The Course on Norm-referenced Academic Assessment: A Survey of Special Education Faculty

Adam B. Lockwood

Kent State University

Ryan L. Farmer

Oklahoma State University

Margaret Schmitt

Western Kentucky University

Karen Sealander

Christopher Lanterman

Northern Arizona University

Megan Adkins

Julie Billiard Schools

Author Note

This is the pre-peer reviewed version of the following article: “The course on norm-referenced academic assessment: A survey of special education faculty,” which has been published in final form at doi.org/10.1002/pits.22616. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This manuscript is not the final version and differs from the final, published version. The authors declare that they have no conflict of interest. We wish to thank all the instructors who completed this survey and who work tirelessly to train special educators. Correspondences concerning this article should be addressed to Adam Lockwood, Kent State University. Email: alockwo2@kent.edu.

Please cite as:

Abstract

Special education teachers play a key role in educational assessment practices, including frequently administering norm-referenced tests (NRTs) of academic achievement. This study examined the course on norm-referenced assessment provided in special education training programs. Data regarding course a) structure; b) assignments; c) test administration requirements; and d) content were collected from one hundred forty-two special education instructors. Results suggest that norm-referenced assessment is not heavily emphasized in special education programs and that the practical application of assessment skills is limited. Additionally, instructors appear to experience a shortage of support and materials compared to other assessment trainers, such as school psychologists. Implications for future research, practice and training are discussed.

Keywords: academic achievement, assessment, special education, training

The Course on Norm-referenced Academic Assessment: A Survey of Special Education Faculty

Assessment plays a critical role in special education. Assessment data are critical for intervention planning and to measure the efficacy of these interventions (Breaux & Lichtenberger, 2016). Assessment allows educators to garner a better understanding of students' capabilities and their unique strengths and areas of needed remediation (Guerriero, et al. 2020). Assessment data are also used during special education evaluations to determine whether a student has a disability and is eligible for the services and protections afforded under the Individuals with Disabilities Education Act (IDEA), 2004. While various educators, including school psychologists and educational diagnosticians, may administer the tests used in the special education eligibility process, recent research by Lockwood et al. (2021) suggests that special education teachers often fill this role as well. However, there is little research examining the training that special education teachers receive in the administration and interpretation of the tests used to make special education eligibility decisions.

Norm-referenced Academic Achievement

Assessment is the process of gathering data to make informed decisions to improve educational outcomes for students. Informal approaches including portfolio, criterion-referenced, and curriculum-based assessments are frequently used by teachers to test students' understanding of instruction and mastery of class content. Formal measures are also widely used in education to compare a student's performance to others of the same grade or age or to compare the student to a specific set of standards. One of the most frequently used types of formal assessment tools are norm-referenced tests (NRTs; Guerriero, et al., 2020). In contrast to informal measures, which are often created by teachers and can be administered as the teacher

sees fit, formal assessments are standardized, meaning they must be administered and scored by a highly trained professional following a strict set of rules. Failing to adhere to these qualifications may result in invalid test scores and interpretations (Breaux & Lichtenberger, 2016). Even when used by experts, such as school psychologists, administration and scoring errors on NRTs are common (see Styck & Walsh, 2016).

The academic achievement of students is of great concern in schools and academic NRTs are frequently used to measure progress, to compare students to others, and to evaluate their progress toward mastery of curriculum content. With regard to special education eligibility evaluations, academic NRTs are intended to measure skill acquisition or fluency corresponding to the seven areas of a specific learning disability (SLD; e.g., oral expression, listening comprehension, written expression, basic reading skill, reading comprehension, mathematics calculation, and mathematics reasoning) though other areas such as social studies or science may also be measured by certain tests (e.g., WJ IV ACH; Mather & Wendling, 2015). NRTs are the most commonly used types academic achievement measures used in determining a SLD (Benson et al., 2020; Maki et al., 2015), which is important as a third of all special education students are served under this label (National Center for Education Statistics, 2020).

Teachers and Tests of Academic Achievement

Until recently, little was known regarding the roles of teachers in NRT administration. This is despite the Council for Exceptional Children's (CEC) insistence that special education teachers be experts in collecting, interpreting, and explaining assessment data (McLeskey, et al., 2017). However, a recent survey of 725 school administrators by Lockwood and colleagues (2021) across the United States found that while school psychologists are the most likely professionals to administer NRTs, special education teachers filled this role in approximately

30% of district evaluations for special education eligibility, and general education teachers filled this role less than 1% of the time. This trend is especially pronounced in Western states where special educators were the primary administrators of NRTs of academic achievement in nearly two-thirds of districts. This study found that the most frequently used measures across assessment personnel included the Kaufman Test of Educational Achievement, Third Edition (KTEA-3; Kaufman & Kaufman, 2014), the Woodcock-Johnson IV Tests of Achievement (WJ-IV ACH; Schrank et al., 2014) and the Wechsler Individual Achievement Test, Third Edition (WIAT-III; Wechsler, 2009).

Unlike the field of school psychology where many studies have examined the fidelity of practitioners (see Styck & Walsh's, 2016 meta-analysis), very little is known regarding the fidelity of special educators in NRT administration. To date, only two published studies have examined the topic. Harrison et al. (2018) examined errors by special education teacher trainees across the Kaufman Test of Educational Achievement, Second Edition (KTEA-II; Kaufman & Kaufman, 2004), WIAT-III, and Woodcock-Johnson III Tests of Achievement (WJ-III ACH; Woodcock et al., 2007) and found that roughly 98% of administration materials contained at least one error. Additionally, Lockwood et al. (2019) examined special education trainees' errors on five of the six subtests of the KTEA-3 and found at least one error on 76%. Both studies determined that clerical (e.g., marking) errors constituted the most common mistakes and suggested that errors could be significantly decreased with proper training, though Lockwood and colleagues noted that no literature examining the typical training of special educators in the area of NRT exists. Additionally, we are unaware of any studies examining practicing special educators' NRT administration fidelity.

