were found between individual and median passages and Broad Written Language for these metrics. This finding is congruent with previous mentioned findings related to the use of production-independent indices for

DOI: 10.13021/jipe.2019.2473

predicting older students' writing performance. Hosp et al. (2014) note such complex indices were not relevant for the students in their study as many students wrote very few words, including several students who only rewrote the provided prompt. However, results from our replication study suggest the relation between %WSC and %CWS may warrant further investigation with postsecondary students with ID/DD. That said, it is important to note these metrics had limitations in terms of their ability to predict performance on Broad Written Language with both metrics failing to do a better job than the traditional CWS.

While the CPM metric was not included at any point in the Hosp et al. (2014) study, we included it (a) because of the aforementioned call by researchers for additional investigation and (b) as a means of exploring whether a simple metric to calculate could do as good of a job as predicting performance as metrics more complex to calculate (e.g., CWS) when making screening decisions. Our results found CPM to be no better at predicting student performance on Broad Written Language than TWW, WSC, and CWS. In addition, CIWS, %WSC, and %CWS were each found to be a better predictor than CPM for this cluster.

Spelling

The traditional metrics investigated by Hosp et al. (2014) were all found to have weak relations with the Spelling test of the WJIII. However, we found the median passage for TWW, WSC, and CWS to have a moderate relation with Spelling and weak to moderate relations for these metrics on individual passages. Further, we found moderate to strong relations for CIWS across individual and median passages.

Similar to the finding related to their relation with Broad Written Language, the accuracy based metrics of %WSC and %CWS were observed to have mostly strong relations with Spelling. Unlike with Broad Written Language, %WSC was found to be a better predictor than all other metrics, including the traditional CWS. Such a finding would appear to make sense as spelling ability would be assumed to be measured by %WSC. A related interesting finding was the observation of the WSC metric not being found to be a better predictor than CWS.

With regard to CPM, most other metrics were found to be better predictors for Spelling (no difference in prediction was found for TWW and WSC). It may not be surprising a metric of punctuation does not predict well to spelling ability, however it is perhaps worthwhile to highlight the finding that there was no difference in prediction for CPM and WSC to Spelling.

Writing Fluency

As with Spelling, Hosp et al. (2014) found weak relations for CBM-WE metrics and the Writing Fluency test from the WJIII. Again, our study largely failed to replicate those findings. TWW, WSC, and CWS on the median passage were all found to have a strong relation with Writing Fluency. Further, on individual passages the metrics were observed to have weak to strong relations. In addition, CIWS on the median passage was observed

DOI: 10.13021/jipe.2019.2473

to have a strong relation with Writing Fluency and on individual passages a moderate relation was observed.

Mostly moderate relations were found for the metrics not previously examined by Hosp et al. (2014). The finding that the traditional CWS was a better predictor than both %WSC and %CWS, as well as CPM, provides evidence for its continued use in CBM-WE.

Limitations

The implications of our findings, while insightful, must be balanced with our study's limitations. As such, we first note our small sample size and its lack of national representation. Regarding sample size, without more students we were unable to examine potential differences in the relation of different CBM-WE metrics and the WJIII for subgroups. Thus, inference of our results to all postsecondary students with ID/DD, may not be appropriate.

Further, the population that is identified as having an ID/DD is a heterogeneous one making generalization from our (or any) study difficult. In addition, our study continued the examination of validity by only investigating evidence of criterion-related validity, just as Hosp et al. (2014) and Hosp et al. (2018). In order to fully evaluate the technical adequacy of CBM-WE with postsecondary students with ID/DD, examination of other facets of the multifaceted nature of validity (Messick, 1989) must be done.

Further still, with emerging evidence of the validity of CBM with this population (Hosp et al., 2014; Hosp et al., 2018), the examination of reliability should also be examined. Though three passages were collected from each student in our study, our small sample size does not allow for calculating alternate-form reliability. Also concerning the use of three passages, it must be noted that story prompts were not counter-balanced during administration. This was due to logistical considerations of the postsecondary program where the data was collected.

