



A 10-Year Retrospective Case Study of the Relation between Four Summative Assessment Measures in a Graduate Speech-Language Pathology Program

Lynne Cameron
Eastern Illinois University, Incameron2@eiu.edu

Laurel Teller
Eastern Illinois University, lteller@eiu.edu

Rebecca Throneburg
Eastern Illinois University, rmthroneburg@eiu.edu

See next page for additional authors

DOI: 10.61403/2689-6443.1304

Follow this and additional works at: <https://ir.library.illinoisstate.edu/tlcsd>

Recommended Citation

Cameron, L., Teller, L., Throneburg, R., Mulvey, N., & Fahy, J. (2024). A 10-Year Retrospective Case Study of the Relation between Four Summative Assessment Measures in a Graduate Speech-Language Pathology Program. *Teaching and Learning in Communication Sciences & Disorders*, 8(1). DOI: <https://doi.org/10.61403/2689-6443.1304>

This Scholarship of Teaching and Learning Research is brought to you for free and open access by ISU ReD: Research and eData. It has been accepted for inclusion in Teaching and Learning in Communication Sciences & Disorders by an authorized editor of ISU ReD: Research and eData. For more information, please contact ISUREd@ilstu.edu.

A 10-Year Retrospective Case Study of the Relation between Four Summative Assessment Measures in a Graduate Speech-Language Pathology Program

Abstract

Assessment is essential to ensure that quality levels of teaching and learning are maintained in graduate programs in Communication Sciences and Disorders (CSD). In this article, we present a ten-year retrospective case study of one CSD master's program approach to address speech-language pathology program-level summative assessment. We evaluated the strength of the relation between three departmental summative measures (i.e., Grand Rounds [Capstone Course] final grade percentages, Written Comprehensive examinations, Oral Comprehensive examinations) and the national Praxis Examination in Speech-Language Pathology (5331). The strongest correlations were between the Grand Rounds final grade percentages, Written Comprehensive examinations, and the Praxis. The weakest correlations were between the Oral Comprehensive examinations and the other examination types. The study findings demonstrate the concurrent validity of Grand Rounds final grade percentages, Written Comprehensive examinations, and the Praxis. Capstone courses should be considered for their benefit in Praxis preparation, whereas oral comprehensive examinations may better serve as formative rather than summative assessment.

Keywords

summative assessment, capstone, written comprehensive examination, oral comprehensive examination

Authors

Lynne Cameron, Laurel Teller, Rebecca Throneburg, Nichole Mulvey, and Jill Fahy

Introduction

Assessment in higher education is essential to ensure quality teaching and learning. Broadly, assessment includes formative assessment, which shapes student learning through continual appraisal of knowledge and skills, and summative assessment, which evaluates student learning outcomes near the conclusion of the educational experience and supports program review (Council on Academic Accreditation in Audiology and Speech-Language Pathology [CAA], 2023a). For accreditation, universities and academic programs must prove through assessment that expected standards of quality are met and maintained. Earning and preserving accreditation are crucial to the future of academic institutions and programs.

In this article, we present a ten-year retrospective case study of one communication sciences and disorders (CSD) graduate program's approach to address speech-language pathology program-level summative assessment to evaluate student learning outcomes for accreditation and program development. We describe the strength of the relation between three departmental summative measures and the national Praxis Examination in Speech-Language Pathology (5331; Praxis II).

Summative Assessment for the Council on Academic Accreditation

The American Speech-Language-Hearing Association's (ASHA) Council on Academic Accreditation in Audiology and Speech-Language Pathology (CAA) is responsible for defining and evaluating standards of accreditation in the field of CSD. CAA (2023b) standards indicate programs "must use a variety of assessment techniques, including both formative and summative methods . . . administered by multiple academic and clinical faculty members" while students are in the graduate program. CAA (2023a) describes summative assessment as, "comprehensive evaluation of learning outcomes, including acquisition of knowledge and skills, at the culmination of an education experience (e. g., course, program)." Like many accrediting bodies, CAA requires audiology and speech-language pathology programs to provide evidence of systematic summative assessment that is used consistently to support student learning and program development; however, CAA does not specify how many summative assessments must be conducted. In the past, the Praxis, which is conducted through the Educational Testing Service (ETS) and is required for licensure and certification, was accepted by the CAA as one summative measure; beginning in January of 2023, however, CAA enforced that the Praxis may no longer serve as a form of summative assessment. Rationale for this decision was based on a lack of validation of the Praxis for the purpose of program-specific summative assessment, consistent with ETS (2021). Per ETS (2021), licensure exams, such as the Praxis, are not meant to assess the breadth of the knowledge and skill in the profession, but rather only address a subset of entry-level content related to safe and effective practice generated by a group of experts and stakeholders in the field. Pass rates on the Praxis, however, remain a student outcome measure that CAA requires all CSD programs to publicly post and reflect on to guide program refinement.

Types of Summative Assessment in Higher Education

Types of summative assessments vary widely among programs. CAA (2023b) provided guidance for how programs can demonstrate compliance with summative assessment of students, including, but not limited to "end of program comprehensive evaluations; final projects, portfolios, and/or

examinations that reflect cumulative work/knowledge over time; capstone projects; final papers; final reports and/or presentations; overall knowledge and skill; rubrics; practical examinations; peer/self-evaluation.” Cunningham and colleagues (1996) surveyed 229 master’s level training programs accredited by the CAA and had a response rate of 90%. At that time, 24% of the programs required both written and oral comprehensive examinations, 49% required only written comprehensive examinations, 5% required only oral comprehensive examinations, 10% of the programs required passing the national certification exam as the only comprehensive examination tool and an additional 6.5% required passing the national certification exam as a condition of graduation in addition to other program comprehensive examination(s), 4% required only a master’s thesis or had no required comprehensive examination, and 2% required only a competency-based portfolio or some “other” form of comprehensive evaluation. The authors did not ask respondents whether the written comprehensive examinations were multiple choice or essay.

Similar illustrations of summative assessment occur across a variety of programs in higher education. Ewell and colleagues (2011) surveyed 2,719 program heads from a sample of regionally accredited, undergraduate degree-granting, two- and four-year public, private, and for-profit institutions in the United States as a follow up to the 2009 National Institute for Learning Outcomes and Assessment (NILOA) and had a response rate of 30%. They reported multiple methods of summative assessment, with the highest reported being capstone course (59%), followed by rubrics on student work (48%), culminating final project (45%), performance assessments (45%), locally developed content exams (38%), local surveys (32%), comprehensive exams (25%), portfolios (24%), standardized content exams (22%), and professional licensure exams (15%). While they identified several frequently-used options for program and summative assessment, Ewell and colleagues (2011) called for case studies at the program level to further identify and understand the variations in assessment methods and the corresponding effects.

Evidence-Based Guidance for Developing Summative Assessments in Higher Education

The task of designing summative assessments with little to no faculty or financial resources dedicated to assessment can be daunting (Emanuel et al., 2013; Ewell et al., 2011). Ewell and colleagues’ call for more program-level research is supported by inconsistent recommendations for best practices in the current assessment literature. For example, some researchers support structured essays, whereas others support multiple-choice assessments (Emanuel et al., 2013; Epstein, 2007; Ewell et al., 2011; Fallon & Emanuel, 2019; Hift, 2014). In his discussion of methods for evaluating professional competence and performance in medicine, Epstein (2007) highlighted the need for written assessment that is absent of “cueing,” such as in multiple-choice questions. Epstein further suggested that structured essays assess higher-order thinking processes better than do multiple-choice questions. Structured essays are more difficult to grade, however, and one frequent barrier in academic programs is that there is no time or money dedicated to summative assessments, and therefore no one to grade these time-consuming examinations (Emanuel et al., 2013; Ewell et al., 2011; Fallon & Emanuel, 2019). Palmer and Devitt (2007) reported significant reliability concerns between faculty in more than half of essay questions scored in a summative medical assessment. Hift (2014) evaluated summative assessment in undergraduate and post-graduate medical education and argued that open-ended summative assessments are not superior to multiple-choice summative assessments in validity or in the

evaluation of higher-order thinking. The author gave excellent examples of context rich multiple-choice items where all choices could be plausible but required analysis, insight, synthesis, and many other higher-order thinking processes to utilize information and successfully draw conclusions to answer the questions accurately. Hift concluded that a well-constructed context-rich multiple-choice item represents a complex problem-solving exercise which activates a sequence of cognitive processes that closely parallel those required in clinical practice. He stated that the evidence does not support the proposition that the open-ended assessment format is superior to the multiple-choice format, at least in exit-level summative assessment, in terms of either its ability to evaluate higher-order cognitive functioning or its validity. Hift even suggested that given the superior reliability and cost-effectiveness of the multiple-choice format, consideration should be given to phasing out open-ended format questions in summative assessment. Fallon and Emanuel (2019) reported positive results in a multiple-choice exam for summative assessment in a graduate-level speech-language pathology program. Fallon and Emanuel modified their program's written comprehensive examination from a mixed format (e.g., essay, short answer, multiple choice) across eight clinical topic areas presented in random order, to a multiple-choice format across ten topic areas presented by topic. Following modifications, Fallon and Emanuel reported scores from the revised multiple-choice examination that were more reflective of Praxis performance, as well as improved student satisfaction and reduced administrative burden.