Assessment Training

Teachers

As previously noted, specifics regarding the training that special education teachers receive in the area of NRT administration is non-existent. However, research over the past 45 years suggests that the training of U.S. teachers in assessment and measurement concepts is inadequate (Goch & French, 2013; Roeder, 1973; Schafer & Lissitz, 1987; Stiggins & Conklin, 1988). For example, Roeder (1973) noted that 31% of teacher training programs required a class devoted to evaluation and only 19% required elementary education students to devote three or more credit hours to the subject. The author also provided a harsh critique stating that “most prospective elementary classroom teachers are better prepared to conduct impromptu art and music lessons than they are to evaluate pupil performance” (p. 141). Fifteen years later Stiggins and Conklin (1988) noted that, while teachers spent as much as a third of their professional time engaging in assessment related activities, over half of teacher training programs provide no training in assessment. Furthermore, no states certification laws required training in this area. Most recently, Greenberg and Walsh (2012) noted that while NRT data are used nationwide by school districts, teachers are generally concerned with criterion-referenced assessments designed to determine if students have mastered specific academic content. Furthermore, their study suggested that only 21% of programs provide coursework that adequately train teachers in assessment literacy (i.e., the different types of standardized and classroom assessment as well as the nomenclature of assessment). They also suggested that less than 1% of programs adequately prepare teacher trainees to describe and display test data. Additionally, only 3% of programs provided an adequate understanding of how to use assessment data for instructional decision making (Greenberg & Walsh, 2012). It is also noteworthy none of these studies had a specific focus on NRT or special education.

School Psychology Training

While little is known regarding the training of special education teachers in NRT, the training of school psychologists in NRTs of cognitive ability is fairly extensive and has remained largely consistent for roughly 35 years (Alfonso et al., 2000; Lockwood & Farmer, 2020; Oakland & Zimmerman, 1986; Miller et al., 2020). The most recent examination of this topic by Lockwood and Farmer (2020) surveyed 127 school psychology faculty members and focused solely on the first course on cognitive NRTs. Results indicated that a great deal of focus and resources are provided to instructors of this class. Instructors generally held a doctoral degree (93%) and had a teaching assistant dedicated to this class (74%). Classes averaged approximately 11 students who dedicated about eight hours a week outside of class to the course. Eighty percent of the class lectures and activities were devoted to cognitive NRTs, and emphasized topics including evidence-based practice and assessing culturally/linguistically diverse (CLD) students. Ninety-nine percent of instructors required trainees to administer tests to a k-12 student and 97% required written reports of these test results. Trainees conducted an average of approximately eight administrations. Roughly 80% of instructors required video recordings of these administrations and 98% completed protocol reviews; 44% indicated that in-person observations are used. Students were generally required to find their own children to test (84%) though 98% of programs provide kits to trainees. Additionally, prerequisite courses often included psychometrics (49%) and statistics (35%). Furthermore, roughly half of programs reported having two or more classes dedicated to cognitive assessment alone (not including classes focused on social emotional, academic achievement or other assessment topics). In sum, a great deal of time is devoted to NRT administration in these programs (Lockwood & Farmer, 2020). School psychologist's training may serve as a point of reference for special education training

with regard to assessment. While useful, this reference is not perfect; school psychologists engage in more high stakes testing (e.g., for special education placement, diagnosis in non-school settings) than teachers, and school psychologist's time is primarily spent completing testing and evaluation-related tasks (Goforth et al., 2021). As such, we would not expect teachers to have as much coursework or practicum experience devoted to testing than school psychologists.

Purpose

Special education teachers frequently administer the NRTs used to make high stakes decisions such as special education eligibility. Furthermore, little is known regarding instruction that these educators receive in NRT administration. The purpose of this study was to fill this gap in the literature by examining the instruction provided to teacher trainees by preparation programs in the U.S. To achieve this goal, a survey based on Lockwood and Farmer's (2020) questionnaire was administered to answer the following research questions:

1. What instruction do special education trainees receive in the area of NRT?
2. To what extent is instruction in NRT administration similar to or different than the practices reported in Lockwood and Farmer's (2020) study?

Method

Approval to conduct this research was obtained from the lead author's institutional review board prior to data collection. All participants provided informed consent electronically. To protect participants' confidentiality no names were gathered and email addresses collected for the participation raffle were garnered separately from the survey responses. Eysenbach's (2004) guidelines for electronic survey research reporting were used to report the methods and results. We report our recruitment methods, all data exclusions, all manipulations, and all measures used

in the study:

Participants

Participants consisted of special education instructors at 142 different teacher training programs. Programs were largely (82%) traditional institutes of higher education that were evenly distributed across the United States and represented 42 states and the District of Columbia. Participant and program data can be found in Tables 1 and 2. Participants were largely White (82%), women (78%) who were an average age of 52 years old. Most (78%) participants held a doctorate degree and reported full-time faculty (63%) status. The most common degree type held was special education (86%). Fifty-four percent reported having a teaching certificate and 39% reported having preK-12 teaching experience.

