

Uncovering the Landscape: A Closer Look at Internships in Educational Measurement

Autumn N. Wild, M.A.  and Brian C. Leventhal, Ph.D. , James Madison University, and Joseph M. Kush, Ph.D.,  IXL Learning

Abstract: *High-impact practices such as internships have long been recognized for enhancing student learning and development. Yet, in educational measurement graduate training, internships remain underexamined. The current three-part study explored the structure, expectations, and implications of graduate-level internships in educational measurement and assessment. Part 1 reviewed 19 internship calls from summer 2024, identifying trends in required skills, application materials, and internship logistics. Part 2 surveyed 22 internship mentors and organizers to better understand specifics pertaining to mentorship, project supervision, the application review/interview process, and internship logistics. Part 3 surveyed 28 current graduate students, examining the value of internships from the intern perspective. Findings suggest that graduate internships in the fields of educational measurement and assessment are valuable and impactful experiences that serve a variety of stakeholders. Study limitations and directions of future research are discussed. Implications are provided for graduate students, graduate programs/faculty, internship organizers/mentors, and the field.*

Keywords: foundational competencies, internships, professional development

In recent decades, high-impact practices (HIPs) have been widely recognized for supporting student learning, engagement, and professional development (Kuh et al., 2008). Although the HIPs framework was initially articulated in the context of undergraduate education, many of the experiences such as intensive feedback, mentorship, applied work, and sustained engagement with authentic problems are also common features of graduate training, particularly in applied and quantitatively oriented fields. Graduate programs in educational measurement and assessment frequently incorporate writing-intensive coursework, collaborative research, cohort-based learning communities, and capstone projects such as theses or dissertations, all of which align with core features of Kuh et al.'s (2008) framework.

One high-impact practice with clear relevance to educational measurement graduate education is the internship experience. Despite their potential to connect academic preparation with professional practice, internships remain under-researched. Internships can provide structured opportunities to apply classroom knowledge in real-world settings, engage with interdisciplinary teams, and receive mentorship from experienced practitioners. In educational measurement specifically, internships may serve as a vital connection between the foundational competencies outlined by Ackerman et al. (2024) and the practical demands of careers in industry. Students develop these competencies, spanning communi-

cation and collaboration, technical and computational skills, and educational measurement expertise, through both curricular and co-curricular structures, with internships explicitly named as one such structure (Ackerman et al., 2024, p. 13). Despite this, there is limited systematic evidence describing how educational measurement internships are structured, what competencies are required for success prior to and during the internship experience, or how these experiences are perceived by multiple stakeholders, including students.

Studies in adjacent disciplines (e.g., statistics, user experience) suggest that internships can support mentorship, applied skill development, and career clarifications but also highlight persistent concerns related to access and equity (Anderson-Cook et al., 2017; Kang & Girouard, 2022; Schnoes et al., 2018). Structural and logistical factors such as compensation, work modality including remote, in-person or hybrid arrangements, mentorship structures, and institutional support may shape who can pursue internships and what students gain from them (Hora et al., 2020). These concerns are particularly salient in fields where internship participation is portrayed as merit based, but may nonetheless be constrained by financial, logistical, and informational barriers (Gaikwad et al., 2024; Jenkins & Yao, 2025; Wilson et al., 2022). To what extent these patterns are present in educational measurement internships remains largely unexplored.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Given the growing attention within the educational measurement community to internships (e.g., Bay et al., 2025; Rowe et al., 2024) and the demand for applied experience in educational measurement careers (e.g., Cheng et al., 2024), there is a clear need for empirical work that systematically describes current internship opportunities and examines internships from multiple stakeholder perspectives. In the present study, we address this gap through a three-part investigation of graduate-level internships in educational measurement and assessment. Through each part of our study (henceforth referred to as Study 1, Study 2, and Study 3), we examine information provided within internship calls, how internship mentors and organizations describe intern selection, supervision, and expectations, and how graduate students perceive internship preparation, access, and value. By integrating evidence across calls for internship positions, professionals, and students, this work provides a descriptive foundation for understanding expectations, barriers, and perceived outcomes within the current internship landscape in educational measurement.