Written Comprehensive Examination and Preparation for License Examination

The primary goal of CSD programs is to prepare students to become competent and effective speech-language pathologists. CSD programs also hope their academic and clinical education gives students the skills necessary to pass summative assessments related to the program's learning objectives, as well as the Praxis, an examination used for licensure and certification in speech-language pathology. Several studies in the healthcare-related literature found trends, correlations, or predictivity between written comprehensive and licensure examination performance, suggesting well-designed written summative assessment may predict and/or advance performance on licensure examinations.

As mentioned previously, Fallon and Emanuel (2019) discovered that modification of their written comprehensive examination to a topic-organized, multiple-choice format, with questions developed by teaching faculty, better reflected students' skills with an 86% pass rate for the new comprehensive examination versus a 45-56% pass rate on the previous comprehensive examination, and 100% first time pass rate for the Praxis. The final version of their assessment included 150 questions, with 15 questions across the following content areas: language development and disorders birth to 5 years; language development and disorders for school-age children; speech-sound disorders in childhood; fluency; dysphagia; cognitive linguistics and aphasia; cognitive linguistics and dementia, traumatic brain injury; adult neuromotor speech; autism spectrum disorders; and augmentative and alternative communication.

Similarities between written comprehensive examination and licensure examination performance are found across other healthcare-related disciplines as well. Kosmahl (2015) reported that physical therapy student performance on a comprehensive examination correlated to National

Physical Therapy Examination scores ($r = .617$; $p < .001$). The comprehensive exam consisted of approximately 100 case-based multiple-choice questions.

In the nursing literature, the Health Education Systems Incorporated (HESI) Exit Exam, often used as a summative assessment in nursing programs, was found to consistently predict National Council Licensure Examination (NCLEX) performance with greater than 98% accuracy (Langford & Young, 2013; Nibert et al., 2005; Shah et al., 2022). The HESI Exit Exam consists of 150 questions, including fill in the blank, multiple choice, drag and drop, and multiple response. Like the NCLEX Exam, scores on the HESI Exit Exam are weighted based on difficulty.

In their systematic review of predictors of success on the North American Pharmacist Licensure Examination (NAPLEX), Park and colleagues (2021) found grade point average and the Pharmacy Curriculum Outcomes Assessment (PCOA) predicted NAPLEX performance. The PCOA consists of 225 multiple-choice questions across four content areas. This format is like the NAPLEX, which also consists of 225 multiple-choice questions. Rowe and Hamilton (2021) discussed similar findings in their study of the relation between an in-house comprehensive examination, PCOA, and NAPLEX. They found correlations between all exams including between the in-house examination and the NAPLEX scaled scores ($r = 0.568$, $p < .001$).

From the above findings, it appears that requiring students to pass a written comprehensive examination that is like the licensure examination, as part of program summative assessment may inform students of areas of strengths and weakness to direct their studies in preparation for the licensure examination. Epstein (2007) suggested that content expectations of summative assessments alone may positively affect student learning because students tend to study more diligently the areas on which they will be tested. Heeneman and colleagues (2015) echoed these findings through student interviews. Their findings revealed that program-assessment provoked learning through assessment preparation, due to the pass or fail nature of the assessment and consequence of remediation or retake. If the comprehensive exam is similar to the licensure examination, there could also be a practice effect (Bridgham & Rothman, 1982), which may improve performance due to repeated practice. Practice tests are offered as a study method for the Praxis through ETS.

Oral Comprehensive Examination

Oral comprehensive examination is another possible option for summative assessment. This method of evaluation allows for assessment of a variety of skills, such as oral fluency, content knowledge, and critical thinking, and thus, allows a breadth and depth of knowledge to be demonstrated by the examinee (McAdams & Robertson, 2012). The oral comprehensive examination, however, can be subjective in nature, and includes several factors that may negatively affect performance. Wellington (2010) discussed several of these factors. In addition to stressors present across all types of high-stakes examinations, Wellington mentioned the added pressures of not thinking straight, becoming defensive, experiencing stage fright, talking too much, and becoming emotional. Torke and colleagues (2010) found that negative performance factors specific to oral comprehensive examinations caused inaccurate assessment of student competencies when compared to written assessments.

Wilbur (2015) reported minimal to no correlation between oral comprehensive examination and other measures of academic performance. Wilbur mentioned that poor inter-rater reliability among faculty groups may have been due to varying faculty experiences and levels and types of training. Likewise, Thomas and colleagues (1993) found poor agreement between examiners. Shenwai and Patil (2013) reported more positive feedback from students and faculty with a structured versus traditional oral comprehensive examination approach, but further research needs to be conducted to gain more objective data to support or refute the success of this design.

Capstone Summative Assessment

Programs across a variety of fields utilize capstone courses and projects as culminating experiences for students. There is minimal research, however, on capstone experiences and outcomes in graduate speech-language pathology programs and other health-related programs.

Even in fields such as occupational therapy where capstones are integrated within curriculum requirements, there is a lack of research and consistency among programs. Krusen and colleagues (2020) completed a scoping review of occupational therapy doctoral capstones and found there was a large gap in definition, intent, process, and outcomes of capstones.

Heidemann and colleagues (2018) discussed their internal medicine residency capstone courses – Residency Preparation Courses (RPC). Assessment for the courses was pass or fail, and the courses included, “interactive didactics, small-group discussions, procedure and code simulations, and simulated cross-cover” (p. 2048). In their study, students provided qualitative feedback related to their experiences. Student comments were positive and highlighted the benefit of synthesizing information near the end of their program when they were motivated and had the knowledge base for application. Assessment was not evaluated in their research; per student report, however, the courses, and timing of the courses, had a positive impact on their learning.

Hirsch and Parihar (2014) described their capstone course as a case-based course consisting of oral and written assignments as well as a final exam similar to the Georgia Board of Pharmacy Practical Examination and the NAPLEX. The pass rate for the course was 95%, and students expressed feeling more prepared for advanced pharmacy practice experiences. No specific objective or outcomes data was presented in their research aside from the overall pass rate for the course.

The capstone is an underrepresented method of summative assessment in the literature. Future research is needed to further investigate the capstone as both a culminating experience and summative assessment in CSD.

Summative Assessment at Eastern Illinois University

The inconsistencies and lack of information regarding summative assessment, as well as the instrumental role of summative assessment in accreditation and program refinement, led the Department of Communication Disorders and Sciences (CDS) Graduate Program at Eastern Illinois University (EIU) to evaluate and analyze its summative assessments and processes. The CDS department at EIU historically conducted summative assessments across four different measures, including Grand Rounds course final grade percentages, Written Comprehensive

examinations, Oral Comprehensive examinations, and although administered through ETS as an outside resource, Praxis II scores, which were also collected and considered in terms of program refinement. The CDS department began to question whether measurement of student learning outcomes and program evaluation could be achieved with fewer summative assessments.

The purpose of this study was to determine the relation among the various summative assessment measures. We hypothesized a strong relation among Grand Rounds final grade percentages, Written Comprehensive examination scores, and Praxis scores, and a weak relation between Oral Comprehensive examination scores and other summative assessment measures. Practically, we hoped outcomes of this study could be used to direct future summative assessment design in the EIU CDS department as well as inform other programs of results to guide their summative assessment planning.

Methods

The Eastern Illinois University Institutional Review Board approved all study methods. Data for this study was collected retrospectively from 260 EIU CDS graduate student records in the ten-year period between 2014-2023. Data included the Grand Rounds final grade percentages, Written Comprehensive examination total score percentage, Oral Comprehensive examination total score percentages, and the Praxis II scores following the 2014 modifications to the format and scoring of the earlier version of the Praxis. Data was excluded from students who did not have all measures, such as those who took the earlier version of the Praxis in 2014 or 2015, those who completed an optional master's thesis and were not required to take the Oral and Written Comprehensive examinations, or the few students who did not report Praxis scores to the university. If a student did not pass the Written Comprehensive examination, Oral Comprehensive examination, or Praxis on the first attempt and a retake was completed, the first attempt score was included for data analysis in the current study.

Grand Rounds. The first summative assessment EIU CDS graduate students encounter is in the Grand Rounds course. Students enroll in on-campus course and clinical practicum experiences at EIU during the Summer 1, Fall 1, Spring 1, and Summer 2 terms. Students take the Grand Rounds course in Summer 2 while also completing the last two disorder courses in CDS and their last on-campus practicum.