Implications and Future Research

If the purpose of a replication study is to confirm what was already observed, our study provides mixed results. If the purpose of a replication study is to "improve scientific research and the validity of its findings" (Cook, 2014 p. 233) there are a plethora of implications from our study for practice and future research. Where our results confirm previous findings of Hosp et al. (2014), steps toward creating general, defensible knowledge about using CBM-WE with postsecondary students with ID/DD are taken; where our results counter, or add nuance, to the findings of previous research, we find areas for further study in order to create such knowledge.

Variability in the relation between individual passages and criterion measures suggests that in order to increase accuracy for making screening decisions about the writing skills of postsecondary students with ID/DD, administering one passage is not sufficient. Such a finding is consistent with other CBM research. Indeed, Hosp et al. (2014) note the administration of only one CBM-WE passage to be a limitation of their study, and called

DOI: 10.13021/jipe.2019.2473

for their research to be replicated. We believe our study is an example of the nature of research building on previous findings while seeking to improve methodology in order to increase understanding. Differences between our findings and Hosp et al. (2014) are likely due to our use of multiple passages providing for a more accurate measurement of students' skills given the increase in stability of measuring students' performance associated with following typical administration guidelines.

Specifically, while Hosp et al. (2014) observed a moderate overall relation between the traditional CWS and CIWS metrics and Broad Written Language, our observations from this study suggest the relation may be stronger. Thus, additional study is necessary in order to examine this relation further. In addition, also related to the relation of CBM-WE metrics and the WJIII, our study included a sample considered normally distributed when examining production-independent metrics. Though not the case for Hosp et al. (2014), this finding is congruent with research using CBM-WE and older students. Such a finding also supports the need for additional exploration. Moreover, it is worth highlighting the relations between all metrics and the Spelling and Writing Fluency tests of the WJIII. In the Hosp et al. (2014) study weak relations were found for all metrics and these tests. However, our results found much stronger relations. In particular, stronger relations were observed for median performance from three administered passages and With regard to Writing Fluency.

In addition, much work has been done developing CBM scores with meaningful cut-scores for predicting student performance. Typically, this has been accomplished by examining the relation between students' CBM performance and state tests for accountability purposes (e.g., Good et al., 2013; Patton, Reschly, & Appleton, 2014). While one would likely be hard pressed to find an argument against having meaningful academic expectations (such as a performing adequately on a state test for accountability) for postsecondary students with ID/DD, identifying such expectations is difficult. That is, while CBM can be used to predict the likeliness of K-12 students meeting expectations on tests of accountability, it is unclear what equivalent outcomes should be used to predict performance to for postsecondary students with ID/DD. Future research should focus on establishing such outcomes.

Absent a clear criterion for outcome performance, however, the use of CBM with postsecondary students with ID/DD still holds promise. For example, in programs such as the one involved in our study, courses focused on basic academic skill instruction should have technically accurate (i.e., valid and reliable) tools for measuring students' progress. CBM is a potential means of doing so. However, to return to Fuchs (2004) Stages of CBM Research, only emerging evidence of the technical adequacy of the static score is currently available (Stage 1) with regard to postsecondary students with ID/DD. Moreover, research examining Stage 2 (technical adequacy of the slope of performance) and Stage 3 (instructional utility) remain to be conducted with this population for CBM-WE, or any other CBM tool.

Despite this, we do not believe programs need wait to use CBM with postsecondary students with ID/DD until additional evidence is collected (although caution should be used when interpreting results prior to this being done). Rather, programs can use CBM

DOI: 10.13021/jipe.2019.2473

to collect proximal data about students' academic skills which may allow program staff to detect changes in academic skill that standardized norm-referenced academic achievement tests are not able to observe. Further, for programs that emphasize academic skill instruction, it is imperative that instructional decisions are based on data in order to maximize the likeliness of students acquiring targeted skills. The field of special education has a long history of supporting data-based decision-making (E. Deno, 1970; Reschly, Tilly, & Grimes, 1999) in K-12 schools, and CBM has played a pivotal role in reform efforts to do so (Burns, Jimerson, VanDerHeyden, & S. Deno, 2016). As students with ID/DD continue to seek out postsecondary education, steps toward implementing CBM practices with programs which serve such students holds promise for improving instructional outcomes of students with ID/DD.

