ASHA American Speech-Language-Hearing Association

LSHSS

Research Note

Online Training Modules for Teaching Assessment Skills to Graduate Student Clinicians

Katharine M. Radville,[a] Emilie C. Larrivee,[a] Lauren S. Baron,[a] Patricia Kelley-Nazzaro,[a] and Joanna A. Christodoulou[a]

[a] MGH Institute of Health Professions, Boston, MA

ARTICLE INFO

Article History:
Received May 10, 2021
Revision received August 13, 2021
Accepted December 13, 2021

Editor-in-Chief: Holly L. Storkel
Editor: Elizabeth D. Peña

https://doi.org/10.1044/2021_LSHSS-21-00068

ABSTRACT

Purpose: Assessment is a crucial skill for speech-language pathologists, who rely on standardized tests to identify characteristics of speech, language, hearing, literacy, and related skill sets. Training in assessment administration is an integral part of graduate education that lays the foundation for appropriate use of these tools. Teaching students to administer standardized assessments is time intensive, can vary depending on student learning pace, and involves both general and test-specific knowledge. The current pilot study investigated the effectiveness of researcher-developed and scalable online training modules for beginning and advanced graduate students for knowledge and self-efficacy outcomes.

Method: Graduate students (n = 61) across four cohorts (two beginning and two advanced) were trained using online, asynchronous assessment training modules. The modules addressed the administration of standardized assessments using written tutorials, video demonstrations, and quizzes. Students completed pre- and posttests on knowledge and self-efficacy for administering and scoring the trained assessments before and after the online training modules.

Results: We found significant pretest to posttest gains in specific assessment-related knowledge for beginning and advanced groups. Following completion of the training modules, beginning graduate students demonstrated increased self-efficacy for general assessment principles, whereas advanced graduate students started and remained high in self-efficacy at posttest.

Conclusions: Using researcher-developed online training modules, both beginning and advanced graduate students improved knowledge for assessment skills. These training modules can be used as a model for developing teaching materials across a range of assessments and related topics that are scalable in the context of remote teaching and learning.

Speech-language pathologists (SLPs) must be able to accurately and efficiently assess individuals in order to describe performance and identify strengths, challenges, and disorders. The American Speech-Language-Hearing Association (ASHA) names assessment as one of the eight domains of service delivery in their Scope of Practice in Speech-Language Pathology (ASHA, 2016b). Accurate and efficient administration of standardized assessment measures is an essential component in the process of differential diagnosis, which generally requires the use of formal, normed measures. Standardized (or normed) assessments allow SLPs to compare individuals to their age- or grade-matched peers, use data to inform diagnoses, describe patterns of strengths and weaknesses, inform goals for intervention, and monitor progress. Standardized assessment scores often inform eligibility for services in both medical and educational settings in high-stakes situations.

Correspondence to Joanna A. Christodoulou: jchristodoulou@mghihp.edu. Katharine M. Radville and Emilie C. Larrivee are co-first authors. **Publisher Note:** This article is part of the Forum: Can You See My Screen? Virtual Assessment in Speech and Language. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Given the importance of formal assessments, ASHA specifically cites competence in their use as an essential function for SLPs (ASHA, 2016b).

SLPs hold an ethical obligation to be competent in the area of assessment. Clinicians are required to participate only in professional activities that are consistent with their training and certification status (ASHA, 2016a). Likewise, student clinicians can only be assigned responsibilities that are consistent with their training (ASHA, 2016a). Unsurprisingly, ASHA (2008) delineates assisting supervisees with using assessment tools and techniques that are specific to their clients as a required aspect of clinical supervision. As such, graduate student clinicians require extensive training and practice in assessment.

Comprehensive assessment training is an integral part of graduate education, both during coursework and during practicum experiences. However, there are clear barriers to thorough instruction in this area. First, formal assessment tools are heterogeneous in nature and measure a wide range of skills (e.g., cognition, language, literacy, and speech production). By definition, these tests are standardized in terms of administration, scoring, and interpretation procedures. Experienced clinicians are expected to determine important components of test use (e.g., start points, basal, ceiling, discontinue rule, appropriate feedback, timing, and materials needed) to ensure accuracy, validity, and reliability. Graduate student clinicians often lack sufficient clinical experience to effectively navigate information in assessment manuals, warranting direct instruction in how to administer standardized assessments correctly. Teaching graduate students to administer standardized assessments is time intensive and requires repeated practice and considerable attention to detail. Traditional training methods require a favorable supervisor–student ratio; challenges in achieving such a ratio can make scalability poor for instructing many students at the same time.

More than 300 communication sciences and disorders (CSD) departments across the United States train over 20,000 speech-language pathology graduate students each year (Council of Academic Programs in CSD & ASHA, 2021). Each department's faculty are responsible for ensuring that students have adequate clinical skills in assessment administration, scoring, and interpretation. Research on teaching and learning in CSD has indicated that beginning clinicians benefit from participation in experiential learning activities (Walden & Gordon-Pershey, 2013) that include scaffolded practice with basic skills (Austin, 2013). More specifically, clinical learning is supported by the use of clinical scripts and protocols such that novice clinicians can complete clinical tasks accurately (Peña & Kiran, 2008). Use of video examples can also facilitate acquisition of clinical skills (Ferguson & Estis, 2018; Peña & Kiran, 2008).

Limited information is currently available related to effective methods for teaching assessment. Lohmander et al. (2020) demonstrated that an asynchronous, computer-based training was effective in improving graduate students' skills for auditory perceptual assessment when rating speech samples from individuals with cleft palate. Ferguson and Estis (2018) similarly demonstrated that brief, online video simulations were both cost effective and improved graduate students' assessment skills related to feeding safety in preterm infants. In a small pilot study of two participants, Irizarry-Pérez et al. (2021) demonstrated that procedural training and practice with peers was effective for training bilingual, graduate student clinicians to administer a Spanish–English oral language screening measure. In a study of audiology and speech-language pathology undergraduate students, Alanazi et al. (2020) examined the effects of simulation training sessions as compared to video-based case scenarios and found that both methods led to improved knowledge and confidence. They concluded that video simulations are a feasible approach for clinical learning. Although others are undoubtedly using video tutorials in their clinical training approaches, to the best of our knowledge, there is a lack of published research on systematic, scalable, and effective training approaches for teaching speech-language pathology graduate students to use a range of standardized assessment measures.

Several factors have elevated the importance of exploring digital contexts for assessment. These factors include (a) the shift to increased use of tele-assessment as a result of COVID-19m, (b) the growth of distance/online graduate programs for SLPs, and (c) varied resources/challenges that are part of online versus in-person learning and assessment. Here, we focus on digital training opportunities to supplement in-person graduate student training. Tele-assessment should be considered not just as the use of tests but also as the practicality of comfort, agility, and experience with online learning platforms to use tests as well. With the onset of the COVID-19 global pandemic, many faculty responsible for training graduate student clinicians were required to pivot to online teaching practices, highlighting an additional advantage of such practices for scalable, consistent, and reliable approaches to supporting graduate student trainees in the digital era. Given the rapid shift to online learning and tele-assessment, we believe that there is the need to provide graduate student clinicians with a framework for standardized test administration that can be used for assessments administered in person or remotely.