The Grand Rounds course in the EIU graduate program was initiated in 2004. This course was always taught during the last semester of on-campus courses and practicum, prior to off-campus full-time educational and medical internships. It initially began as a case-based course for review of low-incidence disorders with multiple faculty teaching for a short period of time in their areas of expertise. It evolved into a case-based course with half of the course having a pediatric focus, and half having an adult acquired medical focus taught by a variety of faculty for a few days each. By 2014, the current longstanding version of the Grand Rounds course had emerged, in which one graduate faculty member with expertise in a variety of pediatric areas teaches half of the course with cases that have many concomitant disorders from birth through school age. A second faculty member with expertise in medical speech-language pathology teaches the other half of the course, with complex multi-dimensional cases from acute care settings and rehabilitation settings. No new information is taught about specific communication disorders; rather, students are challenged to

apply knowledge as they analyze, evaluate, determine, and defend clinical decisions in complex and multifaceted cases. Students must critically evaluate and synthesize information from diverse perspectives. This course facilitated our primary goal of preparing students to independently integrate their knowledge and skills and defend decision making within the type of complex and dynamic cases frequently encountered in educational and medical settings. Near the conclusion of the Grand Rounds course, instructors ask students to reflect on their strengths and weaknesses in the class and in past graduate courses and clinical practicum to develop goals for themselves in internships, which are then discussed with faculty internship coordinators. During the capstone course, students are challenged to respond verbally to spontaneous questions and generate responses based on specific cases with intentional time constraints to prepare for the transition to professional-level expectations. Faculty ask questions and provide scaffolding for students to help them develop higher-level synthesis, insights, and conclusions. Within both the medical and educational sections of the course, students complete case-based individual and small-group collaborative projects, with grades determined through rubric scoring and detailed feedback on aspects of critical thinking. Some assignments must be executed in real time, during class, while other assignments are to be completed outside of class. Assignments are connected to complex cases where clients or patients present with multiple, overlapping concerns, disorders, needs, and contextual influences. Information is presented in the format of medical and school documents. Open-ended question prompts require decision making and defense of responses through oral and written formats.

Students then complete 50-question, complex case-based multiple-choice assessments, for the mid-term and final examinations. As with the cases for projects, those presented in the multiple-choice examinations include multiple concomitant diagnoses, competing needs and priorities, and various relevant and irrelevant factors to consider when executing critical thinking for clinical judgement. Several questions are attached to one case, often with new information being provided prior to subsequent questions, to simulate the dynamic evolution of cases in the medical or educational environment. Exam questions are similar to Hift's (2014) examples with context rich multiple-choice items; all choices could be plausible, and thus, students are required to use analysis, insight, synthesis, and many other higher-order thinking processes to apply knowledge and successfully draw conclusions to answer the questions accurately. Students determine the most appropriate clinical diagnoses, assessment plans, and treatment goals and methodology in series of questions about these complex cases (see the Appendix for a sample case and questions). All projects and examinations are summative in nature within the capstone class, requiring students to utilize and integrate information from their previous graduate courses, find and evaluate new information if needed, and synthesize all the information to draw and defend decision making and conclusions. Therefore, Grand Rounds final grade percentages were used as summative assessment data for this study.

Written Comprehensive Examination. Following Grand Rounds, students enroll in full-time, 14-week medical and educational internship experiences in the Fall 2 and Spring 2 semesters of their graduate program. In December or January following their first full-time internship placement, students return to campus for the Written Comprehensive examination. This Written Comprehensive examination is organized by topic areas from graduate-level courses which are similar to ASHA's Council for Clinical Certification (CFCC) and CAA standard areas. These topic areas include normal development, basic communication science, clinical practicum, research,

developmental and acquired speech sound disorders, developmental language, acquired language, cognitive communication, voice, fluency, dysphagia, audiology and aural rehabilitation, augmentative and alternative communication, and professional regulations and issues. The 200-question multiple choice Written Comprehensive examination was initiated at EIU more than 30 years ago as a comprehensive evaluation of students' knowledge and skills, near the culmination of their graduate education experience. Faculty teaching the courses that correspond to the Written Comprehensive Examination sections develop and periodically revise exam questions. The number of short case-based questions increased to comprise almost half of the items in 2020, but the number of items and the multiple-choice format have remained consistent. Many questions evoke critical thinking through carefully crafted question and response options that require application of coursework to the case to deduce the correct response. In contrast to the Grand Rounds course, the Written Comprehensive Examination offers a more direct comparison of student performance to specific content areas and disorder-specific learning objectives due to the organization of questions by topic and less integration of multiple disorders within one lengthy case-based question prompt (see the Appendix for sample questions). Students receive their Written Comprehensive examination score, and a breakdown of areas of strengths and weaknesses, to use as they continue their next internship placement and study for the Praxis. The Written Comprehensive examination score is the percentage accuracy from the 200-question exam; seventy percent and above is passing. Data collected from the Written Comprehensive examination for this study included the total score percent accuracy.

Oral Comprehensive Examination. While students are on campus in December or January for the Written Comprehensive examination, they also complete an Oral Comprehensive examination. The Oral Comprehensive examination at EIU also started more than 30 years ago. For more than a decade, the format had three university faculty members asking students to respond to a variety of content questions related to the learning objectives from their courses; responses were scored using a rubric rating. Faculty retreat discussions and data analysis indicated that the process induced anxiety and was not functional.

By 2000, a case-based Oral Comprehensive examination was initiated. Students chose a case from their fall internship, prepared a one-page handout that was given to the faculty two days in advance, presented the case to three faculty members, and defended their clinical judgment and decision making for assessment, diagnosis, and treatment of the patient. Students had 15 minutes to present their case without the aid of a script or PowerPoint presentation. The team of three faculty then asked questions for 15 minutes to identify any gaps in the student's application of knowledge, critical thinking, and evidence-based practices. Faculty then used a rubric to rate the student's oral defense in the areas of content, critical thinking, and verbal effectiveness. Students were required to earn 70% or greater of the points from the rubric to pass. After faculty completed their evaluation, the student was asked to return to receive scores and feedback regarding performance across the three areas.

Faculty reviewed data from summative measures annually and periodically discussed concerns that students who struggled with the Oral Comprehensive examination were not highly consistent with those who had been identified as needing assistance in formative assessments or other summative assessments. More detailed rubrics to increase faculty consistency in expectations and scoring were developed. Students were given additional resources about how to provide rationale

and insights in essential clinical decision making such as identifying and describing deficits, answering questions about the diagnosis and prognosis, digging into evidence, and describing choices in treatment goals and methods. Oral clinical case-based presentations with questions from a novel clinical instructor were added as a formative assessment component of two, and then four, on-campus clinical practicum assignments to improve students' skills.

In 2021, the EIU faculty voted to temporarily discontinue the Oral Comprehensive examination for students whose educational program had been modified due to the severe acute respiratory syndrome coronavirus 2 (COVID-19) pandemic. After the main effects of COVID-19 had subsided, the faculty reinstated the oral comprehensive exam for one year and recorded pass or fail results. Ultimately, the faculty then voted to permanently discontinue the Oral Comprehensive examination due to the addition of oral case-based presentations with questioning from novel instructors during clinical practicum assignments and then also in internships, as well as the perceived poor reliability in scoring and other extraneous factors influencing results. Data from the oral comprehensive examination was collected from the rubric total percentages for 166 students from 2014-2020. Oral comprehensive examination percentages were not included in analysis of data from 2021-2023 due to the decision to discontinue this summative assessment.

Praxis II in Speech-Language Pathology (5331). The Praxis II is a validated and reliable assessment required for professional certification by ASHA's CFCC and a criterion for state licensure. The Praxis II was developed and piloted in 2014 and has a score range of 100-200. The ETS technical manual (2021) describes the test development process, which included recruiting professors and clinicians as subject matter experts to complete a job analysis and write items for the test. The Praxis II content, written by subject matter experts nominated by ASHA, includes the following categories: foundations and professional practice; screening, assessment, evaluation, and diagnosis; and planning, implementation, and evaluation of treatment. Items were reviewed for quality and fairness by multiple other content experts. None of the authors have taken this version of the Praxis II and have no direct knowledge of the types of questions on the examination. Review of the sample questions in the ETS Praxis Study Companion for Speech-Language Pathology 5331 indicates that the short cases and other factual questions evaluate recall, understanding, and application of content knowledge (ETS, n.d.).