Recruitment of Participants

A list of all four-year and graduate programs with degree or certificate completers from all 50 states and the District of Columbia for the 2016-2017 school year was provided to the lead researcher by the U.S. Department of Education's data and research department (Title II Research Department, personal communication, August 16, 2019). Duplicate programs (i.e., programs that were listed as both traditional and alternative institutes of higher education [IHE]) were combined resulting in 911 total programs. A Google search was completed to find the email addresses of the department directors/chairs of each program. Directors were contacted during the month of April 2020 using the Tailor Design Method (Dillman et al., 2014) to maximize response rate. Directors/chairs were asked to forward a recruitment message with the survey link embedded to the faculty member(s) responsible for teaching their program's course on individual NRT at their institution. Two follow-up emails were sent. Participants were provided with the option to enter a drawing for one of two \$100 gift cards. A total of 178 respondents clicked on

the survey link. One participant disagreed to provide consent, 21 responded “no” to the question, “Do you teach a class on norm-referenced assessment for K-12 special education eligibility to special education or dual general education/special education teacher trainees?”, which served as an inclusionary criterion. One clicked “yes” but did not complete any additional items. Four other participants completed only 12% of the survey. All of the above respondents were deleted leaving a total of 142 faculty participants for a response rate of 15.6%.

Survey Design

The survey questionnaire was based on Lockwood and Farmer’s (2019) survey of school psychology trainee’s assessment training and designed to gather data regarding: a) the structure of the course; b) course assignments; c) test administration requirements; d) course content; and e) demographics. Items were altered to reflect the roles of special education assessment based on current research (e.g., Harrison et al., 2019; Lockwood and Farmer, 2019) and textbooks (e.g., Overton, 2016; Salvia et al., 2012). The survey was then piloted and reviewed by four content experts who examined the survey for (a) clarity, (b) precision, and (c) accuracy. All four content experts are doctoral-level university faculty members. Three of these experts have experience teaching NRT administration to special education trainees and one teaching academic achievement assessment to school psychology trainees; all have peer-reviewed publications in the area of NRT administration. Items were revised based on the feedback of these content experts. Additionally, the study authors completed the survey prior to distribution to ensure the survey platform functioned accurately, that responses were coded correctly by the digital platform, and that the data were curated in a usable format. The items were administered using Qualtrics, an online survey software platform. The survey consisted of 49 items. Item formats included one text entry, one slider, six matrix tables, and 41 multiple choice questions. Skip logic

was used to minimize completion time ($mdn = 15.0$ minutes). The survey and raw data are available open access at (link blinded for review) and in supplemental materials.

Analysis

Data were downloaded from Qualtrics on May 27, 2020 and analyzed using SPSS v. 27. As this research is exploratory in nature, only descriptive statistics were calculated and reported.

Results

Class Structure and Student Enrollment

Participants indicated that the total credit hours of classes that focus on NRT is most commonly three (53%), followed by six (20%) and four (10%) credits. Twelve percent indicated that more than six, 4% less than three and 1.5% reported five credits hour course are dedicated to NRT. Seventy-eight percent reported that their university offers sections in-person, 32% online and 33% via hybrid modalities. Of these participants, 46% relied solely on in-person instruction, 11% solely hybrid and 9% solely online instruction. An average of 16.86 ($SD = 8.53$; $mdn = 15$) trainees are enrolled in the course each time it is offered, who spend an average of 6.79 ($SD = 6.25$; $mdn = 6$) hours per week outside of class on this course. Most (89%) participants did not report teaching assistant (TA) support for this course.

Roughly 85% of participants indicated that this class is required for students working towards their degree in special education, 49% for dual general/special education students, 8% for general education students; one participant reported that the class is optional. The majority of participants indicated that students take this course during their junior (59%) or senior (42%) year, or as part of their graduate studies (51%); approximately 26% indicated that this course is solely offered to graduate students. When asked to report the amount of time that lapses between course completion and the application of class content (e.g., student teaching) 33% reported

concurrent application while 35% reported one, two (47%), three (30%), and four (8%) semesters. Generally, the only pre/co-requisite was Introduction to Exceptional Learners (or equivalent), which 82% of participants endorsed. Instructor permission was required by 14% of participants; 11% indicated that a psychometrics course and 8% reported that a statistics course are required as pre/co-requisites.

Course Topics and Content

Participants reported that an average of 46% ($SD = 23.9$, $mdn = 41\%$) of their class hours are devoted to NRT administration, scoring and interpretation (vs. curriculum-based measures (CBM), etc.) while only 39% indicated that their program requires more than one course that dedicates greater than 50% of time to NRT use ($m = 1.45$, $SD = .94$, $mdn = 1$). Data related to course content emphases are presented in Table 3. Several areas of substantial emphasis (as defined by a rating of “great emphasis” by greater than 50% of participants) were noted including human development (grades 3-6), assessment of SLD, culturally responsive practice, and use of CBM vs. norm-referenced measures. Conversely, topics including pre-K, human development, gifted assessment, assessment of infants and toddlers, and assessment of multiple and severe disabilities were rarely emphasized (as indicated by $> 50\%$ of participants indicating “limited emphasis” or “not covered”). Various textbooks were used in this course including *Assessing Learners with Special Needs* (36%), *Assessment in Special and Inclusive Education* (26%), *The ABCs of CBM* (17%), *Assessment in Special Education: A Practical Approach* (8%), a book from the *Essential* series (6%), *Assessing Students with Special Needs* (5%), *Assessment of Children & Youth with Special Needs* (5%), and *Classroom Assessment: Principals and Practices* (4%); 6% indicated that some other text is used though there was no overlap amongst these. Journals were also frequently used to cover topics including general assessment issues

(65%), CLD practices (47%), legal/ethical issues (44%), evidence-based assessment (37%), SLD (37%), psychometrics (18%), and behavioral management during testing (15%). Additionally, 9% of participants indicated some other topic (e.g., “CBM” and “response to intervention”).