An effective structure for teaching and learning that can support student clinician training of assessments is Universal Design for Learning (UDL; CAST, 2018). UDL principles endorse multiple opportunities for engagement, representation, and action/expression. For example, providing training that is self-paced and asynchronous

allows for tailored learning engagement. Using multimedia to deliver information (e.g., presentations and video demonstrations) can support dynamic opportunities for representation of content. Offering examples of both what testers should and should not do during administration can support transfer and generalization. Emphasizing mastery-oriented feedback (and using a learn–test–learn–test format) encourages learning material to a high level of both knowledge and comfort that can help student clinicians earn the identity of a competent user of standardized tests.

Self-efficacy, or an individual's perception of and confidence in their ability to perform a specific task (Bandura, 1977), is associated with ability to effectively use knowledge and skills (Bandura, 1997). Within the field of speech-language pathology, higher self-efficacy has been associated with increased confidence in clinical skills (Pasupathy & Bogschutz, 2013), as well as knowledge and skills (Boyer, 2013; Cassidy, 2013). Self-efficacy scales have also been used to assess graduate students' perceptions of readiness for clinical placements and clinical fellowship experiences (Oswalt, 2013) and to assess the effectiveness of graduate student training programs (Victorino & Hinkle, 2019). Given their value for clinical learning and for informing clinical teaching, we used self-efficacy scales (described below) to evaluate graduate students' confidence levels related to assessment before and after participation in the online assessment training modules.

Research Questions and Hypotheses

In this pilot study, we describe training methods that were initially developed to train advanced graduate students to complete pre- and posttesting for a research study. As such, the battery of tests used with the advanced graduate student clinicians included a range of standardized assessments that included both timed and untimed measures and required different response modalities. This presented us with the opportunity to assess the effectiveness of these self-paced, asynchronous, scalable online modules. Although the modules did not provide training in direct tele-assessment, they are specifically relevant to situations in which instructors and graduate student clinicians are not or cannot be in physical proximity to each other. The full training could be conducted without any in-person interaction and could be used to teach students to administer assessments either in-person or via telehealth. Subsequent to using the modules for research study training, we used a subset of the modules to train beginning graduate student clinicians for their clinical practice. We were interested in knowing whether this type of clinical tool learning would be effective for speech-language pathology graduate students. We evaluated the effectiveness of the training modules through quantifying its impacts on graduate student knowledge and self-efficacy. In their clinical training, beginning graduate students then completed one of the training templates to learn additional standardized assessments, though the results of this aspect of clinical learning were not evaluated as a part of this study.

We aimed to evaluate the efficacy of self-paced, online assessment training modules for CSD graduate students. Our first research question was: Do beginning and advanced graduate students show improved assessment-related knowledge after completing the online assessment training modules? We hypothesized that assessment-related knowledge scores would increase from pre- to posttest for both groups. Our second research question was: Do beginning and advanced graduate students report improved self-efficacy after completing the online assessment training modules? We hypothesized that self-efficacy scores would increase for both groups following the trainings. Our third research question was: Is change in assessment knowledge associated with change in assessment self-efficacy from pre- to posttest? We hypothesized a positive association between these variables.

Method

Participants

Graduate student clinicians (N = 61) enrolled in a speech-language pathology training program participated in the online training modules to learn assessments for the purposes of research (advanced students) or as part of their first-year clinical experience (beginning students). The participants were in their first, second, fourth, or sixth semester out of six semesters of the same CSD graduate program. The mean age of participants was 26.1 years old (SD = 4.28). By participant report, the majority were female (female: n = 56, male: n = 5). Race/ethnicity responses were based on census categories and were reported as follows (in order of frequency): 40 White, nine Asian, six Hispanic, three other, two multiple, and one Black/African American. The first- and second-semester students were combined into a single group and classified as "beginning students" (n = 53), and the fourth- and sixth-semester students were combined into a single group and classified as "advanced students" (n = 8). This research was approved with exempt status by the Partners Human Research Committee Institutional Review Board.

All of the participants completed the training procedures between June 2019 and March 2020. They were trained in using assessments that were relevant to their clinical or research-related responsibilities. Therefore, although the nature of the online training modules was consistent across participants, the specific assessments that

participants learned to administer varied between the beginning and advanced students. Although the students had a limited amount of time to complete the entire training, the students were not required to complete each aspect of the training in a specific amount of time, allowing for each task to be paced based on a student's individual needs. The beginning students completed the training modules (including all quizzes) within an average of 9.50 days (SD = 4.25). The advanced students completed the training modules (including all quizzes) within an average of 5.81 days (SD = 4.33), based on when the students first and last accessed the surveys and quizzes. This information was obtained through the time stamps, which were automatically collected by the Google Forms platform and exported along with accuracy data.

Materials and Procedure

The online modules were used to train students in the administration of literacy, language, and cognition subtests from standardized assessments. The training for all students included literacy measures; the training for advanced students also included assessments of oral language and cognition. See Table 1 for the key features of

Table 1. Trained assessments.

Construct	Test	Subtest	Timed or untimed	Response modality
Nonverbal cognitive ability	Kaufman Brief Intelligence Test–Second Edition (KBIT-2)	Matrices (KBIT-2 M)	Untimed	Pointing
Nonverbal processing speed	Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V)	Coding (WISC-V C)	Timed	Drawing symbols
Verbal memory	Wide Range Assessment of Memory and Learning–Second Edition (WRAML-2)	Sentence Memory (WRAML-2 SM)	Untimed	Verbal response (repeating sentences)
Phonological awareness	Comprehensive Test of Phonological Processing–Second Edition (CTOPP-2)	Elision (CTOPP-2 EL)	Untimed	Verbal response (manipulating speech sounds to form a word)
Phonological memory	Comprehensive Test of Phonological Processing–Second Edition (CTOPP-2)	Memory for Digits (CTOPP-2 MD)	Untimed	Verbal response (repeating strings of numbers)
Rapid automatic naming	Rapid Automatized Naming and Rapid Alternating Stimulus Tests (RAN/RAS)	RAN Letters (RAN/RAS L)	Timed	Verbal response (rapidly identify a subset of letters)
Letter identification	Woodcock Reading Mastery Tests–Third Edition (WRMT-III)	Letter Identification (WRMT-III LI)	Untimed	Verbal response (letter names)
Word reading	Woodcock Reading Mastery Tests–Third Edition (WRMT-III)	Word Identification (WRMT-III WI)	Untimed	Verbal response (reading real words)
Decoding	Woodcock Reading Mastery Tests–Third Edition (WRMT-III)	Word Attack (WRMT-III WA)	Untimed	Verbal response (reading nonwords)
Text comprehension	Woodcock Reading Mastery Tests–Third Edition (WRMT-III)	Passage Comprehension (WRMT-III PC)	Untimed	Verbal response (single-word cloze procedure)
Automaticity	Test of Word Reading Efficiency–Second Edition (TOWRE-2)	Sight Word Efficiency (TOWRE-2 SWE)	Timed	Verbal response (reading real words rapidly)
Automaticity	Test of Word Reading Efficiency–2 (TOWRE-2)	Phonemic Decoding Efficiency (TOWRE-2 PDE)	Timed	Verbal response (reading nonwords rapidly)