Test takers must achieve a minimum of 162 on a scale of 100-200 to pass the Praxis II (ASHA, 2023). In 2014-2015, test takers had the option of taking the previous Praxis (5330) which had a score range of 250-990 or the newly revised Praxis II (5331). An 80% or greater pass rate of students enrolled in university CSD programs is required for CAA program accreditation. CAA provides guidance to universities to count students who fail initial attempts but pass within the year as passing because "results should be reported only once for test-takers who took the exam multiple times in a single examination reporting period" (CAA, 2023c). ASHA (2023) recommends that students take the Praxis "no earlier than the completion of their graduate coursework and graduate clinical practicum or during their first year of clinical practice following graduation." EIU graduate students typically take the Praxis II during their final semester prior to graduation when all coursework has been completed, one full-time 14-week internship has been completed, and the second full-time 14-week internship is in progress. For this study, Praxis II total scores were collected to correlate to data from the program's summative assessments.

Data Analysis. Students who took the Praxis II and all departmental summative assessment measures were included in the analyses. Students with missing data were excluded listwise. Means and standard deviations were calculated for the measures for each year and overall. Pearson correlations were employed to determine the relations among the various summative assessment measures. The first set of Pearson correlations evaluated 166 students' Grand Rounds final grade percentages, Written Comprehensive examination total score percentages, Oral Comprehensive examination total score percentages, and Praxis scores. These correlations excluded students from 2021-2023 who either did not participate in Oral Comprehensive examinations or had pass/fail results. The second set of Pearson correlations evaluated 260 students' Grand Rounds final grade percentages, Written Comprehensive examination total score percentages, and Praxis II scores, which included students from 2014-2023.

The number of pass/fails for each measure was calculated for all students since 2014 who took the Praxis II measure and other summative assessments. Related-samples Cochran Q tests were used to evaluate differences in these nominal data.

Results

EIU's mean Praxis II score was 180.37 over the entire period, with mean scores from individual years ranging from 178.38 to 182.03. This is several points above the national mean of 175 -176 for all test takers annually. EIU graduate students' mean Grand Rounds course grade was 87.92%, with yearly averages ranging from a low of 84.99% to a high of 89.74%. EIU graduate students' mean score on the Written Comprehensive examination was 81.13%, with yearly averages ranging from 78.50% to 84.49%. EIU graduate students' mean score on the Oral Comprehensive examination was 80.40%, with yearly averages ranging from 76.96% to 83.46%. The mean scores for the Grand Rounds course were in the upper 80s with small standard deviations in individual performance. The Written Comprehensive examination and the Oral Comprehensive examination both had mean scores in the lower 80s, but standard deviations reflected variations in individual performance that were much larger for the Oral Comprehensive exam. Results are presented in Table 1.

For the 166 students who participated in all summative assessment measures between 2014 and 2020, the highest correlation was between the department's objective Written Comprehensive examination and the Praxis II ($r = .675, p < .001$), which was expected as they are both objective comprehensive exams covering knowledge, recall, application, and insight into many aspects within the scope of practice. The final grade percentage in the capstone Grand Rounds course had a moderately high relation with scores on both the Written Comprehensive examination ($r = .531, p < .001$) and the Praxis II ($r = .500, p < .001$). The Oral Comprehensive examination had the weakest relation with the other measures, with no significant relation with the Written Comprehensive examination ($r = .149, p = .056$), and a small but significant relation with the Grand Rounds final grade percentage ($r = .215, p = .005$) and the Praxis II ($r = .216, p = .005$). Results are summarized in Table 2.

Table 1

Mean scores and standard deviations (in parenthesis) for graduate students who took the Praxis II exam and other departmental summative measures from 2014 to 2023

Year (n)	Praxis II	Grand Rounds	Written	Oral
	M (SD)	M (SD)	M (SD)	M (SD)
2014 (1)	185.00	88.00	83.00	84.00
2015 (21)	178.57 (8.50)	84.99 (2.77)	80.67 (5.47)	79.57 (10.62)
2016 (28)	181.96 (6.72)	89.70 (3.47)	82.32 (5.14)	83.46 (10.64)
2017 (28)	181.50 (7.65)	89.74 (2.83)	83.45 (5.25)	79.93 (9.15)
2018 (35)	181.94 (9.64)	88.52 (3.03)	84.49 (5.77)	80.03 (10.86)
2019 (24)	178.38 (9.98)	87.96 (2.80)	81.96 (5.89)	76.96 (14.83)
2020 (29)	182.03 (7.56)	85.93 (4.26)	80.83 (4.26)	81.69 (11.97)
2021 (29)	178.55 (7.64)	87.81 (2.95)	79.57 (5.53)	-- **
2022 (29)	178.76 (8.40)	88.34 (3.18)	78.50 (7.48)	pass or fail***
2023 (34)	180.65 (7.73)	87.57 (2.54)	78.57 (6.79)	--****
Overall *	180.37 (8.24)	87.92 (3.38)	81.13 (6.13)	80.40 (11.36)

Note. * (n=260) and (n=166 for Oral), ** Oral Comprehensive examination was suspended this year due to COVID-19. *** Oral Comprehensive examination was graded as pass or fail this year. **** Oral Comprehensive examination was discontinued.

Table 2

Pearson correlations for 166 students participating in all summative assessments between 2014 and 2020

Assessment	Praxis II	Grand Rounds	Written
Grand Rounds	.500**		
Written	.675**	.531**	
Oral	.216*	.215*	.149

Note. * $p < .01$, ** $p < .001$

The Oral Comprehensive examination was suspended during COVID-19 in 2020, reinstated for one year with pass or fail scores recorded, and then permanently discontinued as a summative measure in 2023. Relations among the other measures were calculated for the longer period from 2014 to 2023, with 260 students. The pattern and strength of relations was consistent with the smaller set of years above. The highest correlation was between the department's objective Written Comprehensive examination and the Praxis II ($r = .673, p < .001$). The Capstone Grand Rounds final grade percentage had a moderately high relation to scores on both the Written Comprehensive examination ($r = .518, p < .001$) and the Praxis II ($r = .513, p < .001$). Results are summarized in Table 3.

Table 3

Pearson correlations for 260 students participating in all summative assessments between 2014 and 2023

Assessment	Praxis II	Written
Grand Rounds	.513**	.518**
Written	.673**	

Note. ** $p < .001$

For certification, graduate students and recent graduates can take the Praxis numerous times until they pass. ASHA's CAA for program accreditation requires an 80% or greater pass rate of students enrolled in university CSD programs for CAA program accreditation, but a person who passes the exam after multiple attempts in the same year is considered to have passed in this number, similar to someone who passes the first time. For EIU's summative comprehensive examination assessments, if students fail the first time, they are given a second attempt later in the same semester and graduation is not delayed. If students fail a second time, graduation must be postponed and an alternative assessment method used. In Table 4 below, we present the number of graduate students each year who did not achieve a passing score on the first attempt on the Praxis, the Written Comprehensive examination, the Oral Comprehensive examination, or received a C or lower in the Grand Rounds course. We also present overall first attempt pass rates for each measure in the bottom row of Table 4. Six students of the 260 in the sample failed the Praxis II on the first attempt; one of these students also failed the Oral Comprehensive examination on the first attempt. Ten of the 260 students failed the Written Comprehensive examination on the first attempt; eight of these students were in the three most recent years. The mean pass rate for EIU's written comprehensive exam was 99% in 2015-2020 before COVID-19 and reduced to 91% in the three most recent years from 2021 to 2023. Twenty-seven students failed the Oral Comprehensive examination on the first attempt, with at least two students failing each year. Although previous tables showed that the mean score on the Oral Comprehensive examination was only one point lower than the Written Comprehensive examination, the fail rate on the first attempt was more than two and a half times for the Oral Comprehensive exam compared to the Written Comprehensive exam. The overall pass rate for the first attempt of the Oral Comprehensive examination was 86%. EIU's Praxis II first attempt pass rate remained high during the 2021-2023 period at 99%, even though the Written Comprehensive examination pass rate dropped from 98% to 91% during that time. The national pass rate (162 or higher) for all test takers ($n = 76,069$) on the Speech Pathology Praxis II reported by ETS between 2014 and the summer of 2021 ranged from 80-83% each year

(ETS, 2023). In the two most recent years, the national pass rate on the Praxis II dropped to 76-77% for all test takers, September 2021 to August 2023 ($n = 26,168$; ETS, 2023).

Only two EIU students in total failed more than one assessment on the first attempt. No students received grades of C or lower in the Grand Rounds capstone course. All 27 students who failed the first attempt of the Oral Comprehensive examination passed on the second attempt. All 10 of the students who failed the first attempt of the Written Comprehensive examination passed on the second attempt. Five of the six students who failed the Praxis II on the first attempt, passed on the second attempt. One student received a 160 on the first Praxis II attempt in January, a 161 on the second Praxis II attempt in February, and a 175 on the third Praxis II attempt in March (162 is passing).