Specifically, this work was guided by the following research questions:

1. What baseline competencies do students need to secure internships, and how are these assessed or communicated?
2. How can graduate programs better prepare students, and how can students prepare themselves, to be competitive applicants for internships that align with their professional goals?
3. What barriers exist to equitable access to internships, including those related to format (e.g., remote, hybrid), compensation, mentorship, and institutional support?
4. How do different stakeholders, including students, internship mentors, and host organizations, perceive the expectations, structure, and value of graduate internships?

Study 1: Review of Calls

Our initial study examined graduate internship opportunities within the fields of educational measurement and assessment. For this review, we collected and evaluated a sample of internship calls for applications (henceforth referred to as calls) using a set of criteria.

Methods. The research team consisted of one doctoral student and two faculty members. “Internship opportunities” included non-academic positions tailored to graduate students in the fields of educational measurement and assessment. Attention was given to internship opportunities marketed as providing experience in quantitative methods, psychometrics, assessment, measurement, and/or educational data analysis. Internship calls were solicited from various sources, including the National Council on Measurement in Education (NCME) website (NCME, 2025), the NCME Graduate Student Issues Committee (GSIC), LinkedIn, and email correspondence with affiliated organizations. Calls considered for inclusion were for opportunities taking place within the United States in the summer of 2024 (May 1 through August 31). Paid and unpaid opportunities were considered, as well as full-time and part-time positions. The final analytic sample included 19 internship calls (see Appendix A).

With the final sample of calls determined, the research team created a list of 63 criteria to evaluate the calls (see Appendix B). These 63 criteria fell into five primary categories: Application Material, Applicant Qualifications, Internship Logistics, Internship Experiences, and Other Important Features. These criteria were determined based on the research questions, the research team’s familiarity with internship calls, prior experience with internship applications, previous NCME discussion sessions on internships (e.g., Alahmadi et al., 2023), and shared professional interests/questions relevant to the field.

Once the evaluation criteria were finalized, each member of the research team independently rated the 19 calls against the criteria before meeting to adjudicate discrepancies. During this meeting, the three raters compared their initial ratings and discussed all differences. Notably, most discrepancies arose from the graduate student member’s interpretations of the internship calls, particularly when internship descriptions blurred the distinction between required and preferred qualifications (e.g., coursework, software skills). This pattern is noteworthy because graduate students are the primary audience for these calls, yet the expectations were often communicated in ways that presupposed prior experience with how to interpret internship calls rather than providing explicit guidance. For any criterion without full initial agreement, we defined consensus as alignment between two of the three raters, though discussion continued until all three team members agreed on the final rating. These adjudication discussions highlighted the need for organizations to communicate requirements more explicitly and consistently so that prospective applicants, especially those early in their training, can accurately assess their preparedness.

Results.

Application material. Internship calls sometimes required letters of recommendation (number of calls(n) = 5) or a list of references (n = 2). Most calls requested a curriculum vita (CV) or résumé, with a CV being more commonly required (n = 9 CV required; n = 2 résumé required; n = 4 either accepted). Transcripts were also commonly requested (n = 7), with a majority accepting unofficial copies (n = 6). One internship call explicitly required an official transcript, which may have an implied additional cost to applicants. Only one internship required a writing sample, while another listed it as optional. A cover letter or statement of interest/purpose was typically mandatory (n = 15 required; n = 1 requested). These statements often asked applicants to address their fit for the position, project preferences, research interests, relevant experiences, and career goals. Most internship calls were for general positions (n = 14), with a few targeting specific tracks (n = 2) or projects (n = 3). A small number of postings (n = 2) required applicants to submit research ideas, and an equal amount simply requested them. Two internship calls noted that a knowledge test would be part of the application process. Notably, no application required a diversity statement.

Applicant qualifications. Internship calls frequently required applicants to have knowledge in broad topics such as item response theory (IRT) and classical test theory (CTT), with additional preference given to general

psychometrics and measurement theory. However, only six internship calls explicitly listed required coursework (e.g., CTT, IRT, research design), while five listed preferred courses.