Note. KBIT-2 M = Kaufman Brief Intelligence Test–Second Edition Matrices (Kaufman & Kaufman, 2004); WISC-V C = Wechsler Intelligence Scale for Children–Fifth Edition Coding (Wechsler, 2014); WRAML-2 SM = Wide Range Assessment of Memory and Learning–Second Edition Sentence Memory (Sheslow & Adams, 2013); CTOPP-2 EL = Comprehensive Test of Phonological Processing–Second Edition Elision (Wagner et al., 2013); CTOPP-2 MD = Comprehensive Test of Phonological Processing–Second Edition Memory for Digits (Wagner et al., 2013); RAN/RAS L = Rapid Automatized Naming and Rapid Alternating Stimulus Tests Letters (Wolf & Denckla, 2005); WRMT-III LI = Woodcock Reading Mastery Tests–Third Edition Letter Identification (Woodcock, 2011); WRMT-III WI = Woodcock Reading Mastery Tests–Third Edition Word Identification (Woodcock, 2011); WRMT-III WA = Woodcock Reading Mastery Tests–Third Edition Word Attack (Woodcock, 2011); WRMT-III PC = Woodcock Reading Mastery Tests–Third Edition Passage Comprehension (Woodcock, 2011); TOWRE-2 SWE = Test of Word Reading Efficiency–Second Edition Sight Word Efficiency (Torgeson et al., 2012); TOWRE-2 PDE = Test of Word Reading Efficiency–Second Edition Phonemic Decoding Efficiency (Torgeson et al., 2012).

each assessment, including construct, presence of a timed condition, and response modality.

The online training modules included the same elements for each assessment and included components of UDL. These elements included a slideshow overview of the measure and an overview document outlining specific administration requirements (i.e., materials, scoring procedures, starting points by age or grade, basals and ceilings, reversal and discontinue rules, and feedback rules). This information was presented in a consistent format to help the graduate students through consistent patterns and develop a framework for learning additional standardized measures in the future. In addition to these written materials, students watched two brief (2- to 4-min) demonstration videos per assessment. The first video demonstrated correct test administration, and the second video depicted inaccurate administration. For each video of inaccurate assessment, graduate students identified common testing errors. Materials were provided to all students through the use of a virtual platform (excluding test manuals, protocols, and stimulus books), which was organized by test. Although the training modules themselves were fully online, students had access to traditional, paper-based test manuals, protocols, and stimulus books. Access to physical testing materials was offered to ensure that students were able to practice with each measure in the manner in which they would each eventually administer the tests (in person). Lastly, the graduate student clinicians were assigned to practice test administration with one another. All students were required to complete a knowledge quiz on each assessment to mastery. Students received their scores immediately, and an e-mail was automatically sent following their participation that indicated what their responses were and whether their response was correct or incorrect. Participants tracked and reported their progress toward completing the training on a spreadsheet to ensure full, consistent participation. Each participant completed all required aspects of the training with the goal of mastery. Before and after the online training (but before in-person training components), participants completed the same knowledge quizzes and self-efficacy scales. Students were asked to confirm that they would complete the pre and post knowledge quizzes without consulting another person, training materials, or notes. Refer to Figures 1 and 2 for depictions of the training procedures for each cohort of participants. Although not part of this study, fidelity of the test administration and scoring that followed the training were directly monitored by licensed SLPs to ensure accuracy.

Knowledge Quizzes

To assess the effectiveness of the online training modules for improving assessment-related knowledge, participants completed the same knowledge quizzes (see Appendix A) before and after participating in the training.

Figure 1. Advanced students training method.

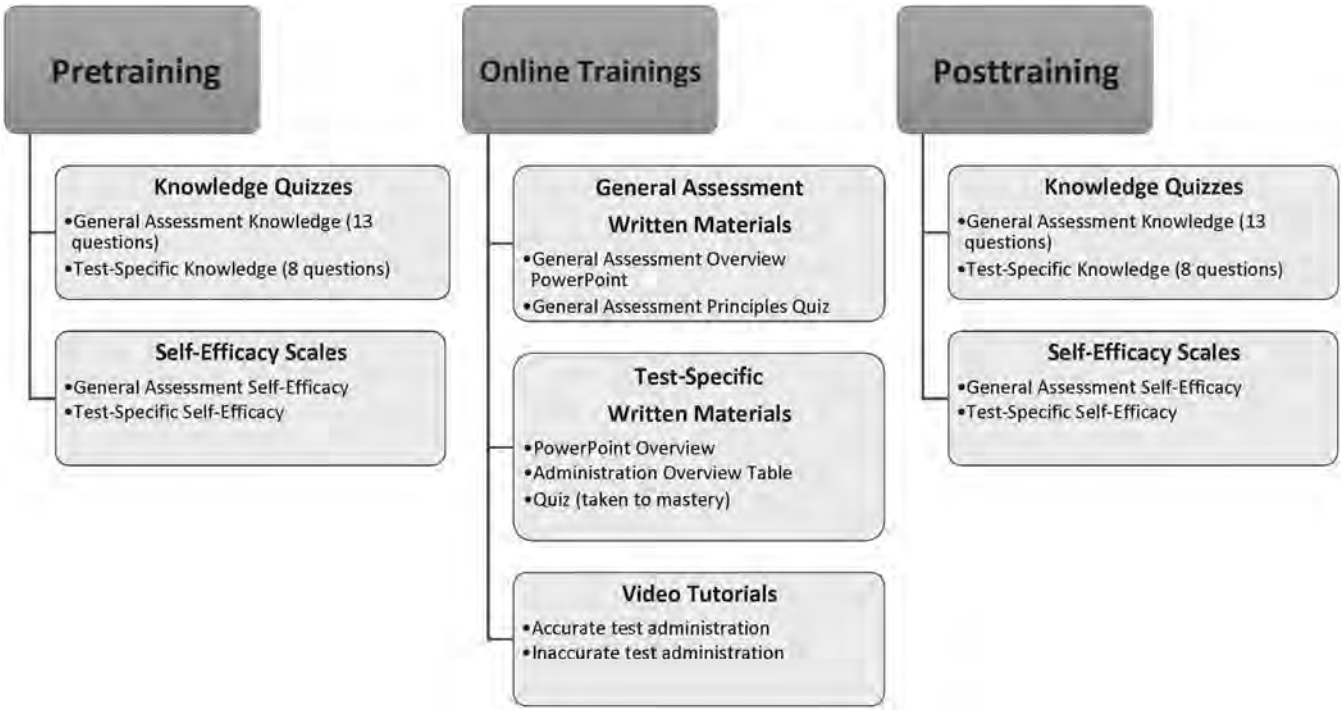
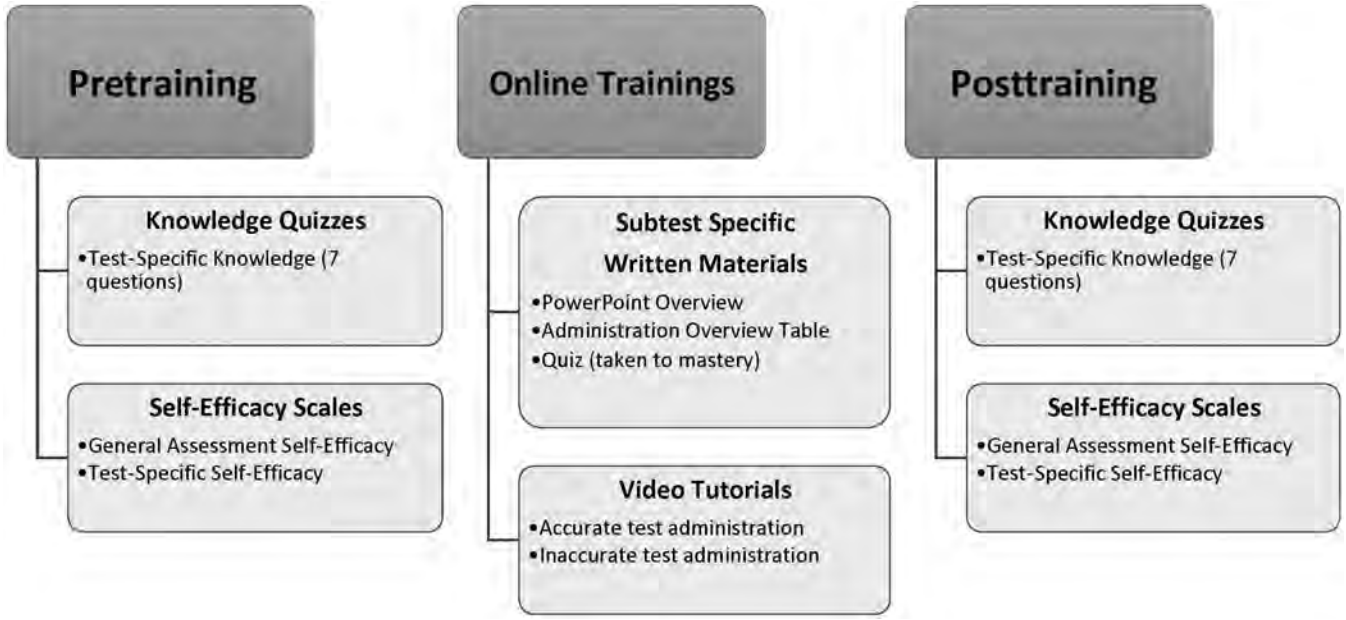