Table 4

Number of Students Who Failed the First Attempt

Year (n)	Praxis II	Written	Oral	Grand Rounds
2015 (21)	2	0	4	0
2016 (28)	0	0	4	0
2017 (28)	0	1	3	0
2018 (35)	1*	0	7	0
2019 (24)	2	1	3	0
2020 (29)	0	0	4	0
2021 (29)	1	2	--**	0
2022 (29)	0	3***	2	0
2023 (34)	0	3	--**	0
Overall	6	10	27	0

Note. *The student who failed Praxis II also failed the Oral Comprehensive examination. **Oral Comprehensive examination was not administered. ***One student who failed the Praxis II also failed the Oral Comprehensive examination.

A related-samples Cochran Q test indicated there was a significant difference in the number of pass or fail results between the assessment measures ($p < .001$). Pairwise comparisons with significance values adjusted by the Bonferroni correction for multiple tests indicated that the Oral Comprehensive examination had significantly more fails than each of the other measures (Written Comprehensive examination, $p < .001$; Praxis II, $p < .001$; Grand Rounds, $p < .001$). None of the other pairwise comparisons for fail rate were statistically different from each other (Written

Comprehensive examination and Praxis II, $p = .591$; Written Comprehensive examination and Grand Rounds, $p = .282$; Praxis II and Grand Rounds, $p = .591$)

Discussion

The current study compared correlations and pass rates among three program-specific summative measures as well as the national Praxis II examination. While CAA no longer accepts the Praxis as a program-specific summative measure because it is not validated for this purpose, the Praxis was included in this study because it is a validated and reliable measure developed by content matter experts that provides a broad view of students' general preparation for entry-level clinical practice. All CSD programs must post students' Praxis pass rates as an outcome measure, and all CSD students must pass the exam for state licensure and national certification. It is, therefore, valuable to include the Praxis pass rates for consideration when evaluating program outcomes and summative assessment measures.

Results of this study were consistent with our hypothesis that Grand Rounds course final grade percentages, Written Comprehensive examination scores, and Praxis II scores would be more strongly correlated than Oral Comprehensive examination scores. In fact, we found weak correlations of less than .22 between Oral Comprehensive examination scores and other summative assessments. These results, and other concerns described below, strengthened our rationale for eliminating the Oral Comprehensive examination as a high-stakes summative assessment, and led us to suggest that the Oral Comprehensive examination may be more appropriate as a formative assessment. Given evidence of a significant correlation between Grand Rounds final grade percentages and the Written Comprehensive examination, we considered if there was a continued need for both Grand Rounds and the Written Comprehensive examination; due to their unique contributions to students' education and preparation and strategic timing within the program sequence, however, we argue that both hold value as summative assessment. The Praxis remains a requirement of CFCC and CAA and provides an additional perspective from a validated source related to students' entry-level professional knowledge. While the Praxis cannot be included as a program summative assessment for individual students, CAA indicates that it provides an appraisal of student outcomes that should be considered during program evaluation and requires that pass rates be posted on program websites.

Our data revealed weak correlations between the Oral Comprehensive examination and other summative assessments, in addition to the Praxis II. It also indicated that the Oral Comprehensive examination had a significantly higher number of fails on the first attempt compared to the other measures. Past research indicated poor inter-rater reliability of the oral comprehensive examination, and lack of correlation of oral comprehensive examination results with other academic measures (Thomas et al., 1993; Torke et al., 2010; Wilbur, 2015). Faculty in CDS at EIU have frequently discussed the stressors present across all types of high-stakes examinations, and like Wellington (2010), have suggested that the Oral Comprehensive examinations have added psychological pressures for some students. Past yearly evaluation of Oral Comprehensive examination results compared to other formative and summative results by individual students align with Torke and colleagues (2010), who found that negative performance factors specific to oral comprehensive examinations sometimes caused inaccurate assessment of student competencies when compared to written assessments. The literature about comprehensive oral

examinations typically involves questions or cases chosen by the faculty. Historically, EIU used questions and cases chosen by faculty for the Oral Comprehensive examination; however, as previously discussed, the Oral Comprehensive examination was modified by 2000 to allow students to choose a case from their internship placement on which to present. Students at EIU have consistently reported preferring this method of oral comprehensive examination over a traditional format because it gave them more power to choose a case that they felt most comfortable with, and the students knew the case that they were preparing for in advance. Challenges with this method at EIU were differences in the complexities of the cases the students chose and the length of time the students saw the clients in different settings. For example, graduate students who had a medical internship in acute care in the fall semester may have only seen the client they were presenting about in the Oral Comprehensive examination for two sessions when the client was very ill. Other graduate students in medical rehabilitation settings may have seen the client for a few weeks to a month, and graduate students who had an educational internship in a school setting were more likely to have seen a client for 20-25 sessions during the 14-week internship. Faculty were aware that the possibilities for assessment and treatment varied for clients seen for different amounts of time in various settings. It was possible for students to score very highly on complex cases that they only saw for a limited time, but these students often had to respond to more hypothetical questions, than another graduate student who may have presented a case of a client with only a basic functional speech sound disorder who was treated for 12-14 weeks. These experiences, in addition to findings in the literature, suggest that oral comprehensive examinations often lack the validity and reliability needed to function as a summative assessment. Wilbur (2015) mentioned that poor inter-rater reliability for oral comprehensive examinations among faculty groups negatively impacted reliability, possibly due to varying faculty experiences and levels and types of training. Faculty at EIU also noticed this factor over time with Oral Comprehensive examinations. In response, EIU faculty groupings integrated senior with new faculty and calibration trainings and rubrics with more detailed sample responses were created to increase reliability and consistency in expectations and scoring. While not directly evaluated, we noticed through observation that reliability and case complexity was often a confounding factor during our Oral Comprehensive examinations, even with more detailed rubrics to guide expectations and calibration attempts.

In alignment with Epstein's (2007) call for summative assessment to be psychometrically sound, we suggest that oral comprehensive examinations may be better suited for formative assessment. Student stress, faculty group reliability and consistency, higher initial fail rates, and the weak relation between oral results and other formative and summative results were reasons that CSD faculty at EIU permanently discontinued Oral Comprehensive examinations as a summative assessment. Epstein (2007) stated that a distinction should be made between assessments that are suitable only for formative use and those that have sufficient psychometric rigor for summative use. We believe the skills in oral clinical case presentation and the ability to critically think and answer questions about clients are important skills, but we moved this task into four formative assessment tasks within on-campus, and two additional formative assessment tasks in internship clinical assignments.

Results of the current study mirror findings documented in previous research in terms of the correlation between written comprehensive examinations and licensure examinations. Like several studies (Fallon & Emanuel, 2019; Kosmahl, 2015; Nibert et al., 2002), the scores of our Written

Comprehensive examination correlated most strongly to scores students achieved on the Praxis II; the initial attempt fail rates were low and not significantly different between these two measures.

EIU's Written Comprehensive examination has always been comprised of 200 objective multiple-choice questions designed by faculty. Over time with revisions, the questions have included more brief written cases with multiple follow-up objective questions about the cases. Questions are organized by topics, and therefore, case questions are typically focused on diagnosis, assessment, or treatment decisions related just to that topic area. If the questions on the Praxis II are similar to the practice questions in the ETS 5331 study guide, it appears the Praxis II is also a multiple-choice objective exam comprised of many short case-based questions and other factual questions that evaluate recall, understanding, and application of content knowledge (ETS, n.d.). EIU had a slightly larger number of students fail the first attempt of the Written Comprehensive examination in the years during and immediately following COVID-19 (2 to 3 students per year) compared to prior to COVID-19 (0-1 student per year), and the mean score on the test dropped by a few points during these years. Prior to COVID-19, the mean pass rate for EIU's Written Comprehensive examination was 99% in 2015-2020, and reduced to 91% in the three most recent years from 2021 to 2023. EIU's Praxis II first attempt pass rate remained high during the 2021-2023 period at 99%. Nationally, the pass rate for all test takers on the Praxis II was 80-83% from the implementation of the 5331 test version in 2014, but the national pass rate on the Praxis II also declined in that 2021-2023 period to 76-77%. In addition to student learning changes related to COVID-19, such as open-note examinations and online courses, EIU also revised its Written Comprehensive Examination in 2019, so that almost half of the questions were case-based. Faculty who monitored the written examination reported that most of the students are taking the full three hours to complete the exam in recent years, compared to a larger majority of students finishing early in previous years. For EIU program evaluation, the question is if instructional or examination quality changed during that period, or if the nature of students' studying and ease at taking closed-note and closed-book exams changed due to COVID-19. Another possibility is that the increased proportion of case-based questions requires a different speed of processing and integrating information, than when the Written Comprehensive exam had fewer case-based questions.