Conclusion

This study continued a line of research examining the potential use of CBM with postsecondary students with ID/DD by seeking to replicate previous findings of the Hosp et al. (2014). Replication serves a vital role in establishing general knowledge about effective assessment and evaluation practices. In the case of our study, certain findings from initial research hold true but other results – possibly due to differences in our present sample compared to the initial sample – were not replicated. Differences across samples likely reflects addressing limitations from the original study, the heterogeneous make-up of postsecondary students with ID/DD, and the challenge of accurately "placing" an individual within a specific disability category. Therefore, additional research continues to be warranted prior to accepting definitive answers for using CBM-WE with postsecondary students with ID/DD.

References

Amato, J. M., & Watkins, M. W. (2009). The predictive validity of CBM writing indices for eighth-grade students. *The Journal of Special Education, 44*, 195-204.

Applebee, A. N., & Langer, J. A. (2011). "EJ" extra: A snapshot of writing instruction in middle schools and high schools. *The English Journal*, *100*(6), 14-27.

Burns, M. K., Jimerson, S. R., VanDerHeyden, A. M., & Deno, S. L. (2016). Toward a unified response-to-intervention model: Multi-tiered systems of support. In S.R. Jimerson, M.K. Burns, & A.M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (pp. 719-732). Springer, Boston, MA US.

Burns, M. K., Petersen-Brown, S., Haegele, K., Rodriguez, M., Schmitt, B., Cooper, M., ... & VanDerHeyden, A. M. (2016). Meta-analysis of academic interventions derived from neuropsychological data. *School Psychology Quarterly*, *31*(1), 28.

Campbell, H., Espin, C. A., & McMaster, K. (2013). The technical adequacy of curriculum-based writing measures with English learners. *Reading and Writing*, *26*(3), 431-452.

Cook, B. G. (2014). A call for examining replication and bias in special education research. *Remedial and Special Education*, *35*(4), 233-246.

Deno, E. (1970). Special education as developmental capital. *Exceptional Children, 37,* 229-237.

DOI: 10.13021/jipe.2019.2473

Espin, C. A., De La Paz, S., Scierka, B. J., & Roelofs, L. (2005). The relationship between curriculum-based measures in written expression and quality and completeness of expository writing for middle school students. *The Journal of Special Education*, *38*(4), 208-217.

Espin, C. A., Weissenburger, J. W., & Benson, B. J. (2004). Assessing the writing performance of students in special education. *Exceptionality*, *12*(1), 55-66.

Fletcher, J. M., & Miciak, J. (2017). Comprehensive cognitive assessments are not necessary for the identification and treatment of learning disabilities. *Archives of Clinical Neuropsychology*, *32*(1), 2-7.

Francis, G. (2012). The psychology of replication and replication in psychology. *Perspectives on Psychological Science*, *7*(6), 585-594.

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, *33*(2), 188-193.

Gansle, K. A., Noell, G. H., VanDerHeyden, A. M., Naquin, G. M., & Slider, N. J. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternate measures for curriculum-based measurement in writing. *School Psychology Review*, *31*(4), 477-497.

Germann, G. (2010). Thinking of yellow brick roads, emerald cities, and wizards. In M. R. Shinn & H. M. Walker (Eds.), Interventions for achievement and behavior problems in a three-tier model including RTI (pp. xiii-xxxv). Washington, DC: NASP Publications.

Good, R., Kaminski, R., Dewey, E., Wallin, J., Powell-Smith, K., & Latimer, R. (2013). *DIBELS Next Technical Manual.* Eugene, OR: Dynamic Measurement Group.

Graham, S., Capizzi, A., Harris, K. R., Hebert, M., & Morphy, P. (2014). Teaching writing to middle school students: A national survey. *Reading and Writing*, *27*(6), 1015-1042.

Gransee, B., Weissenburger, J. W., Johnson, C. L., & Christensen, P. (2009). Curriculum-based measures of writing for high school students. *Remedial and Special Education*, *30*(6), 360-371.

Grigal, M. Dwyre, A., & Davis, H. (2006). *Transition service for students aged 19-21 with intellectual disabilities in college and community setting: Models and implications for success*. Minneapolis: University of Minnesota, National Center on Secondary Education and Transition.