Figure 2. Beginning students training method.



Advanced graduate students completed a 13-question quiz that evaluated their baseline knowledge of general test administration principles and an eight-question quiz that assessed their knowledge of how to administer specific assessments.

Beginning graduate students completed a seven-question quiz that evaluated their knowledge of specific assessment measures but did not complete the general assessment knowledge quiz. After completing the training but prior to any in-person instruction, all participants took the same knowledge quiz a second time (i.e., post-test). This timeline helped to ensure that student learning was not bolstered by the in-person feedback from the researcher or licensed clinician.

Self-Efficacy Scales

Prior to participating in any aspect of the online training, all participants rated their self-efficacy for general, assessment-related skills (see Appendix B). For this purpose, we constructed a questionnaire that used a 4-point Likert scale (*not yet confident*, *minimally confident*, *moderately confident*, and *very confident*). For advanced students, the survey included 10 questions that assessed confidence for general assessment principles: establishing rapport, maintaining client engagement, behavior management, providing feedback, administering assessments accurately, collecting data, modifying timing/pacing, implementation of supervisor feedback, and ethical accountability during assessment. Therefore, advanced students' self-efficacy scores were based on a total of 40 possible points. Beginning students completed a similar, nine-question survey that assessed confidence for each of the same principles with the exception of incorporation of supervisor feedback. Therefore, beginning students' self-efficacy scores were based on a total of 36 possible points. A second self-efficacy scale asked all participants (beginning and advanced) to rate their confidence for each specific assessment measure that they were assigned to learn via the online training modules. These assessment-specific self-efficacy scales used the same 4-point Likert scale. As with the knowledge quizzes, all participants took each self-efficacy scale a second time after completing the online training modules, but prior to receiving any in-person instruction.

Analysis

To evaluate the efficacy of self-paced, online assessment training modules for CSD graduate students, we collected data related to assessment knowledge and self-efficacy before and after students participated in the online training. Quiz results are reported as items correct out of the total number of survey questions as a decimal in analyses or as a percentage in the text. Analyses were conducted separately for beginning and advanced students due to uneven samples sizes and graduate training differences between the two groups. To address whether beginning and advanced graduate students showed improved assessment-related knowledge after completing the online assessment training modules (Research Question 1), we conducted paired-samples t tests comparing mean general assessment knowledge in advanced students before and after the training modules. We conducted paired-samples t tests comparing

each cohort's assessment-specific knowledge before and after participation in the training; all reported p values (text and tables) reflect use of Bonferroni corrections.

To examine whether beginning and advanced graduate students reported improved self-efficacy after completing the online assessment training modules (Research Question 2), we conducted paired-samples t tests to compare mean self-efficacy before and after participation in the training for both the beginning and advanced students. Also for both groups, we conducted paired-samples t tests to examine differences in confidence for specific assessment measures before and after participation in the training. We chose to use paired-samples t tests instead of an omnibus test due to our goal to evaluate change in knowledge and self-efficacy for each assessment measure separately. We used Hedges' g to measure effect size for all comparisons. We corrected Hedges' g for analyses for advanced students given the small sample size. We interpreted effect sizes using the standard metric of 0.2 as a small effect size, 0.5 as a medium effect, and 0.8 as a large effect (Cohen, 1977). Lastly, we conducted Pearson correlations to analyze the relationship between change in assessment knowledge and change in assessment-related self-efficacy for both beginning and advanced students (Research Question 3).

Results

Knowledge scores before and after the training for beginning and advanced student cohorts yielded favorable outcomes. Figure 3 shows percentage accuracy with error bars representing standard deviations by group for pre and post. Beginning students' test-specific knowledge scores (seven-question quiz) were significantly lower before (M = 18.60%, SD = 24.36%) as compared to after (M = 85.71%, SD = 19.21%) participation in the training, t = −15.38, p < .0001, g = 3.06. (Beginning students did not participate in the general assessment knowledge quiz.) Similarly, advanced students demonstrated statistically significant increases in general knowledge (13-question quiz) and test-specific knowledge (eight-question quiz) following the training. Their general knowledge scores were significantly lower before (M = 78.33%, SD = 10.54%) as compared to after (M = 95.00%, SD = 9.92%) participation in the training, t = −3.54, p = .02, g = 1.23. Advanced students' test-specific knowledge was also significantly lower before (M = 28.13%, SD = 21.91%) as compared to after (M = 85.94%, SD = 12.38%), t = −6.33, p < .001, g = 4.25.

We evaluated change in self-efficacy for both cohorts following participation in the training. Beginning students' general self-efficacy (out of 36 possible points) was significantly lower before (M = 21.11, SD = 6.17) as compared to after (M = 26.21, SD = 5.16) participation in the training, t = −6.78, p < .0001, g = 0.90. Beginning students also demonstrated significantly increased confidence for the administration of each of three trained assessment measures (see Table 2). Advanced graduate students did not demonstrate increased general self-efficacy, though their confidence (out of 40 possible points) was initially high (M = 29.88, SD = 4.55) and remained so following the training (M = 31.00, SD = 4.34), t = −1.39, p = 1.00, g = 0.19. However, the advanced graduate students demonstrated significantly increased confidence for all 12 trained assessments (see Table 3).