Our Grand Rounds capstone course was like that of Hirsch and Parihar (2014) in terms of case-based content, assignments, and evaluation. While Hirsch and Parihar did not evaluate the correlation between their capstone course and licensure examination, our study found that performance in our capstone course had significant moderately high correlations with performance on the Praxis II. Positive student feedback and positive faculty impressions both indicate that the Grand Rounds course is a beneficial learning experience for students prior to internships. Occasionally, students at EIU have suggested that the Written Comprehensive examination may not be necessary since the examinations within the Grand Rounds course are comprehensive in nature. This thought is like those suggested by Cunningham and colleagues (1996), and indeed, the program might have the flexibility to do this within CAA's guidance for summative assessment options. However, CAA does state that summative assessment should include a comprehensive evaluation of learning outcomes, including knowledge and skills, at the culmination of the educational program. We reported the Grand Rounds course grades as a summative measure in this paper; we are unsure if CAA would consider it a true summative measure, however, because CAA's documentation guidance repeatedly states, "assess student performance at the END of the program" and lists "end of program comprehensive exams" and capstone projects as options.

The Written Comprehensive examination, which is administered in the final semester of the program, seems to be more at the culmination of the graduate education program, whereas the Grand Rounds experience is at the culmination of the on-campus portion of the program. EIU faculty also believe there is benefit in studying and taking the Written Comprehensive examination and then the Praxis after one full-time internship is completed. More of the material may make greater sense to students after seeing many more clients in a full-time medical or school setting. EIU also recognizes that our graduate students may not be the strongest test takers. In the years when the Graduate Record Examination (GRE) was required, our admitted class average was typically approximately 298, when many other CSD programs were requiring at least a 300 GRE score as an application minimum. We do believe the series of experiences starting with the Grand Rounds course, then full-time internships post coursework with a Written Comprehensive examination between the internships, and then preparing for the Praxis is meaningful. This series of experiences helps EIU students critically evaluate and integrate information to succeed in the summative assessments. All these questions inspire more investigation into the use and benefits of a capstone course as part of summative assessment, and the timing of other summative assessments in relation to required internships and certification and licensure examinations.

Grand Rounds offers students an opportunity for more in-depth and complex, integrated cases, whereas the Written Comprehensive examination provides a broad summative measure of student learning outcomes that can be directly compared to content learning objectives. The timing of Grand Rounds facilitates a transition between knowledge to practice, as this occurs prior to internship placements. Written Comprehensive examinations take place after students have had a culminating experience through internship and can serve as a culminating summative assessment near the end of their educational program. Thus, we conclude that both Grand Rounds and the Written Comprehensive examination serve a unique purpose crucial to student preparation and evaluation.

Limitations. Multiple factors may have affected results of the study. Results included data from the COVID-19 pandemic years during which learning was modified to respond to health and safety rules to reduce disease transmission. This alternative approach to learning deviated from our typical teaching practices. Additionally, the Grand Rounds course, Oral Comprehensive examination, and Written Comprehensive examination were modified over the years as program analysis and refinement occurred and as faculty changes influenced the dynamic relationship between teaching and summative assessment. We continue to consider other functional measures that could serve as summative assessments to evaluate the degree to which our students have acquired the knowledge and skills to be effective clinicians. Since all our students have full-time, 14-week internships in medical and educational settings, we regularly review ratings and feedback from the internship supervisors. The feedback from internship supervisors, as well as graduate student exit surveys, helps us to continue to modify course content and practicum education for program assessment. As a summative assessment measure to compare across individual students, however, numerical ratings between supervisors appear to lack reliability, and these do not meet CAA guidelines that summative assessment should be administered by a range of program faculty and supervisors to evaluate students' progress.

Future Research Directions. In general, there is a gap in research on summative assessment in the healthcare-related professions, and specifically in speech-language pathology. Medical educators developed a consensus statement for criteria for good assessment (Norcini et al., 2011) that includes (a) validity or coherence (measure what you hope to), (b) reproducibility or consistency (reliability), (c) equivalence (assessments administered at different institutions or testing cycles would have similar outcomes), (d) feasibility (efficient and cost effective), (e) educational effect (student who takes assessment is motivated to learn), (f) catalytic effect (the assessment provides outcomes that are fed back to the program for better teaching and learning), and (g) acceptability (to teachers and learners). Future researchers should survey programs in speech-language pathology to better understand current trends in summative assessment and how well assessments meet criteria for good assessment stated above or elsewhere. More detail is needed to describe summative assessments and capstone experiences as well as validate summative assessments at the program-level in speech-language pathology. Our research suggests the oral case-based examination should be considered for its value as a formative compared to a summative assessment; further research is needed to determine the best use of the oral comprehensive examination.

Author Disclosures

The authors have no relevant financial or nonfinancial relationships to disclose.

References

- American Speech-Language-Hearing Association. (2023). *About the speech-language pathology Praxis exam*. <https://www.asha.org/certification/praxis/about-the-speech-language-pathology-praxis-exam/>
- Bridgham, R. G., & Rothman, A. I. (1982). The effects of taking a practice examination on scores in the Qualifying Examination of the Medical Council of Canada. *Medical Education*, 16(4), 219-222. <https://doi.org/10.1111/j.1365-2923.1982.tb01252.x>
- Council on Academic Accreditation in Audiology and Speech-Language Pathology. (2023a). *Documenting compliance*. <https://caa.asha.org/reporting/documenting-compliance/>
- Council on Academic Accreditation in Audiology and Speech-Language Pathology. (2023b). *Guidance to programs: How to demonstrate compliance with Standard 5. 2- Program assessment of students (summative assessment)*. <https://caa.asha.org/siteassets/files/2023-standard-5.2-guidance-to-programs.pdf>
- Council on Academic Accreditation in Audiology and Speech-Language Pathology. (2023c). *Student achievement measures*. <https://caa.asha.org/reporting/student-achievement-measures/>
- Cunningham, D. R., Purvis, G. O., Baker, B. M., & Windmill, I. M. (1996). Comprehensive examinations: A tradition challenged. *NSSLHA Journal*, 23, 90-93. <https://doi.org/10.1044/nsshla.23.90>
- Educational Testing Service. (n.d.). *The Praxis study companion: Speech-language pathology 5331*. <https://www.ets.org/pdfs/praxis/5331.pdf>
- Educational Testing Service. (2021). *Technical manual for the Praxis® Tests and related assessments*. <https://www.ets.org/pdfs/praxis/technical-manual.pdf>
- Educational Testing Service. (2023). ETS data manager summary report for 5331 by year for all test takers Sep 2014 to Aug 2023. [Data Set]. ETS.
- Emanuel, D. C., Robinson, C. G., & Korczak, P. (2013). Development of a formative and summative assessment system for AuD education. *American Journal of Audiology*, 22(1), 14-25. [https://doi.org/10.1044/1059-0889\(2012/12-0037\)](https://doi.org/10.1044/1059-0889(2012/12-0037))
- Epstein, R. M. (2007). Assessment in medical education. *New England Journal of Medicine*, 356(4), 387-396. <https://doi.org/10.1056/NEJMra054784>
- Ewell, P., Paulson, K., & Kinzie, J. (2011). *Down and in: Assessment practices at the program level*. National Institute for Learning Outcomes Assessment. <https://files.eric.ed.gov/fulltext/ED535126.pdf>
- Fallon, K. A., & Emanuel, D. C. (2019). Journey of change: A summative assessment transformation. *Teaching and Learning in Communication Sciences & Disorders*, 3(1), Article 6. <https://doi.org/10.30707/TLCS3.1Fallon>
- Heeneman, S., Oudkerk Pool, A., Schuwirth, L. W., van der Vleuten, C. P., & Driessen, E. W. (2015). The impact of programmatic assessment on student learning: Theory versus practice. *Medical Education*, 49(5), 487-498. <https://doi.org/10.1111/medu.12645>
- Heidemann, L. A., Walford, E., Mack, J., Kolbe, M., & Morgan, H. K. (2018). Is there a role for internal medicine residency preparation courses in the fourth year curriculum? A single-center experience. *Journal of General Internal Medicine*, 33(12), 2048-2050. <https://doi.org/10.1007/s11606-018-4620-6>