Course Activities

Sixty-six percent of participants indicated that trainees are required to administer NRTs to k-12 students, while 69% are required to practice on classmates, and 16% use a simulated administration. Twelve percent indicated that students do not administer tests using any of these modalities. Of the trainees that are required to administer to a k-12 student, 49% are responsible for locating these examinees, 36% obtain these examinees from their school practicum site, 5% are obtained by the instructors, 5% a mix of the instructor and trainees, 1% from an affiliated site, and the remaining 3% from some “other” source.

Trainees are observed administering NRTs in person (43%), via video (34%) or via audio (6%) recordings; roughly 27% of participants indicated that trainees are not observed administering NRTs and 12% do not review protocols for errors (which corresponds to the 12% that do not require any form of administration). Test kits are generally (73%) provided by the college department though 11% are borrowed from their school practicum sites, 7% obtain kits some other way (e.g., “from the library,” “a mix of all of these”); 9% of students do not access a test kit at all, which suggests that 3% of those students who do not administer a test still review test kits. Thirty-one percent indicated that their trainees primarily observe others administer assessments via video, 23% in person, and 21% both in person and by video. Sixteen percent reported that their trainees are not provided with any modeling and 9% indicated that some “other” (e.g., “class demonstration,”) form of modeling occurs.

Participants indicated that test administration, protocol reviews and observations are not

the only course activities used. Ninety percent indicated that their students are required to write a report and 89% noted that the instructor provides a specific template for this purpose. Sixty-five percent require a class presentation, 60% utilize quizzes, 54% a final exam, 42% a journal article review, and 27% some “other” assignment (e.g., “case studies,” “discussion board posts,” and “eligibility role playing”). Furthermore, 76% of participants indicated that trainees are taught to calculate derived scores and 53% reported that students are taught to use scoring software.

Tests Used During Coursework

Sixty-six percent indicated that trainees administer the WJ-IV ACH; 47% the KTEA-3; 41% the KeyMath-3; 38% the WIAT-3; 36% the Test of Written Language, Fourth Edition (TOWL-4); 28% the WRMT-III; and 11% reported that trainees administer an outdated version of one of these measures. Participants indicated that trainees conduct an average of 2.31 ($SD = 3.27$, $mdn = 2$) practice administrations (e.g., with a classmate or teaching assistant) and an average of 2.33 ($SD = 3.38$, $mdn = 2$) administrations with a k-12 student. See Table 4 for more information regarding the reported tests used.

Instructors indicated that several factors were of great consideration when selecting a specific test to be taught in their class. One reason appeared to be test availability (68%). However, other considerations included whether the test is widely used regionally (67%) and nationally (60%), expectations of field placements (53%), to provide trainees with knowledge of diverse measures (52%) and validity across diverse populations (50%). See Table 5 for more details regarding test selection considerations.

Discussion

Special education teachers are frequently called upon to administer the NRTs used to determine eligibility for special education services. These measures are difficult to administer

with fidelity and administration and scoring errors are ubiquitous. However, little is known about the training that is provided to special education teachers in the area of NRT administration. This is in contrast with the field of school psychology, which has examined the training of practitioners in use of NRTs for decades. The aim of this study was to examine current training practices in special education, with training practices in school psychology used as a point of reference and contrast these findings with the most recent study of school psychologists conducted by Lockwood and Farmer (2020).

Similarities and Differences

A few similarities in the training of school psychologist and special education teachers in NRTs emerged. First, we should establish that the survey conducted by Lockwood and Farmer (2020) specifically asked school psychology faculty about the cognitive assessment course whereas we are interested in achievement assessment courses. This distinction means that some aspects of the courses would not be expected to generalize (e.g., focus on interpretation strategies with cognitive tests; focus on intellectual disability) to the participants in this study. However, there are several key similarities. Both require a significant time commitment from students outside of class and are generally taught by doctoral level practitioners with degrees in their respective fields. Both courses tend to be three credit hours. Instructors of both classes generally require report writing and provide templates to do so. Additionally, both courses emphasize assessment of school age students and issues related to CLD learners.

A number of significant differences between the current training of school psychologists and special education teachers were noted. While school psychology programs have traditionally provided instruction almost exclusively in person, less than half of special education NRT courses do so solely using this modality. Special education assessment trainees are much less

likely to have prerequisites knowledge of psychometrics and statistics than their school psychology counterparts and are much more likely to be undergraduate students. Additionally, almost half of participants indicated that more than a semester passes between completion of this class and fieldwork, compared to less than one in five school psychology instructors.

School psychology programs also appear to focus more on experiences with children and adolescents. While virtually all school psychology assessment courses report requiring their trainees to administer NRTs to k-12 students, only about two-thirds of special education assessment instructors do. Special education instructors also require about one-fourth the number of NRT administrations in this class than their school psychology counterparts. Greater than one in ten instructors require no administration practice at all. Additionally, while courses from school psychology and education programs generally engage in report writing, special education trainees appear to write reports using “mock” data far more often than their school psychology counterparts.