We evaluated the relationship between change in assessment knowledge and change in assessment-related

Figure 3. Advanced and beginning students' knowledge scores before and after the training.

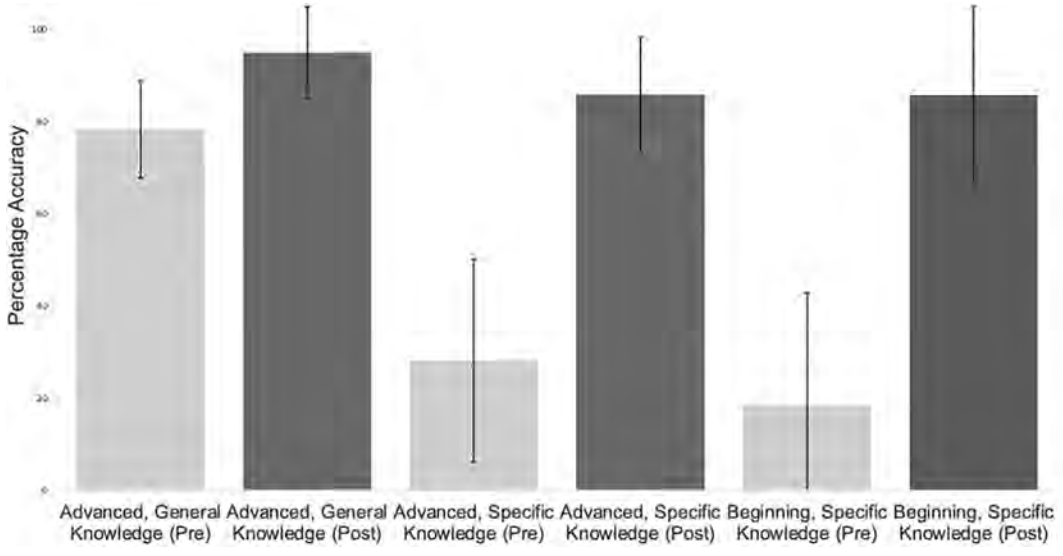


Table 2. Assessment-specific self-efficacy: beginning students.

Measure	*n*	*M* (pre)	*SD* (pre)	*M* (post)	*SD* (post)	95% CI (post–pre)	*t*	*p*	*g*
WRMT-III WI	53	1.30	0.54	2.74	0.66	[–1.67, –1.21]	–12.72	< .0001	2.40
WRMT-III WA	53	1.32	0.55	2.68	0.64	[–1.60, –1.12]	–11.54	< .0001	2.28
WRMT-III PC	53	1.28	0.53	2.74	0.71	[–1.70, –1.21]	–11.89	< .0001	2.32

Note. Confidence was self-reported using a 4-point scale. Reported *p* values include Bonferroni corrections for multiple (three) analyses. CI = confidence interval; WRMT-III WI = Woodcock Reading Mastery Tests–Third Edition Word Identification; WRMT-III WA = Woodcock Reading Mastery Tests–Third Edition Word Attack; WRMT-III PC = Woodcock Reading Mastery Tests–Third Edition Passage Comprehension.

self-efficacy for each group. For beginning students, change in test-specific knowledge (measured on a single survey for the Word Identification, Word Attack, Passage Comprehension subtests of the Woodcock Reading Mastery Tests–Third Edition) was significantly and positively correlated with change in general self-efficacy ($r = .29$, $p = .04$). For advanced students, change in self-efficacy (which was minimal) was not significantly correlated with change in general assessment knowledge ($r = -.48$, $p = .22$) or change in test-specific knowledge ($r = .67$, $p = .07$).

Discussion

Training graduate students in the core competency of using formal assessments is deemed an essential skill for SLPs by ASHA (2016b). With high-stakes decisions such as medical care or access to special education services based in part on standardized assessments, it is essential that speech-language pathology graduate students have access to test training that fosters accuracy, reliability, and validity. To alleviate the time-intensive demands of assessment training for faculty and foster mastery of core concepts for students, digital training tools can be a valuable resource (Alanazi et al., 2020; Ferguson & Estis, 2018; Peña & Kiran, 2008; Walden & Gordon-Pershey, 2013). We offer a pilot study of digital training modules for CSD and related disciplines that integrates self-efficacy and principles from UDL.

This pilot study aimed to address the need for scalable, systematic assessment training methods for beginning and advanced speech-language pathology graduate students by evaluating the effectiveness of online training modules for improving assessment-related knowledge and self-efficacy. The approach aimed to provide individual

Table 3. Assessment-specific self-efficacy: advanced students.

Measure	*n*	*M* (pre)	*SD* (pre)	*M* (post)	*SD* (post)	95% CI (post–pre)	*t*	*p*	*g*
KBIT-2 M	8	1.00	0.00	2.63	1.06	[–2.51, –0.74]	–4.33	.04	1.64
WISC-V C	8	1.00	0.00	3.25	0.46	[–2.64, –1.86]	–13.75	.001	5.18
WRMT-III LI	8	1.50	0.54	3.38	0.52	[–2.17, –1.58]	–15.00	.001	2.69
WRMT-III WI	8	1.88	0.64	3.38	0.52	[–1.95, –1.05]	–7.94	.001	1.94
WRMT-III WA	8	2.00	0.54	3.25	0.71	[–1.84, –0.66]	–5.00	.02	1.44
WRMT-III PC	8	1.88	0.64	3.38	0.52	[–1.95, –1.05]	–7.94	.001	1.94
TOWRE-2 SWE	8	2.00	0.76	3.38	0.52	[–2.14, –0.61]	–4.25	.05	1.60
TOWRE-2 PDE	8	1.75	0.89	3.38	0.52	[–2.39, –0.86]	–5.02	.02	1.69
CTOPP-2 EL	8	1.88	0.64	3.63	0.52	[–2.34, –1.16]	–7.00	.003	2.26
CTOPP-2 MD	8	1.88	0.64	3.38	0.52	[–1.95, –1.05]	–7.94	.001	2.26
WRAML-2 SM	8	1.13	0.35	3.63	0.52	[–3.12, –1.87]	–9.35	.001	4.24
RAN/RAS L	8	1.25	0.46	3.5	0.54	[–2.84, –1.66]	–9.00	.001	3.39