- Hift, R. J. (2014). Should essays and other “open-ended”-type questions retain a place in written summative assessment in clinical medicine? *BMC Medical Education*, 14(1), Article 249. <https://doi.org/10.1186/s12909-014-0249-2>
- Hirsch, A. C., & Parihar, H. S. (2014). A capstone course with a comprehensive and integrated review of the pharmacy curriculum and student assessment as a preparation for advanced pharmacy practice experiences. *American Journal of Pharmaceutical Education*, 78(10), Article 192. <https://doi.org/10.5688/ajpe7810192>
- Kosmahl, E. M. (2005). Factors related to physical therapist license examination scores. *Journal of Physical Therapy Education*, 19(2), 52-56. <https://doi.org/10.1097/00001416-200507000-00007>
- Krusen, N. E., Murphy-Hagan, A., & Foidel, S. (2020). The purpose of capstone in an entry-level clinical doctorate: A scoping review. *Journal of Occupational Therapy Education*, 4(4), Article 4. <https://doi.org/10.26681/jote.2020.040404>
- Langford, R., & Young, A. (2013). Predicting NCLEX-RN success with the HESI Exit Exam: Eighth validity study. *Journal of Professional Nursing*, 29(2), S5-S9. <https://doi.org/10.1016/j.profnurs.2012.06.007>
- McAdams, C. R., III, & Robertson, D. L. (2012). An informed look at doctoral vivas (oral examinations) in the preparation of counselor educators. *Counselor Education and Supervision*, 51(3), 176-188. <https://doi.org/10.1002/j.1556-6978.2012.00013.x>
- Nibert, A. T., Young, A., & Adamson, C. (2008). Predicting NCLEX success with the HESI Exit Exam: Fourth annual validity study. *CIN: Computers, Informatics, Nursing*, 26(5), 28S-34S. <https://doi.org/10.1097/01.NCN.0000336439.16918.8b>
- Norcini, J., Anderson, B., Bollela, V., Costa, M. J., Duvivier, R., Galbraith, R., Hays, R., Kent, A., Perrott, V., & Roberts, T. (2011) Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher*, 33(3), 206-214. <https://doi.org/10.3109/0142159X.2011.551559>
- Palmer, E. J., & Devitt P. G. (2007). Assessment of higher order cognitive skills in undergraduate education: Modified essay or multiple choice questions? *BMC Medical Education*, 7(49), Article 49. <https://doi.org/10.1186/1472-6920-7-49>
- Park, S. K., Phillips, J., & Pavuluri, N. (2021). Systematic review of predictors of success for the North American Pharmacist Licensure Examination. *American Journal of Pharmaceutical Education*, 85(10), 1086-1100. <https://doi.org/10.5688/ajpe8591>
- Rowe, E. L., Pittman, J. M., & Hamilton, B. S. (2021). Use of an in-house comprehensive exam as a predictor for academic success. *Currents in Pharmacy Teaching and Learning*, 13(6), 643-651. <https://doi.org/10.1016/j.cptl.2021.01.036>
- Shah, M., Fuller, B., Gouveia, C., Mee, C. L., Baker, R. S., & San Pedro, M. O. Z. (2022). I . *Journal of Professional Nursing*, 39, 131-138. <https://doi.org/10.1016/j.profnurs.2022.01.010>
- Shenwai, M. R., & Patil, K. B. (2013). Introduction of structured oral examination as a novel assessment tool to first year medical students in physiology. *Journal of Clinical and Diagnostic Research*, 7(11), 2544-2547. <https://doi.org/10.7860/JCDR/2013/7350.3606>
- Thomas, C. S., Mellso, G., Callender, K., Crawshaw, J., Ellis, P. M., Hall, A., MacDonald, J., Silfverskiold, P., & Romans-Clarkson, S. (1993). The oral examination: A study of academic and non-academic factors. *Medical Education*, 27(5), 433-439. <https://doi.org/10.1111/j.1365-2923.1993.tb00297.x>

- Torke, S., Reem, R. A., Ramnaryan, K., & Asha, K. (2010). The impact of viva voce examination on students' performance in theory component of final summative evaluation in physiology. *Journal of Physiology and Pathophysiology*, *1*(1), 10–12.
- Wellington, J. (2010). Supporting students' preparation for the viva: Their pre-conceptions and implications for practice. *Teaching in Higher Education*, *15*(1), 71-84. <https://doi.org/10.1080/13562510903487867>
- Wilbur, K. (2015). Summative assessment in a doctor of pharmacy program: a critical insight. *Advances in Medical Education and Practice*, *6*(2015), 119-126. <https://doi.org/10.2147/AMEP.S77198>

Appendix

Sample Questions from Grand Rounds and Written Comprehensive Examination

Sample Questions from Grand Rounds Capstone Exams * Indicates the correct response.	Cognitive Requirements/Demands
<p>CASE #1: CHARLIE</p> <p>Charlie, a 24-year-old male, was admitted to the ER on Saturday, 6/6/23, s/p head-on MVA. He was unresponsive on-scene, extricated, intubated, and airlifted to a Level I Trauma Center. Charlie remained unresponsive to painful stimuli in ER. He was found to have a bilateral fronto-temporal depressed skull fx; intracerebral hemorrhage; elevated ICP; compound fracture L femur; heavy facial trauma to bilateral orbits. Trauma surgeons completed placement of EVD and craniotomy/removal of L fronto-temporal and superior parietal skull fragments. Surgeons were unable to save L eye but repaired L femur. By Sunday morning, Charlie was transferred to ICU post-operatively for monitoring, recovery, where he remained intubated x3 days. On Wednesday, Charlie was successfully extubated and showed some signs of emerging consciousness/responsiveness, although we are not yet certain. Charlie is NPO w/NG tube. Wednesday afternoon we get orders, "speech to eval and tx".</p> <ol style="list-style-type: none"> 1. Given your chart review, you predict Charlie may ultimately present with which of the following diagnoses: <ol style="list-style-type: none"> a. Aphasia only b. Motor speech disorder only c. Cog-Comm disorder only d. *Cog-Comm disorder, motor speech disorder, and potentially also dysphagia e. Both aphasia and a motor speech disorder 2. As you prepare to see Charlie for the first time, you anticipate the most 	<p>When approaching the case history information, students must be able to recognize, recall (remember), and interpret (understand) the medical terminology and abbreviations within the chart, the impact of injuries upon anatomy and physiology, and diagnostic features of different types of clinical disorders. Additionally, they must be able to discriminate between relevant versus irrelevant details and organize information (analyze) within a cognitive framework to support critical thinking related to traumatic brain injury in general, and specifically to this particular patient's brain injury circumstances. Finally, students must judge the information related to suspected diagnoses and areas of deficit (evaluate) and hypothesize the most likely outcome (create) prior to and during their thought processes to select the most plausible multiple-choice response to the question. This exemplifies cognitive processing across lower- and higher-order levels of Bloom's taxonomy.</p> <p>When responding to questions, students must again recognize, recall (remember), and interpret (understand) the medical terminology related to speech-language pathology topics and then infer (understand) which response could be possible based on the case and available options. They must then compare (understand) the options and differentiate (analyze) which option is most likely. Ideally, students then check (evaluate) their response to ensure alignment between case information and background knowledge of the option and their hypothesis (create) based on the case.</p> <p>Examples of breakdown across areas of Bloom's Revised Taxonomy may occur as follows:</p>

<p>pressing needs and goals for your evaluation. You determine that your plan bedside is to:</p> <ol style="list-style-type: none"> Complete a clinical swallow eval to determine what diet Charlie can tolerate safely *Determine Charlie's level of consciousness, and the nature and consistency of any responsiveness to stimuli Evaluate Charlie's orientation and reliability for answering Y/N questions None of the above A, B, and C <p>3. Given your analysis of Charlie's and the team's immediate needs, which of the following assessment tools would you use during your Wednesday afternoon bedside visit:</p> <ol style="list-style-type: none"> Montreal Cognitive Assessment Mann Assessment of Swallowing Ability Cognitive Linguistic Quick Test (CLQT) *Coma Recovery Scale – Revised Bedside Form, Western Aphasia Battery – Revised <p><i>New information on Thursday morning:</i> Last night Charlie regained consciousness, but he is agitated, thrashes in bed, pulls at his tubes, and has had to be physically restrained to prevent himself from removing surgical dressings and IVs. Charlie inconsistently responds to nurses' simple directives. He vocalizes but produces unintelligible, confused, nonmeaningful utterances.</p> <p>4. Given Charlie's new status, what is your plan for your inpatient session with Charlie today?</p> <ol style="list-style-type: none"> Evaluate Charlie's level of alertness, nature and consistency 	<p><i>Remember:</i> If a student fails to recall terms or abbreviations, such as bilateral fronto-temporal depressed skull fx, intracerebral hemorrhage, elevated ICP, compound fracture L femur, and heavy facial trauma to bilateral orbits, they may not recognize the diffuse nature of the injury with the most significant areas of impact near the anterior portion of the cerebral cortex.</p> <p><i>Understand:</i> If students recall terms but do not classify, infer, and interpret them correctly, they may have gaps in their critical thinking. For example, if a student associated the bilateral orbits with the occipital lobe, they may incorrectly understand the most significant areas of damage.</p> <p><i>Apply and Analyze:</i> If students fail to capture and apply their knowledge to the specific, unique details of the case, they risk providing robotic, generic responses that fail to address the individuality of a patient's needs. For example, if students focus solely on assessment tools that could be used for a patient with a traumatic brain injury, they might elect to use an evaluation tool, such as the CLQT. While the CLQT can be used to evaluate a number of cognitive processes in patients with traumatic brain injury, that particular tool is not appropriate for our patient in his initial days. Indeed, he is not even yet fully conscious. Students must also demonstrate analysis as they distinguish between relevant and irrelevant information and recognize the implications of certain details. Inaccurate integration of information can result in faulty conclusions. For example, if a student fails to find the connection between the multiple sites of lesion, diffuse brain damage, and cognitive communication disorder, they may incorrectly diagnose acquired disorders that result from more localized lesions (e.g., aphasia, right hemisphere dysfunction) instead of a cognitive communication disorder.</p>
---	---

- of responsiveness, and degree of orientation
- b. Determine if Charlie has a reliable response to basic Yes/No questions and comprehension for simple, concrete requests
 - c. Conduct a thorough, bedside/clinical assessment of oral motor and swallowing ability
 - d. *A. and B. only
 - e. A., B., and C.