One striking difference is the amount of resources that are provided in the school psychology course vs. special education course. For example, while three quarters of school psychology instructors reported having a TA assigned to this course, only roughly one in 10 special education instructors had any support. This is despite special education instructors reporting class sizes that are one and a half times larger than their school psychology counterparts. Special education instructors also appear to have less access to tests and were less likely to provide kits, requiring students to obtain them from school districts far more often than school psychology courses do. Additionally, roughly one in ten special education instructors indicated using an outdated version of a NRT. This disparity in resources is also evident in the justification that special education instructors gave for choosing tests, indicating that test

availability was the number one reason why they chose to teach a specific measure. This is in contrast to school psychology instructors who indicate that test use in their course is most heavily influence by issues such as field expectations, psychometrics and validity for use with CLD students. These differences may originate with how members of the professions differently value various responsibilities, and the costs and benefits of engaging in testing for each profession. As such, while we see this comparison as useful for providing a point of reference, we do not view training provided in school psychology programs as a goal for special education programs or as a criterion of quality.

Limitations and Future Research

One limitation of this study involves the use of survey methods, which rely on retroactive self-report. Due to this, participants may have been influenced by a normative bias, social desirability, may have recalled information inaccurately, or a combination of these response patterns. However, as the survey was administered anonymously to current instructors, we believe normative bias and social desirability to be minimal. Another possible limitation includes our relatively low response rate of 16%, which may have introduced non-response bias (Dillman et al. 2014). To address these limitations, future research may wish to increase response rates and examine permanent products from the special education such as course rubrics and syllabi (see Miller et al., 2020 for an example in another field). Nevertheless, as this is the only study on this topic, we believe it provides valuable insight for future research and implications for teacher training programs. Additionally, it is of note that multiple survey researchers (e.g., Massey & Tourangeau, 2013; Perneger et al., 2005; Pike, 2012) have suggested that there is little difference between responders and non-responders and therefore low response rates likely have a minimum influence on conclusions derived from surveys (Perneger et al., 2005). Most notably a recent

simulation study by Fosnacht et al. (2017) found that response rates as low as 5%-10% provided similar survey estimates to response rates of 50% or greater as long as the survey data were collected from sample frames greater than 500 (our sample frame was 911).

Another limitation of our study involves the comparisons we made to research in school psychology, most notably the research of Lockwood and Farmer (2020). These authors examined NRTs of cognitive ability and we focused on NRTs of academic achievement, which are not directly comparable. However, due to the myriad similarities between of NRTs of academics and cognitive ability in practice and publication, and the congruence of administration and scoring errors observed on NRTs of academic and cognitive ability (Harrison et al. 2018; Lockwood et al., 2019) we believe that the comparison is fair (if not 100% compatible). Additionally, because school psychologists receive training in multiple areas of assessment, this comparison likely underestimates the discrepancy of training provided to special education teachers vs. school psychologists. To further understand training, future research may wish to examine the entirety of special education graduates' assessment training (e.g., instruction across all courses, fieldwork, etc.).

The final limitation of this study concerns the current COVID-19 pandemic, which has changed the way that instruction is provided. Most notably, universities moved classes entirely online during spring of 2020, a trend that has largely continued to the present. Additionally, the use of online instruction may persist after COVID-19 has abated (Gallagher & Palmer, 2020). Participants in this study were instructed that “(a)s times are unusual with COVID-19, please answer these questions to reflect how you would normally (not currently) conduct this course,” as we wanted to get an understanding of how the majority of current teachers, who attended college prior to COVID-19, were trained. Future research examining the changes in assessment

training during and after COVID-19 would appear necessary to determine how this public health crisis has affected teacher preparation.

Implications

There are multiple implications for training that can be garnered from our data. First, it would appear that more experience administering NRTs to k-12 students may benefit trainees as they develop mastery of NRTs. Recent research on cognitive assessment suggests that most school psychology trainees require six or more practices to develop NRT administration proficiency (Cormier et al., 2019) and may provide a good rule of thumb for special educator NRT class instructors. Second, all trainees should be recorded via video so that they can self-reflect on their performance and so instructors can provide individualized, corrective feedback and praise (Mrazik et al., 2012). Third, it appears necessary to provide all students with video recordings of testing procedures so that they can learn through modeling. As many classes are administered online, providing access to videotaped administrations may be a good approach as it allows remote (but secure) access to modeling, and for the student to watch administrations repeatedly. Fourth, up-to-date NRT kits should be used. Tests are revised either when (a) test content or language becomes outdated or (b) when re-norming of the test provides a better comparison sample due to changes in curriculum, in the representation of the sample, or so forth. As such, using outdated tests may impact the validity of test score interpretations (AERA et al., 2014). Additionally, NRTs are often altered when they are updated, and it is necessary for students to practice using the most current versions that they will be using in the school. These implications appear especially important in Western universities, as special education teachers in Western states are the primary administrators of academic NRTs (Lockwood et al., 2021).

In order to make these changes, more support for instructors would appear necessary as

special education assessment instructors are asked to do a lot with limited resources. Like their school psychology trainer counterparts, it would appear beneficial to reduce class sizes to around 10 students and provide each trainer with an experienced, dedicated graduate level TA. Additionally, programs must have access to up-to-date NRTs. Though we did not ask why older versions are used it appears logical that instructors are using outdated measures because they do not have access to more recent test kits. To ensure this support is provided, the CEC (2020) may wish to require that more program support and resources are provided to assessment courses in order to meet accreditation standards. As a point of comparison, accredited school psychology programs must have a ratio of no more than 12 trainees to every full-time faculty member and must provide adequate resources that support instruction (e.g., test kits; NASP, 2010). Similar requirements would appear appropriate for academic achievement assessment classes, though we acknowledge that budgetary constraints may not allow for this. One possible, partial solution that instructors may wish to explore is to apply for test publisher programs that provide free test materials (including kits) to training programs.