Note. Self-efficacy was self-reported using a 4-point scale. Reported *p* values include Bonferroni corrections for multiple (12) analyses. Hedges' *g* has been corrected due to the small sample size ($n = 8$). KBIT-2 M = Kaufman Brief Intelligence Test–Second Edition Matrices (Kaufman & Kaufman, 2004); WISC-V C = Wechsler Intelligence Scale for Children–Fifth Edition Coding (Wechsler, 2014); WRMT-III LI = Woodcock Reading Mastery Tests–Third Edition Letter Identification (Woodcock, 2011); WRMT-III WI = Woodcock Reading Mastery Tests–Third Edition Word Identification (Woodcock, 2011); WRMT-III WA = Woodcock Reading Mastery Tests–Third Edition Word Attack (Woodcock, 2011); WRMT-III PC = Woodcock Reading Mastery Tests–Third Edition Passage Comprehension (Woodcock, 2011); TOWRE-2 SWE = Test of Word Reading Efficiency–Second Edition Sight Word Efficiency (Torgeson et al., 2012); TOWRE-2 PDE = Test of Word Reading Efficiency–Second Edition Phonemic Decoding Efficiency (Torgeson et al., 2012); CTOPP-2 EL = Comprehensive Test of Phonological Processing–Second Edition Elision (Wagner et al., 2013); CTOPP-2 MD = Comprehensive Test of Phonological Processing–Second Edition Memory for Digits (Wagner et al., 2013); WRAML-2 SM = Wide Range Assessment of Memory and Learning–Second Edition Sentence Memory (Sheslow & Adams, 2013); RAN/RAS L = Rapid Automatized Naming and Rapid Alternating Stimulus Tests Letters (Wolf & Denckla, 2005).

pacing, learning at any location, and a learn–test–learn–test cycle to help students focus on important knowledge and skills through use of online quizzes to foster mastery. We hypothesized that assessment-related knowledge and self-efficacy scores would increase for both groups following the trainings. We also hypothesized that there would be a positive association between change in knowledge and change in self-efficacy. Study findings indicated that the training facilitated gains in general and specific assessment knowledge for advanced graduate students. The training also supported gains in specific assessment knowledge for beginning graduate students (we did not assess general assessment knowledge for beginning students). Effect sizes for each pre- and posttraining comparison of assessment were large based on standard interpretation of Hedges' *g* (Hedges, 1981). These findings indicate that, in addition to reaching statistical significance, the participation in the training had clinically relevant impacts on student knowledge.

The training modules also facilitated increased general self-efficacy for beginning graduate students and increased test-specific self-efficacy for both beginning and advanced graduate students. Although advanced students did not demonstrate improved general confidence, they did report increased confidence for each specific assessment measure. These findings are unsurprising given that the advanced students had participated in assessment-related instruction through previous coursework and clinical opportunities prior to participating in the training. Their increased confidence for specific measures provides promising information supporting the utility of this type of training for experienced assessors who require training for novel, unfamiliar measures. As with our results for pre- and posttraining knowledge comparisons, effect sizes were large for all self-efficacy-related analyses. This finding indicates that participation in the training had meaningful impacts on student self-efficacy.

Change in knowledge and self-efficacy were positively and significantly correlated for beginning graduate students but not for advanced graduate students. For beginning students, these results are in line with prior research demonstrating that self-efficacy is related to improved clinical knowledge and skills (Boyer, 2013; Cassidy, 2013). For advanced graduate students, generally elevated self-efficacy ratings before and after the training (i.e., ceiling effect) limited exploration of the relationship between change in general knowledge and change in self-efficacy. Overall, the current findings highlight utility of online assessment training modules for supporting beginning graduate students' learning and self-efficacy and advanced graduate students' learning. We offer that clinical graduate studies should advance and reinforce the bidirectional relation between self-efficacy and knowledge.

Standardized and scalable training has considerable potential to support graduate-level instruction in the area of assessment. It has potential benefits as an in-class teaching tool and applications for both clinical and research settings. Given its largely asynchronous nature, it allows a considerable degree of flexibility for both instructors and students in terms of timing and location. It is notable that both beginning and advanced students demonstrated improved confidence and knowledge with minimal synchronous, face-to-face time demands. As such, the training modules proved to be both a scalable and effective method for instruction of portable, transferable clinical skills. In addition to scalability, the training modules are reusable, expandable, and could be completed in parts or as a whole. The modules can be used flexibly, either during remote learning or to augment in-person learning. In addition to CSD faculty members developing these training modules, test publishers could also enhance the effective use of their products by making this material readily available.

The training modules could be examined as components to be used to address specific training needs across settings and disciplines. The utility of training modules for improving graduate students' self-efficacy and knowledge was also promising in terms of generalizability. It is also possible that this type of training tool could be useful for training clinicians in other professions (e.g., audiology, special education, clinical psychology, and occupational therapy). For example, clinical psychology training emphasizes the need for extensive training in standardized assessment administration, but limitations have been well documented regarding wide variation in teaching methods (Childs & Eyde, 2002; Clemence & Handler, 2001) and how extensively assessment-related skills are taught (Mihura et al., 2017).

Limitations

Although the findings of this study are promising and have practical implications for clinical education, there are several limitations. First, use of a control group and random assignment would increase the rigor of the study. To address this, a next step would be to include a control group and compare gains in knowledge and confidence between students who participated in the online training to students who participated in more traditional, in-person training. Additionally, although the short time between our pre- and posttraining measures (less than 10 days, on average, for both groups) makes it unlikely that participants learned about general assessment principles or test-specific assessment from sources outside the trainings (e.g., coursework or other clinical experiences), this possibility cannot be ruled out. Next, our sample size for advanced graduate students was small. Extending this investigation to a larger cohort of students would be both feasible, given the strong scalability of the training modules,

and useful in terms of demonstrating its effectiveness. Lastly, although our results indicate benefits related to general assessment knowledge, test-specific knowledge, general self-efficacy related to assessment, and confidence for administering specific assessments, the study did not directly quantify improved or accurate use of assessments following the training. Although students were supervised during assessment administration and scores were checked to ensure accuracy (as is standard when training graduate students), this was not systematically measured and was not part of this study. As such, we are not able to confirm that gains in assessment knowledge and self-efficacy directly supported improvements to clinical practice. A logical extension of this research would be to systematically and directly measure administration and scoring abilities after the trainings, such as with the use of fidelity rubrics (e.g., Irizarry-Pérez et al., 2021). This would allow for analysis of the relationship between the trainings and generalization to successful completion of clinical tasks.

Conclusions

Our findings provide novel information regarding the utility of a clinical teaching tool that is both scalable and replicable. This work is applicable to teaching graduate students to administer a range of assessment measures across constructs and response modalities. Although we used these tools to train in-person assessment, the same components could be used to support telepractitioners and/or for telepractice training. Given the importance of assessment as a clinical skill across a variety of content areas, this study has implications for instructors who teach graduate students with varied levels of experience and informs the scholarship of teaching and learning of essential knowledge and skills in speech-language pathology.

Author Contributions

Katharine M. Radville: Conceptualization (Equal), Formal analysis (Lead), Methodology (Equal), Project administration (Equal), Resources (Supporting), Writing – original draft (Equal), Writing – review & editing (Equal). **Emilie C. Larrivee:** Conceptualization (Equal), Investigation (Equal), Methodology (Equal), Project administration (Lead), Resources (Lead), Writing – original draft (Equal), Writing – review & editing (Supporting). **Lauren S. Baron:** Conceptualization (Supporting), Methodology (Supporting), Project administration (Supporting), Resources (Supporting), Supervision (Supporting), Writing – review & editing (Supporting). **Patricia Kelley-Nazzaro:** Conceptualization (Supporting), Methodology (Supporting), Project administration (Supporting), Supervision (Supporting), Writing – review & editing (Supporting). **Joanna A. Christodoulou:** Conceptualization (Equal), Methodology (Lead), Project administration (Equal), Resources (Equal), Supervision (Lead), Writing – original draft (Supporting), Writing – review & editing (Lead).