Grigal, M., Hart, D., & Weir, C. (2012). A survey of postsecondary education programs for students with intellectual disabilities in the United States. *Journal of Policy and Practice in Intellectual Disabilities*, *9*(4), 223-233.

Grigal, M., Hart, D., & Weir, C. (2013). Postsecondary education for people with intellectual disability: Current issues and critical challenges. *Inclusion*, *1*(1), 50-63.

Heward, W. L., Alber-Morgan, S. R., & Konrad, M. (2017). *Exceptional children: An introduction to special education* (11th ed.). Pearson Education, Inc. Boston, MA.

Hosp, J. L., Hensley, K., Huddle, S. M., & Ford, J. W. (2014). Using curriculum-based measures with postsecondary students with intellectual and developmental disabilities. *Remedial and Special Education*, *35*(4), 247-257.

Hosp, J. L., Ford, J. W., Huddle, S. M., & Hensley, K. K. (2018). The importance of replication in measurement research: Using curriculum-based measures with postsecondary students with developmental disabilities. *Assessment for Effective*

DOI: 10.13021/jipe.2019.2473

Intervention, *43*(2), 96-109.
Hosp, M. K., Hosp, J. L., & Howell, K. W. (2016).). *The ABCs of CBM: A practical guide to curriculum-based measurement* (2nd ed.). New York: Guilford.
Jasny, B. R., Chin, G., Chong, L., & Vignieri, S. (2011). Again, and again, and again…. *Science*, *334*(6060), 1225-1225.
Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology*, *101*(1), 136.
Malecki, C. K., & Jewell, J. (2003). Developmental, gender, and practical considerations in scoring curriculum-based measurement writing probes. *Psychology in the Schools*, *40*(4), 379-390.
Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. R. Shinn (Ed.), *The Guilford school practitioner series. Curriculum-based measurement: Assessing special children* (pp. 18-78). New York, NY, US: Guilford Press.
McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing: A literature review. *The Journal of Special Education*, *41*(2), 68-84. doi:10.1177/00224669070410020301.
Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*(1), 172.
Mercer, S. H., Martínez, R. S., Faust, D., & Mitchell, R. R. (2012). Criterion-related validity of curriculum-based measurement in writing with narrative and expository prompts relative to passage copying speed in 10th grade students. *School Psychology Quarterly*, *27*(2), 85-95.
Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*(2), 5-11.
National Commission on Writing [NCW] (2005, July). *Writing: A powerful message from State government.* Retrieved from www.collegeboard.com
Papay, C. K., & Bambara, L. M. (2011). Postsecondary education for transition-age Students with intellectual and other developmental disabilities: A national survey. *Education and Training in Autism and Developmental Disabilities, 46*, 78-83.
Parker, R., Tindal., G., & Hasbrouck, J. (1991). Countable indices of writing quality: Their suitability for screening-eligibility decisions. *Exceptionality: A Special Education Journal*, *2*(1), 1-17.
Patton, K. L. S., Reschly, A. L., & Appleton, J. (2014). Curriculum-based measurement as a predictor of performance on a state assessment: Diagnostic efficiency of local norms. *Educational Assessment*, *19*(4), 284-301.
Plotner, A. J., & Marshall, K. J. (2014). Navigating university policies to support postsecondary education programs for students with intellectual disabilities. *Journal of Disability Policy Studies*, *25*(1), 48-58.
Reschly, D. J., Tilly III, W. D., & Grimes, J. P. (1999). *Special Education in Transition: Functional Assessment and Noncategorical Programming*. Sopris West, Longmont, CO 80507.
Riley-Tillman, T. C., & Burns, M. K. (2009). *Single case design for measuring response to educational intervention.* New York: Guilford.
Riverside. (2011). *Woodcock-Johnson III Normative Update*. Retrieved from http://www.riversidepublishing.com/products/wjIIIComplete/details.html#reliability
Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston,