New information from your Thursday morning inpatient visit: Charlie attends for up to 3 seconds at a time and continues to attempt to remove arm restraints. Charlie often yells out, but his speech appears to be severely dysarthric, and his vocal quality is wet. You offer some one-step requests (e.g., squeeze my hand, shake your head, close your eyes, open your mouth), but Charlie does not demonstrate intentional or corresponding responses. You are also unable to determine any sort of reliable means of responding to Yes/No queries.

5. Charlie's mother arrives. She states that she brought in his favorite frozen drink this morning and asks you when he will start eating. You respond with which one of the following:

- a. Charlie is going to need long-term alternative nutrition. I'm recommending the medical team consider placement of a PEG.
- b. Charlie is not yet able to safely try any foods or liquids; I will be monitoring his cognition and oral motor/sensory skills to determine when we can safely evaluate his potential for eating.
- c. Charlie should not have anything just yet by mouth, not even his favorite drink. He may not be able to safely manipulate or

Evaluate and Create: If students do not evaluate their response in accordance with the case, they may at the very least overlook details that could influence their response. For example, if a student focuses heavily on the cognitive aspects of the injury and fails to evaluate those details to connect the case with concomitant dysphagia, they may only include cognitive-linguistic diagnoses in their response. Further, if a student fails to evaluate their presumptions or previous conclusions, in light of new information, they may fail to revise plans or priorities in keeping with the patient's evolving status. This can result in the student planning and carrying out tasks which are no longer timely or relevant, putting the patient at risk and failing to offer the medical team the most useful insight possible. Students must also demonstrate defensible judgment as they appraise the information not only from the chart, but from other constituents (e.g., other medical staff, family). Astute clinicians must recognize the need to integrate multiple sources of rapidly changing information, search for inconsistencies or gaps in their conclusions, and deliberately employ careful, critical thinking to reconcile those inconsistencies. For example, nursing comments describing a patient's attempts to communicate must be infused and compared to foundational knowledge of the specifics of the patient's brain injury and evaluated for deficits in cognitive, linguistic, motor speech, voice (or all of the above), in order to provide the most appropriate and useful recommendations to others, for how to communicate with a patient. This same process applies to on-the-spot decision making for answering questions from family members who are often eager to make things 'better' in ways the patient may not yet be able to participate in or tolerate. Careful judgment is necessary to provide not only defensible recommendations, but to do so in a way that is in tune with family members.

control food and liquids, and this could put him at risk for ...

- d. All of the above
- e. *B. and C. only

New information from nursing staff: The nurse states that Charlie has been saying “oh—oh—oh,” since 3 a.m. Thursday. She says he seems to be calling for her but every time she goes into his room, “he won’t tell me what he wants”.

6. The nurse asks if you can give Charlie a communication board so he can point and spell out things. You explain/state that:

- a. Charlie is a perfect candidate for an alphabet board; I'm bringing one in this afternoon.
- b. Charlie is a perfect candidate for a speech generating device; I'm bringing on in this afternoon.
- c. *For now, try asking Charlie simple, short Yes/No questions about his pain; speak slowly and clearly and wait several seconds to observe for any response.
- d. For now, try asking Charlie open-ended questions, such as "What happens when you try to move your right hand?" and let him explain.
- e. For now, ask Charlie open-ended questions and repeat them every few seconds.

New information, new orders: It’s Friday, and you have new orders to eval Charlie’s swallowing. The physiatrist wants to decide about a PEG soon. However, the overnight shift reported that Charlie had complications around 4 a.m. (early Friday). It appears that he had a series of seizures which have now been controlled with meds. He also presents with a suspected deep vein thrombosis/clot in L leg and an accumulation of fluid in the pericardial space around his heart.

7. Your visit this morning will focus on which one of the following:

- a. Ongoing cognitive therapy and family education; ignore the swallow orders, as they are not relevant.
- b. Re-evaluate Charlie's cognitive, oral motor, and language status, given new issues.
- c. Determine if you can, or cannot, do this clinical swallow eval today, given Charlie's new medical circumstances. Document accordingly.
- d. Proceed with the clinical swallow eval as requested by the physician yesterday.
- e. *B. and C.


8. Which of the following cognitive processes are you most concerned about for Charlie, at the moment?

- a. Sustained attention to simple stimuli
- b. Basic processing of auditory and visual stimuli
- c. Reasoning and executive functions
- d. *A. and B.
- e. A. B. and C.

New week and new updates: It was a long week, and you're glad to be home Saturday and Sunday. You return to work on Monday and discover that Charlie has become more calm and focused. He no longer requires restraints and nursing notes seem to indicate Charlie has been asking for his girlfriend, but still seems fairly confused. PT notes state that Charlie is A&Ox2.

9. Which of the following assessment tools will you bring with you this Monday morning?

- a. BEST

<ul style="list-style-type: none"> b. Western Neuro Sensory Stimulation Profile c. BURNS Right Hem Inventory d. *CLQT e. BDAE Short Form <p>10. As of Monday morning, what Rancho Los Amigos Scale of Cognitive Functioning Level do you anticipate Charlie to be functioning at?</p> <ul style="list-style-type: none"> a. Level II b. Level IV c. *Level V d. Level IX e. Level X 	
<p>Sample Questions from Written Comprehensive Exam (Topic = Motor Speech) * Indicates the correct response.</p>	<p>Cognitive Requirements/Demands</p>
<p>You are evaluating Hector, a 25-year-old with a complaint of “slurred speech” and “difficulty talking”. During the oral mechanism exam you observe the following:</p>  <ul style="list-style-type: none"> 1. Based on the picture above what is the most appropriate interpretation <ul style="list-style-type: none"> a. Hector appears to have appropriate tongue strength as he is able to protrude his tongue anteriorly. b. *Hector appears to have left sided tongue weakness c. Hector appears to have right sided tongue weakness. 2. Where in the motor system is the most likely damage? <ul style="list-style-type: none"> a. Contralateral upper motor neuron b. Ipsilateral upper motor neuron c. Contralateral lower motor neuron d. *Ipsilateral lower motor neuron 	<p>When approaching the case history information, students would need to recall (remember) motor speech disorders and the types of upper motor neuron and lower motor neuron damage that can contribute to motor speech disorders, interpret (understand) case history details correctly within the context of their motor speech content knowledge, differentiate between relevant and irrelevant details to attribute (analyze) the signs and symptoms to the correct motor speech disorder, check (evaluate) their thinking to ensure accurate connections, and hypothesize (create) the motor speech disorder and associated details based on the case information.</p> <p>Examples of breakdown across areas of Bloom’s Revised Taxonomy may occur as follows:</p> <p>Remember: For the first question, students may not recognize the need to orient themselves to the patient’s left versus their left as they are facing the patient, which would cause the student to incorrectly identify weakness on the right versus left side of the tongue. For the second question, students need to recall signs and symptoms of motor neuron damage.</p>

<p>e. There is no evidence of damage as the tongue is able to protrude anteriorly.</p> <p>3. Based on the small amount of information you have at this point, which motor speech disorder is most likely?</p> <ul style="list-style-type: none">a. *Flaccid dysarthriab. Spastic dysarthriac. Hypokinetic dysarthriad. Unilateral upper motor neuron dysarthriae. Apraxiaf. We have no evidence at this point of a possible motor speech disorder.	<p>Understand: For the second question, students may have difficulty accurately categorizing the signs and symptoms into upper versus lower motor neuron damage, which could cause them to incorrectly answer that the patient is experiencing contralateral upper motor neuron damage.</p> <p>Apply and Analyze: If students are unable to use their knowledge to organize signs and symptoms and discriminate among the various motor speech disorders, they will lack the cognitive ability needed to select the most likely type of motor speech disorder.</p> <p>Evaluate and Create: If students fail to check all relevant details against their hypothesis and integrate signs and symptoms in the case information with the various motor speech disorders, they may overlook details that lead to the correct diagnosis. For example, if students hypothesize an upper motor neuron lesion and fail to revise their hypothesis when reviewing the image of the individual when lingual atrophy is displayed, they will select the incorrect type of dysarthria, unilateral upper motor neuron dysarthria versus flaccid dysarthria.</p>
---	--