Acknowledgments

The authors thank Annie Fox for analysis consultation and the MGH Institute of Health Professions Communication Sciences and Disorders students who completed the training modules. The authors appreciate the generosity of colleagues who supported the development of online training modules with templates and content, sourced from NIH DC010784 (PI: Gray) and IES R305F100002 (PI: Justice).

References

Alanazi, A. A., Mohamud, M. S., & AlSuwailem, S. S. (2020). The effect of simulation learning on audiology and speech-language pathology students' self-confidence related to early hearing detection and intervention: A randomized experiment. *Speech, Language and Hearing, 23,* 1–14. https://doi.org/10.1080/2050571X.2020.1846839

American Speech-Language-Hearing Association. (2008). *Knowledge and skills needed by speech-language pathologists providing clinical supervision*. https://www.asha.org/policy/KS2008-00294/

American Speech-Language-Hearing Association. (2016a). *Code of ethics*. https://www.asha.org/Code-of-Ethics/

American Speech-Language-Hearing Association. (2016b). *Scope of practice in speech-language pathology*. https://www.asha.org/policy/SP2016-00343/

Austin, L. (2013). Scaffolding early clinical learning for students in communication sciences and disorders. *SIG 11 Perspectives on Administration and Supervision, 23*(3), 86–91. https://doi.org/10.1044/aas23.3.86

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Boyer, V. (2013). Graduate students working with English language learners: Impact on self-efficacy and knowledge acquisition. *SIG 10 Perspectives on Issues in Higher Education, 16*(2), 63–70. https://doi.org/10.1044/ihe16.2.63

Cassidy, C. H. (2013). The relationship between perceived supervisory roles, working alliances, and students' self-efficacy in speech-language pathology practicum experiences. *SIG 11 Perspectives on Administration and Supervision, 23*(3), 92–109. https://pubs.asha.org/doi/full/10.1044/ass23.3.92

CAST. (2018). *Universal Design for Learning Guidelines Version 2.2*. http://udlguidelines.cast.org

Childs, R. A., & Eyde, L. D. (2002). Assessment training in clinical psychology doctoral programs: What should we teach? What do we teach? *Journal of Personality Assessment, 78*(1), 130–144. https://doi.org/10.1207/S15327752JPA7801_08

Clemence, A. J., & Handler, L. (2001). Psychological assessment on internship: A survey of training directors and their

expectations for students. *Journal of Personality Assessment, 76*(1), 18–47. https://doi.org/10.1207/S15327752JPA7601_2

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Routledge.

Council of Academic Programs in Communication Sciences and Disorders & American Speech-Language-Hearing Association. (2021). *Communication sciences and disorders (CSD) education survey national aggregate data report: 2019–2020 academic year*. https://www.asha.org/siteassets/uploadedfiles/csd-education-survey-national-aggregate-data-report.pdf

Ferguson, N., & Estis, J. (2018). Training students to evaluate preterm infant feeding safety using a video-recorded patient simulation approach. *American Journal of Speech-Language Pathology, 27*(2), 566–573. https://doi.org/10.1044/2017_AJSLP-16-0107

Hedges, L. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Education Statistics, 6*(2), 107–128. https://doi.org/10.2307/1164588

Irizarry-Pérez, C. D., Lugo-Neris, M. J., Bedore, L. M., & Peña, E. D. (2021). Establishing diagnostic skills in novice bilingual clinicians: A scaffolded approach. *Teaching and Learning in Communication Sciences & Disorders, 5*(1), 1–17. https://doi.org/10.30707/TLCSD5.1.1624982519.487027

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test–Second Edition*. Pearson.

Lohmander, A., Klinto, K., Shalling, E., Portela, A. S., Johansson, K., & McAllister, A. (2020). Students take charge of learning: Using e-learning in perceptual assessment in speech-language pathology. *Scandinavian Journal of Educational Research, 65*(3), 468–480. https://doi.org/10.1080/00313831.2020.1716064

Mihura, J. L., Roy, M., & Graceffo, R. A. (2017). Psychological assessment training in clinical psychology doctoral programs. *Journal of Personality Assessment, 99*(2), 153–164. https://doi.org/10.1080/00223891.2016.1201978

Oswalt, J. (2013). Students' self-perceptions of knowledge and skills pre and post externship. *SIG 11 Perspectives on Administration and Supervision, 23*(3), 110–120. https://doi.org/10.1044/aas23.3.110

Pasupathy, B., & Bogschutz, R. (2013). An investigation of graduate speech-language pathology students' SLP clinical self-efficacy. *Contemporary Issues in Communication Sciences and Disorders, 40*(Fall), 151–159. https://doi.org/10.1044/cicsd_40_F_151

Peña, E. D., & Kiran, S. (2008). In defense of cookbooks. *Topics in Language Disorders, 28*(3), 242–258. https://doi.org/10.1097/01.TLD.0000333599.30468.9d

Sheslow, D., & Adams, W. (2013). *Wide range assessment of memory and learning*. Pro-Ed.

Torgeson, J. K., Rashotte, C. A., & Wagner, R. K. (2012). *Test of Word Reading Efficiency–Second Edition*. Pro-Ed.

Victorino, K. R., & Hinkle, M. S. (2019). The development of a self-efficacy measurement tool for counseling in speech-language pathology. *American Journal of Speech-Language Pathology, 28*(1), 108–120. https://doi.org/10.1044/2018_AJSLP-18-0012

Wagner, R. K., Torgeson, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing–Second Edition*. Pro-Ed.

Walden, P. R., & Gordon-Pershey, M. (2013). Applying adult experiential learning theory to clinical supervision: A practical guide for supervisors and supervisees. *SIG 11 Perspectives on Administration and Supervision, 23*(3), 121–144. https://doi.org/10.1044/aas23.3.121

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children–Fifth Edition*. Pearson.

Wolf, M., & Denckla, M. B. (2005). *RAN/RAS: Rapid Automatized Naming and Rapid Alternating Stimulus Tests*. Pro-Ed.

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests–Third Edition*. Pearson.

Appendix A (p. 1 of 3)

Knowledge Quizzes

General Assessment Knowledge: Advanced Graduate Student Clinicians

Answer each question on this entire quiz without looking at any notes so you can test your own knowledge.

I confirm that I will participate without consulting another person, materials, or notes.

Select the appropriate option for each blank space, respectively.

1) The basal establishes (a) _____ for the test, which (b) _____ for each test.	For blank (a): a starting point	For blank (a): baseline	For blank (b): varies	For blank (b): is the same
2) The ceiling establishes (a) _____ for the test, which (b) _____ for each test.	For blank (a): an ending point	For blank (a): a starting point	For blank (b): varies	For blank (b): is the same

Select the best option for the questions below.