DOI: 10.13021/jipe.2019.2473

MA: Pearson.
Tindal, G., & Parker, R. (1989). Assessment of written expression for students in compensatory and special education programs. *The Journal of Special Education*, *23*(2), 169-183.
Wagner, M., Newman, L., Cameto, R., Garza, N., & Levine, P. (2005). *After high school: A first look at the postschool experiences of youth with disabilities. A report from the National Longitudinal Transition Study-2 (NLTS-2)*. Menlo Park, CA: SRI International.
Watkinson, J. T., & Lee, S. W. (1992). Curriculum-based measures of written expression for learning-disabled and nondisabled students. *Psychology in the Schools*, *29*(2), 184-191.
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock Johnson Tests of Achievement*. Itasca, IL. Riverside.
Woods-Groves, S, Hua, Y., Therrien, W.J., Kaldenberg, E. R., Hendrickson, J. M., Lucas, K.G., & McAninch, M. J., (2014). An investigation of strategic writing instruction for postsecondary students with developmental disabilities. *Education and Training in Autism and Developmental Disabilities*, *49(*2), 248-262.
Woods-Groves, S., Hua, Y., Therrien, W. J., Kaldenberg, E. R., Kihura, R. W., & Hendrickson, J. M. (2015). An investigation of the efficacy of an editing strategy with postsecondary individuals with developmental disabilities. *Education and Training in Autism and Developmental Disabilities*, *50(1),* 95-108.
Ysseldyke, J., & Marston, D. (1999). Origins of categorical special education services in schools and a rationale for changing them. *Special education in transition: Functional assessment and noncategorical programming*, 1-18.

DOI: 10.13021/jipe.2019.2473

Table 1
Descriptive Statistics for CBM-WE Metrics and the WJIII

Metric / Criterion	Passage / Median	Mean	SD	Skewness	Kurtosis
TWW	1	35.50	15.72	.87	.93
	2	41.86	20.47	.30	-.20
	3	33.73	13.94	.10	-.80
	Median	35.59	13.08	-.07	-.71
WSC	1	33.36	15.75	.78	.41
	2	39.59	20.91	.39	-.90
	3	32.32	14.03	.03	-.92
	Median	33.32	13.24	-.04	-.77
CWS	1	28.86	15.85	.39	-.43
	2	36.32	23.18	.62	.18
	3	26.27	15.29	.57	-.56
	Median	29.36	15.18	.06	-.79
CIWS	1	17.91	19.00	.73	.02
	2	26.46	25.22	.67	.24
	3	14.59	19.74	.57	-.74
	Median	19.18	18.39	.20	-.81
%WSC	1	92.84%	7.39%	-1.26	1.59
	2	92.0%	7.83%	-1.47	1.96
	3	94.80%	7.26%	-1.52	1.63
	Median	92.62%	6.91%	-1.47	2.54
%CWS	1	72.45%	21.04%	-.34	-.84
	2	71.45%	23.35%	-1.19	.57
	3	69.83%	24.76%	-.33	-1.38
	Median	71.42%	22.12%	-.68	-.57
CPM	1	2.91	2.62	.863	.045
	2	3.86	2.87	.030	-1.16
	3	2.64	1.94	.010	-1.19
	Median	2.91	2.16	-.028	-1.36
WJ III	Broad	69.59	17.05	-.35	-.86
	Spelling	77.182	25.35	-.58	-.03
	Writing Fluency	74.86	15.35	-.67	.71

Note. *N* = 22; CBM-WE = curriculum-based measurement for written expression; WJIII = Woodcock Johnson Tests of Academic Achievement-Third Edition; SD = standard deviation; TWW; total words written; WSC = words spelled correct; CWS = correct writing sequences; CISW = correct minus incorrect writing sequences; CPM = correct punctuation marks; Broad = Broad Written Language.

DOI: 10.13021/jipe.2019.2473

Table 2

Bivariate Correlations for CBM-WE Metrics and the Broad Written Language Cluster of the WJIII

Metric	Passage 1	Passage 2	Passage 3	Median
TWW	.442**	.702***	.322	.524*
WSC	.505*	.720***	.398	.620**
CWS	.704***	.745***	.537**	.749***
CIWS	.766***	.755***	.534*	.769***
%WSC	.627**	.732***	.774***	.854***
%CWS	.840***	.754***	.687***	.810***
CPM	.537**	.572**	.493*	.515*

Note. $N = 22$; CBM-WE = curriculum-based measurement for written expression; WJIII = Woodcock Johnson Tests of Academic Achievement-Third Edition; TWW; total words written; WSC = words spelled correct; CWS = correct writing sequences; CISW = correct minus incorrect writing sequences; CPM = correct punctuation marks.
* $p < .05$; ** $p < .01$; *** $p < .001$.