Software proficiency was a common requirement, with eight internship calls specifying particular tools, most notably R ($n = 7$) and Python ($n = 6$). An additional five calls noted preferred, but not required, software experience (e.g., WinSteps). Notably, most calls mentioned software only in general terms, requiring experience or familiarity without specifying particular skills or libraries.

Effective communication skills were also emphasized. Oral communication skills were required in six calls, written communication skills in four, with both oral and written communication listed as preferred qualifications in two others. Again, no specific skills within these domains were articulated. Applicants were expected to demonstrate the ability to collaborate effectively while also working independently. These competencies were often expected to be highlighted in the application materials such as the cover letter and CV.

Experience requirements varied. Most ($n = 15$) internships were open only to doctoral students, with 10 specifically requiring applicants to at least be in their second or third year. A few opportunities ($n = 2$) were open to students at the B.A. or M.A. level. While most calls targeted students in psychometrics explicitly, some welcomed applicants from disciplines such as mathematics, industrial-organizational psychology, and developmental or educational psychology. Only one internship call explicitly required prior research experience (e.g., analyzing large-scale data), and one other listed research experience as preferred. Additional considerations included a demonstrated interest in certification testing, a preference for dissertating applicants, and encouragement for students from underrepresented groups to apply.

Internship logistics. The majority of internship calls ($n = 15$) specified a monetary stipend. Of these, 2 offered hourly wages (\$25/hour and \$32/hour, both for 8-week experiences), whereas 13 provided a total stipend; notably, both the minimum (\$6,000; remote) and the maximum (\$15,000; in-person) amounts corresponded to 8-week experiences ($M = \$9,900$, $SD = \$2,300$). Most calls ($n = 14$) did not specify expected weekly work hours; however, among those that did ($n = 5$), the range was 30 to 40 hours per week. Nearly all internships began in early June, with 11 noting flexibility in the exact start date. Duration of experiences typically ranged from 8 to 10 weeks, though the shortest was 4 weeks and the longest 12.

Work modalities varied across postings. While three internships did not specify a modality, others offered in-person ($n = 6$), virtual ($n = 5$), or hybrid ($n = 4$) formats. Most in-person or hybrid internships offered relocation support, commonly around \$2,000.

Only one internship had an application deadline before January 1 of the internship year. Other deadlines were distributed across January ($n = 3$), February ($n = 8$), and March ($n = 6$). A majority ($n = 12$) of postings provided details about the application review process, such as timeline or interview format. Of these 12, 7 explicitly indicated that an interview would be a part of the application experience. Additionally, nine internship calls hinted at the intended cohort size, ranging from a single intern to multiple participants.

Internship experiences. The majority of internships reviewed were research-focused ($n = 13$), while a smaller number emphasized operational experiences ($n = 4$). Several calls ($n = 7$) included examples of research projects, either as potential topics for incoming interns or as summaries of past intern work. Although most internships indicated that duties would be conducted independently ($n = 13$), many also mentioned the availability of mentorship ($n = 11$). Four calls highlighted unique experiences offered to interns, such as a paid trip to the host office or provision of an official work laptop.

In terms of outcomes, most internships ($n = 11$) expected or encouraged the completion of a final product, such as a conference proposal or a written deliverable for internal or external use. However, only three calls explicitly mentioned support for conference participation, with funding ranging from \$1,000 to the full coverage of travel expenses.

Other important features. Several internship calls reflected a commitment to diversifying the field. Two opportunities were specifically designated for Black, Brown, or Indigenous students, while others broadly encouraged applications from individuals from underrepresented backgrounds. International applicants were explicitly welcomed in five postings, provided they were eligible to work in the United States.

Applicants with interests in psychometrics operations, certification testing, education improvement, and non-profit work were particularly valued across multiple calls. Additionally, most calls ($n = 13$) included information about the host organization's mission, and 11 explicitly highlighted a commitment to diversity, equity, and inclusion within their organizational values.