3) Assessors may provide feedback to the child on each question about the correctness of answers.	True		False	
4) The item where an assessor begins test administration is called the:	Basal	Ceiling	Discontinuation rule	Starting point
5) If you are unsure of how to score a particular item, what should you do?	Make your best guess concerning the score of the item	Ask the question again at the end of the test	Provide the student with the correct response	Write exactly what the child answered, leave the score blank
6) Assessors may modify the wording and procedures on standardized tests.	True		False	
7) When removing a child for testing, what are the key pieces of information to inform the teacher with?	When you want to test the child and what test(s) will be administered	Who you would like to test, where you are taking the child, and for how long	Your name and the assessments you will be administering	Who you would like to test and why
8) It is best to use only uppercase letters when filling out assessment record forms.	True		False	
9) What is the recommended writing utensil for assessors?	Felt tip pen	#2 Pencil with 0.9-mm lead	Black or blue ballpoint pen	Red ballpoint pen
10) The audio recorder should be stopped between each measure and paused while students are reading or listening to a story.	True		False	
11) Select the one piece of information that should NEVER be included in any audio recording?	Child's ID number	Child's first and last name	Assessor name	Date of administration
12) Why is it important to follow each measure's protocol as it relates to prompting?	Tests must be given the same way every time with the goal being the same test results regardless of whom administers	Inappropriate prompting will result in not being able to use/score those responses	The use of prompts allows the assessor to get more information from the child without providing the child additional guidance or information on the question	All of the above
13) What things should an assessor do to decrease the amount of missing data?	Check to make sure that all items on the record form are marked clearly and accurately	Follow the administration and scoring protocols as written for each measure	Make sure that the date is filled in completely and accurately	All of the above

Appendix A (p. 2 of 3)

Knowledge Quizzes

Test Specific Knowledge: Advanced Graduate Student Clinicians

Answer each question on this entire quiz without looking at any notes so you can test your own knowledge.

I confirm that I will participate without consulting another person, materials, or notes.

Select the best option for the questions below.

Question	True	False	I don't know
On the KBIT-2 Matrices subtest, an examinee may indicate their answer by pointing to the picture or saying its letter. If a child offers the name of a picture (e.g., bone), prompt them to respond by pointing or saying the letter.			
On the CTOPP-2 Elision subtest, you should provide the student with the letter name (i.e., say cat, now say cat without saying "k").			
On the Memory for Digits subtest of the CTOPP-2, if examinee wants a trial repeated, you can repeat the answer once during the items where you provide feedback (Items 1–4).			
On the RAN/RAS, if the examinee asks to start over the section, they may do so 1 time, and only the second attempt is recorded.			
On the WRAML-2 Sentence Memory subtest, you can provide emphasis to parts of the sentence you believe are most important and most complex.			
On the WRMT-III, you should go above the ceiling to complete the items for the remainder of each page.			
For both subtests of the TOWRE-2, you should prompt the student by pointing if they stop at the end of a column and tell them to "go on" if they pause on a word for more than 5 s.			
On the WISC-V Coding subtest, you should have an eraser available and provide it to the student if they ask for one.			

Note. KBIT-2 = Kaufman Brief Intelligence Test–Second Edition (Kaufman & Kaufman, 2004); CTOPP-2 = Comprehensive Test of Phonological Processing–Second Edition (Wagner et al., 2013); RAN/RAS = Rapid Automatized Naming and Rapid Alternating Stimulus Tests (Wolf & Denckla, 2005); WRAML-2 =Wide Range Assessment of Memory and Learning–Second Edition (Sheslow & Adams, 2013); WRMT-III = Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); TOWRE-2 = Test of Word Reading Efficiency–Second Edition (Torgeson et al., 2012); WISC-V = Wechsler Intelligence Scale for Children–Fifth Edition (Wechsler, 2014).

General Assessment Self-Efficacy: Beginning Graduate Student Clinicians

Please provide a confidence-level rating for the following skills using these response options:	(1) Not yet confident	(2) Minimally confident	(3) Moderately confident	(4) Very confident
Establishes rapport while recognizing the needs, values, and preferred mode of communication and cultural linguistic backgrounds of clients and relevant others				
Keeps client actively engaged during assessment				
Effectively manages clients' behavior and attention				
Provides appropriate feedback based on the assessment guidelines				
Administers assessment procedures correctly				
Collects data effectively in a nonobtrusive manner				
Modifies timing/pacing of activities to effectively target session objectives				
[Raw] scores assessment correctly [based on procedures delineated in manual]				
Demonstrates accountability to ethical practice, confidentiality, and policy and procedures that protect and respect clients' interests				

Appendix A (p. 3 of 3)

Knowledge Quizzes

Assessment-Specific Self-Efficacy: Beginning Graduate Student Clinicians

Please provide a confidence-level rating for administration (including basals/ceilings, feedback, discontinue rule, etc.) and raw scoring for the following assessments:	(1) Not yet confident	(2) Minimally confident	(3) Moderately confident	(4) Very confident
WRMT-III: Word Identification				
WRMT-III: Word Attack				
WRMT-III: Passage Comprehension				

Note. WRMT-III = Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011).

Appendix B

Self-Efficacy Scales

General Assessment Self-Efficacy: Advanced Graduate Student Clinicians

Please provide a confidence-level rating for the following skills using these response options:	(1) Not yet confident	(2) Minimally confident	(3) Moderately confident	(4) Very confident
Establishes rapport while recognizing the needs, values, and preferred mode of communication and cultural linguistic backgrounds of clients and relevant others				
Keeps client actively engaged during assessment				
Effectively manages clients' behavior and attention				
Provides appropriate feedback based on the assessment guidelines				
Administers assessment procedures correctly				
Collects data effectively in a nonobtrusive manner				
Modifies timing/pacing of activities to effectively target session objectives				
[Raw] scores assessment correctly [based on procedures delineated in manual]				
Demonstrates accountability to ethical practice, confidentiality, and policy and procedures that protect and respect clients' interests				
Implements supervisor feedback accurately and effectively				

Assessment-Specific Self-Efficacy: Advanced Graduate Student Clinicians

Please provide a confidence-level rating for administration (including basals/ceilings, feedback, discontinue rule, etc.) and raw scoring for the following assessments:	(1) Not yet confident	(2) Minimally confident	(3) Moderately confident	(4) Very confident
CTOPP-2: Elision				
CTOPP-2: Memory for Digits				
KBIT-2: Matrices				
TOWRE-2: Sight Word Efficiency				
TOWRE-2: Phonemic Decoding Efficiency				
WISC-V: Coding				
WRMT-III: Letter Identification				
WRMT-III: Word Identification				
WRMT-III: Word Attack				
WRMT-III: Passage Comprehension				
WRAML-2: Sentence Memory				

Note. CTOPP-2 = Comprehensive Test of Phonological Processing–Second Edition (Wagner et al., 2013); KBIT-2 = Kaufman Brief Intelligence Test–Second Edition (Kaufman & Kaufman, 2004); TOWRE-2 = Test of Word Reading Efficiency–Second Edition (Torgeson et al., 2012); WISC-V = Wechsler Intelligence Scale for Children–Fifth Edition (Wechlser, 2014); WRMT-III = Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011); WRAML-2 = Wide Range Assessment of Memory and Learning–Second Edition (Sheslow & Adams, 2013).