Table 3

Bivariate Correlations for CBM- WE Metrics and the Spelling Test of the WJIII

Metric	Passage 1	Passage 2	Passage 3	Median
TWW	.347	.662**	.307	.502^
WSC	.406^	.681***	.387^	.582**
CWS	.572**	.694***	.526*	.690***
CIWS	.617**	.701***	.520*	.702***
%WSC	.615**	.764***	.836***	.896***
%CWS	.751***	.734***	.701***	.783***
CPM	.372^	.519*	.408^	.395^

Note. $N = 22$; CBM-WE = curriculum-based measurement for written expression; WJIII = Woodcock Johnson Tests of Academic Achievement-Third Edition; TWW; total words written; WSC = words spelled correct; CWS = correct writing sequences; CISW = correct minus incorrect writing sequences; CPM = correct punctuation marks.
^ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

DOI: 10.13021/jipe.2019.2473

Table 4

Bivariate Correlations for CBM-WE Metrics and the Writing Fluency Test of the WJIII

Metric	Passage 1	Passage 2	Passage 3	Median
TWW	.230	.750***	.607**	.752***
WSC	.264	.740***	.648**	.761***
CWS	.425**	.740***	.693***	.791***
CIWS	.545**	.681***	.690***	.710***
%WSC	.435*	.642**	.612*	.583**
%CWS	.500*	.649**	.568**	.638**
CPM	.550**	.547**	.671**	.560**

Note. *N* = 22; CBM-WE = curriculum-based measurement for written expression; WJIII = Woodcock Johnson Tests of Academic Achievement-Third Edition; TWW; total words written; WSC = words spelled correct; CWS = correct writing sequences; CISW = correct minus incorrect writing sequences; CPM = correct punctuation marks.
* $p < .05$; ** $p < .01$; *** $p < .001$.

Table 5

Differences in prediction for CBM-WE metrics and the WJIII, Meng's z

Criterion	Metric	TWW	WSC	CWS	CIWS	%WSC	%CWS	CPM
WJ III Broad Written Language	TWW	0	-3.81**	-2.86**	-1.99*	-2.58**	-1.99*	0.07
	WSC		0	-2.14*	-1.43	-2.13*	-1.53	0.86
	CWS			0	-0.40	-1.29	-0.94	1.60
	CIWS				0	-1.11	-0.71	2.07*
	%WSC					0	0.78	2.53*
	%CWS						0	2.17*
WJ III Spelling	TWW	0	-3.10**	-2.27*	-1.51	-3.33**	-1.85^	0.76
	WSC		0	-1.68^	-1.06	-3.05**	-1.51	1.44
	CWS			0	-0.22	-2.61**	-1.31	1.80^
	CIWS				0	-2.58**	-1.28	2.23*
	%WSC					0	2.12*	3.75**
	%CWS						0	2.58**
WJ III Writing Fluency	TWW	0	-0.45	-0.61	0.39	1.19	0.81	1.69^
	WSC		0	-0.58	0.52	1.37	0.94	1.88^
	CWS			0	1.62	1.99*	2.09*	1.71^
	CIWS				0	1.17	1.01	1.19
	%WSC					0	-0.65	0.13
	% CWS						0	0.50

Note. *N* = 22; CBM-WE = curriculum-based measurement for written expression; WJIII = Woodcock Johnson Tests of Academic Achievement-Third Edition; TWW; total words written; WSC = words spelled correct; CWS = correct writing sequences; CISW = correct minus incorrect writing sequences; CPM = correct punctuation marks.
^ $p < .10$; * $p < .05$; ** $p < .01$.

DOI: 10.13021/jipe.2019.2473

About the Authors

[1]Jeremy W. Ford, Ph.D., NCSP is an assistant professor in the Early & Special Education Department Boise State University. Email: jwford@boisestate.edu

[2]Erica R. Kaldenberg is the Assistant Director of UI REACH in the College of Education, University of Iowa. Email: erica-kaldenberg@uiowa.edu

DOI: 10.13021/jipe.2019.2473