There are multiple implications for practice as well. As only a third of participants indicated that trainees take the course while student teaching, most special educators experience a gap between when the NRTs are learned and when they will administer them with any regularity. Due to the complicated and novel nature of NRTs and because drift is common (Gilmore & Campbell, 2009), special education teachers who have not administered a NRT for a term or longer should refamiliarize themselves with test materials (e.g., manual, kit, protocols, scoring software) to ensure administration fidelity. Additionally, many of the NRTs that are commonly used in school districts (e.g., KTEA-3, WIAT-3; Lockwood et al., 2021) are not taught by the majority of instructors. Therefore, it may also be necessary for special education

teachers to seek out the mentorship of an experienced administrator (e.g., master teacher, school psychologist) for help understanding the idiosyncrasies of any novel NRT that they administer. Relatedly, because only two thirds of trainees practice administering a NRT to a k-12 student during this class, it is recommended that any teacher who has not had this opportunity ask an experienced educator (e.g., master teacher, school psychologist) to observe them administer their first few NRTs and provide feedback regarding their performance. Furthermore, research (i.e., Cormier et al., 2019) suggests that six or more practice administrations are required to develop proficiency with NRT administration; because our participants reported that special education trainees administer, on average two NRTs during this course, special education trainees may be entering the field without having mastered this skill. For this reason, it may be necessary to provide in-service trainings to new special education teachers on NRT administration. Additionally, future research may wish to examine the frequency of administration and scoring errors on NRTs by practicing special education teachers to determine if proficiency has been obtained.

Conclusion

Assessment plays a critical role in special education eligibility, determining students' unique strengths and needs, and to aid in intervention implementation. Assessment data is frequently obtained through the use of NRTs and special educators frequently administer, score, and interpret these NRTs. However, our research suggests that the majority of training programs offer three or less credits dedicated to assessment and even in these courses, NRTs do not comprise the sole content of the course. Additionally, large class sizes and limited support were found. Special education programs are encouraged to provide more support to faculty who teach this course and consider increasing their emphasis on this topic so that trainees are prepared to

meet the demands of their field.

References

- Alfonso, V. C., LaRocca, R., Oakland, T. D., & Spanakos, A. (2000). The course on individual cognitive assessment. *School Psychology Review, 29*. 52-64.
<http://dx.doi.org/10.1080/02796015.2000.12085997>
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (2014). *Standards for educational and psychological testing*. AERA.
<https://www.apa.org/science/programs/testing/standards>
- Benson, N.F., Maki, K.E., Floyd, R.G., Eckert, T.L., Kranzler, J.H. & Fefer, S.A. (2020). A national survey of school psychologists' practices in identifying specific learning disabilities. *School Psychology*. <https://doi.org/10.1037/spq0000344>
- Breaux, K. C., & Lichtenberger, E. O. (2016). *Essentials of KTEA-3 and WIAT-III assessment*. John Wiley & Sons.
- Cormier, D. C., Van Norman, E. R., Cheong, C., Kennedy, K. E., Bulut, O., & Mrazik, M. (2019). Developing proficiency in standardized cognitive assessment scoring: How much is enough?. *Canadian Journal of School Psychology, 34*(3), 215-233.
<http://dx.doi.org/10.1177/0829573518765010>
- Council for Exceptional Children. (2020, July). *Initial practice-based professional preparation standards for special educators*. <https://exceptionalchildren.org/standards/initial-practice-based-professional-preparation-standards-special-educators>
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, mail, and mixed-mode surveys: The tailored design method*. (4th ed.). Wiley & Sons.
- Eysenbach, G. (2004). Improving the quality of web surveys: The Checklist for Reporting

- Results of Internet E-Surveys (CHERRIES). *Journal of Medical Internet Research*, 6(3), e34. <https://doi.org/10.2196/jmir.6.6.e34>
- Fosnacht, K., Sarraf, S., Howe, E., & Peck, L. K. (2017). How important are high response rates for college surveys?. *The Review of Higher Education*, 40(2), 245-265. <https://doi:10.1353/rhe.2017.0003>
- Gallagher, S. & Palmer, J. (2020). *The pandemic pushed universities online. The change was long overdue*. Harvard Business Review. <https://hbr.org/2020/09/the-pandemic-pushed-universities-online-the-change-was-long-overdue>
- Gilmore, L., & Campbell, M. (2009). Competence in intelligence testing: A training model for postgraduate psychology students. *The Educational and Developmental Psychologist*, 26(2), 165-173. <https://doi.org/10.1375/aedp.26.2.165>
- Goforth, A. N., Farmer, R. L., Kim, S. Y., Affrunti, N., Naser, S. C., & Lockwood, A. B. (2021, February). *School psychology: A national perspective from the 2020 membership survey*. Virtual Presentation at the National Association of School Psychologists 2021 Annual Convention, Virtual.
- Gotch, C. M., & French, B. F. (2013). Elementary teachers' knowledge and self-efficacy for measurement concepts. *The Teacher Educator*, 48(1), 46-57. <http://dx.doi.org/10.1080/08878730.2012.740150>
- Greenberg, J., & Walsh, K. (2012). What teacher preparation programs teach about k-12 assessment: A review. *National Council on Teacher Quality*.
- Guerrriero, Houser, & McGinley (2020). *The special educator's guide to assessment: A comprehensive overview by IDEA disability category*. Sage.
- Hammill, D. D., Larsen, S. C. (2009). *Test of written language* (4th ed.). Pro-Ed.

- Harrison, G. L., Goegan, L. D., & Macoun, S. J. (2018). Common examiner scoring errors on academic achievement measures. *Canadian Journal of School Psychology, 34*(2), 98-112.
<https://doi.org/10.1177/0829573518763484>
- Individuals with Disabilities Education Improvement Act (2004).
<https://www.congress.gov/105/plaws/publ17/PLAW-105publ17.pdf>
- Kaufman, A. S., & Kaufman, N. L. (2014). *Kaufman Test of Educational Achievement, third edition (KTEA-3)*. Pearson.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman test of educational achievement, second edition (KTEA-II)*. Pearson.
- Lockwood, A., Farmer, R., Bohan, K., Winans, S., & Sealander, K. (2021). Academic achievement test use and assessment practices: A national survey of special education administrators. *Journal of Psychoeducational Assessment*. Advance online publication.
<https://doi.org/10.1177/0734282920984290>
- Lockwood, A., & Farmer, R. (2020). The course on cognitive assessment: Two decades later. *Psychology in the Schools, 57*(2), 265-283. doi.org/10.1002/pits.22298
- Lockwood, A., Sealander, K., Gross, T.J., & Lanterman, C. (2019). Teacher trainees' administration and scoring errors on the Kaufman Test of Educational Achievement. *Journal of Psychoeducational Assessment, 38*(5), 551-563.
doi.org/10.1177/0734282919871144
- Maki, K. E., Floyd, R. G., & Roberson, T. (2015). State learning disability eligibility criteria: A comprehensive review. *School Psychology Quarterly, 30*(4), 457-469.
<http://dx.doi.org/10.1037/spq0000109>
- Massey, D. S., & Tourangeau, R. (2013). Where do we go from here? Nonresponse and social

- measurement. *The ANNALS of the American Academy of Political and Social Science*, 645(1), 222–236. <https://doi.org/10.1177/0002716212464191>
- Mather, N., & Wendling, B. J. (2015). *Essentials of WJ IV tests of achievement*. John Wiley & Sons.
- McLeskey, J., Barringer, M-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., Lewis, T., Maheady, L., Rodriguez, J., Scheeler, M. C., Winn, J., & Ziegler, D. (2017, January). *High-leverage practices in special education*. Council for Exceptional Children & CEEDAR Center.
- Miller, L.T., Bumpus, E.C. & Graves, S.L. (2020). The state of cognitive assessment training in school psychology: An analysis of syllabi. *Contemporary School Psychology*. <https://doi.org/10.1007/s40688-020-00305-w>
- Mrazik, M., Janzen, T. M., Dombrowski, S. C., Barford, S. W., & Krawchuk, L. L. (2012). Administration and scoring errors of graduate students learning the WISC-IV: Issues and controversies. *Canadian Journal of School Psychology*, 27, 279-290. <http://dx.doi.org/10.1177/0829573512454106>
- National Association of School Psychologists. (2010). *Standards for graduate preparation of school psychologists*. National Association of School Psychologists.
- National Center for Education Statistics. (2020, May). *The condition of education at a glance*. <https://nces.ed.gov/programs/coe/>
- Oakland, T. D., & Zimmerman, S. A. (1986). The course on individual mental assessment: A national survey of course instructors. *Professional School Psychology*, 1, 51-59. <https://doi.org/10.1037/h0090497>
- Overton, T. (2016). *Assessing learners with special needs: An applied approach*. (8th ed.).

Prentice Hall.

Perneger, T. V., Chamot, E., & Bovier, P. A. (2005). Nonresponse bias in a survey of patient perceptions of hospital care. *Medical Care*, 43(4), 374–380. <https://doi.org/10.1097/01.mlr.0000156856.36901.40>

Pike, G. R. (2012). NSSE benchmarks and institutional outcomes: A note on the importance of considering the intended uses of a measure in validity studies. *Research in Higher Education*, 54(2), 149–170.

Roeder, H. H. (1973). Teacher education curricula—Your final grade is F. *Journal of Educational Measurement*, 10(2), 141–143. <http://dx.doi.org/10.1111/j.1745-3984.1973.tb00791.x>

Salvia, J., Ysseldyke, J., & Witmer, S. (2012). *Assessment: In special and inclusive education*. Cengage Learning.

Schafer, W. D., & Lissitz, R. W. (1987). Measurement training for school personnel recommendations and reality. *Journal of Teacher Education*, 38(3), 57–63. <http://dx.doi.org/10.1177/002248718703800312>

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock-Johnson IV Tests of Cognitive Abilities*. Riverside.

Stiggins, R. J., & Conklin, N. F. (1988). Teacher training in assessment. *Northwest Regional Education Laboratory*.

Styck, K. M., & Walsh, S. M. (2016). Evaluating the prevalence and impact of examiner errors on the Wechsler scales of intelligence: A meta-analysis. *Psychological Assessment*, 28(1), 3. <http://dx.doi.org/10.1037/pas0000157>

Wechsler, D. (2009). *Wechsler individual achievement test* (3rd ed.). Pearson.

Woodcock, R. W., Schrank, F. A., Mather, N., & McGrew, K. S. (2007). *Woodcock-Johnson III*

tests of achievement. Riverside.

Table 1*Participant Data*

	<i>%</i>	<i>n</i>
Degree		
Doctoral (Ed.D., Ph.D., or Psy.D.)	77.5%	110
Masters (M.A., M.Ed., M.S.)	54.9%	78
Educational Specialist	21.8%	31
Certifications & Licensures		
Certified Teacher (prek-12)	53.5%	76
Licensed/Certified School Psychologist	12.0%	17
Licensed Psychologist	4.9%	7
Other Qualifications		
Public or Private School Experience	54.9%	78
Pre-k-12 Teaching Experience	39.4%	56
Faculty Status		
Full-time Faculty	62.7%	89
Part-time Faculty	12.0%	17

Note. $n = 142$. Participants were instructed to “mark all that apply” and therefore frequencies will not sum to 100%.

Table 2*Represented Program Data*

	<i>%</i>	<i>n</i>
Department Type		
Special Education	67.6%	96
General Education	15.5%	22
Educational Psychology	2.1%	3
Other	10.6%	15
No Response	4.2%	6
Program Type		
Traditional IHE	77.5%	110
Traditional & Alternative IHE	13.4%	19
Alternative IHE	3.5%	5
Alternative non-IHE	0.7%	1
No Response	4.9%	7
Terminal Degrees Offered*		
Bachelor's	76.8%	109
Master's	78.9%	112
Educational Specialist	25.4%	36
Doctorate	32.4%	46
Teaching Certificate	59.2%	84
Other	7.7%	11

Note. IHE = institute of higher education. *Participants could select multiple terminal degrees offered and therefore frequencies will not sum to 100%.

Table 3*Emphasis Placed on Topics Featured in Academic Assessment Coursework*

Topic/Subtopic	Percentage of Sample Indicating Amount of Emphasis Placed on Course Topics				
	Great Emphasis	Moderate Emphasis	Limited Emphasis	Not Covered	Mean (SD) ^a
Human Development					
Pre-k	13.0%	23.7%	39.7%	23.7%	2.74 (0.97)
Grades k-2	49.2%	41.7%	6.1%	3.0%	1.63 (0.74)
Grades 3-6	59.5%	33.6%	4.6%	2.3%	1.50 (0.70)
Jr. High/middle school	35.6%	46.2%	7.6%	10.6%	1.93 (0.93)
High school	29.0%	43.5%	14.5%	13.0%	2.11 (0.97)
Assessing Diverse Populations					
Culturally and/or linguistically diverse	45.2%	45.9%	8.1%	0.7%	1.64 (0.66)
Emotional/behavioral disturbance	29.1%	53.0%	15.7%	2.2%	1.91 (0.73)
Giftedness	3.0%	23.1%	47.0%	26.9%	2.98 (0.79)
Infants and preschoolers	10.4%	11.9%	48.5%	29.1%	2.96 (0.91)
Intellectual and developmental disabilities	44.0%	39.6%	12.7%	3.7%	1.76 (0.82)
Multiple and severe disabilities	14.9%	31.3%	40.3%	13.4%	2.52 (0.91)
Specific learning disabilities	68.4%	24.8%	3.8%	3.0%	1.41 (0.71)
Current Issues Related to Achievement Assessment					
Culturally responsive practice	57.5%	38.1%	4.5%	0.0%	1.47 (0.58)
Effect of labeling in education	50.7%	37.3%	11.9%	0.0%	1.61 (0.69)
Error analysis	35.8%	46.3%	16.4%	1.5%	1.84 (0.75)
Item analysis	25.4%	53.7%	19.4%	1.5%	1.97 (0.71)
Quantitative v. qualitative assessment	29.9%	52.2%	13.4%	4.5%	1.93 (0.78)
Use of CBM v. norm-referenced measures	54.1%	37.8%	6.7%	1.5%	1.56 (0.69)

Note. a = Great emphasis = 1, moderate emphasis = 2, limited emphasis = 3, and no coverage = 4. Lower means indicate greater overall

emphasis. Numbers have been rounded and may not equal 100%. CBM = curriculum-based measurement

Table 4*Percentage of sample indicating whether a test was administered by students*

Test	%	n
Comprehensive Mathematical Abilities Test (CMAT)	4.9%	7
Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)	15.5%	22
Developmental Reading Assessment, Third Edition (DRA3)	18.3%	26
Dyslexia Early Screening Test, Second Edition (DEST-2)	3.5%	5
Kaufman Test of Academic Achievement, Third edition (KTEA-3)	46.5%	66
KeyMath-3 Diagnostic Assessment	40.8%	58
Process Assessment of the Learner, Second Edition: Diagnostics for Reading and Writing (PAL-II)	1.4%	2
Process Assessment of the Learner, Second Edition: Diagnostics for Math (PAL-II Math)	0.7%	1
Test of Written Language, Fourth Edition (TOWL-4)	35.9%	51
Wechsler Individual Achievement Test, Third Edition (WIAT-3)	38.0%	54
Wide Range Achievement Test, Fifth Edition (WRAT5)	15.5%	22
Woodcock-Johnson IV Test of Academic Achievement (WJ-IV ACH)	66.2%	94
Woodcock Reading Mastery Tests, Third Edition (WRMT-III)	27.5%	39
An Older Version of One of the Above	11.3%	16

Note. Participants were instructed to “mark all that apply” and therefore frequencies will not sum to 100%.

Table 5*Factors Influencing Test Selection*

Reason	Amount of Consideration Given to Factors Determining Test Selection			
	Great	Moderate	Limited	No Consideration
Availability of Test Kits	67.6%	23.7%	4.3%	4.3%
Widely Used Test (Regionally)	66.9%	23.7%	6.5%	2.9%
Widely Used Test (Nationally)	60.3%	26.2%	9.2%	4.3%
Expectations of Field Placements	53.2%	21.6%	12.9%	12.2%
So that Students Have Knowledge of Diverse Measures	51.8%	38.8%	7.2%	2.2%
Validity Across Diverse Populations	49.3%	40.7%	5.7%	4.3%
Psychometric Properties of Scores	45.3%	33.1%	15.8%	5.8%
Instructor Familiarity with Test	29.5%	36.7%	23.7%	10.1%

Note. Numbers have been rounded and may not equal 100%.