Study 2: Internship Survey—Mentors and Organizers

Building on the findings from the review of internship calls, in Study 2 we sought to gain deeper insights into the internship experience by directly surveying individuals who worked with interns in some capacity (e.g., mentor, supervisor) during the experience. This approach allowed for the exploration of perspectives and contextual factors that could not be captured through the analysis of internship calls alone.

Methods. We developed the survey based on our research questions, findings from Study 1, informal discussions within and beyond our research team, and our professional experience in the field. We disseminated the survey using Qualtrics, designing it to include up to 80 items depending on branching logic. Though most of the questions used Likert response scales, we also employed multiple choice, open response, and select-all-that-apply items. The questions covered seven thematic areas: Demographics/Sorting Variables, Mentorship, Project, Application, Interview, Qualities, and Logistics (see Appendix C). We distributed the survey to designated contacts listed in the internship calls from Study 1, contacts listed on the NCME website (NCME, 2025), members of the Psychometric Internship Collective, and contacts affiliated with the Educators of Measurement SIGIMIE. We also used snowball sampling techniques to expand the survey's reach, encouraging participants to share with colleagues that met at least one of the following criteria:

- organizer of an internship program in industry (with or without direct supervision of graduate-level interns),
- professional who has mentored at least one graduate-level intern since summer 2021, and/or
- professional who has supervised at least one graduate-level intern since summer 2021.

After receiving approval from the Institutional Review Board, we launched the survey on January 25, 2025 and sent a reminder email to initial contacts on February 13, 2025. We closed the survey on February 28, 2025. Of the responses received, eight were removed due to incomplete responses and an additional two were removed due to lack of consent, resulting in a final analytic sample of 22 respondents.

Respondents primarily worked in K–12 assessment, higher education assessment, and licensure/certification contexts. Represented organizations ranged in size from 11–25 employees to more than 100 employees. Of the 22 respondents, 20 identified as internship mentors, 10 of whom also indicated they served as internship organizers. One additional respondent identified solely as an internship organizer, and one respondent reported previously chairing an internship program. All respondents indicated they had supervised at least one intern project since 2021, with nearly all respondents also reporting involvement in reviewing applications ($n = 21$) and interviewing candidates ($n = 20$).

Results. To support interpretation of program-level patterns and address questions about organization representation, we first situate the results within the organization contexts reported by respondents.

Among respondents affiliated with K–12 assessment organizations employing more than 100 individuals ($n = 7$), one internship program was described as research-focused, one as operations-focused, and five as combining research and operations. Of these five dual-focus programs, only two reported the same stipend amount. The two higher education assessment organizations of similar size each reported distinct compensation levels. Among licensure and certification respondents from organizations with 26–100 employees ($n = 2$), one internship program focused exclusively on research while the other combined research and operations. For the remaining organizations employing more than 100 individuals ($n = 7$), 4 respondents reported being from research-focused internships and 3 from dual-focus, with compensation amounts differing across programs. Within the organization categorized as “Other” ($n = 4$), two responses were clearly distinguishable, while the remaining respondents could not be confidently separated as representing different organizations.

Taken together, we conservatively estimate that these data represent individuals from approximately 19 unique internship programs. Because organizations may offer multiple internship programs with different emphases, these estimates reflect distinct internship programs rather than unique organizations. In addition, because respondents reported on internship experiences spanning the previous 3 years, program characteristics and compensation structures may have changed over time. Accordingly, differences in reported stipends may reflect program evolution rather than distinct internship offerings. Given our sampling techniques were broad and not targeted, it is also possible that some internship programs described by respondents did not appear among the internship calls reviewed in Study 1.

Mentorship. In this section, we explored the dynamics of mentorship, including preparation, challenges, and rewards associated with guiding interns. Only 6% of respondents reported receiving formal training from their organization on how to be an effective mentor, while 67% indicated they had not received any training. Despite this, approximately 89% felt they were adequately prepared to mentor interns. Most respondents ($n = 14$) agreed or strongly agreed they had access to the necessary resources to mentor effectively. The same number reported experiencing professional growth through the mentoring process. All respondents described mentoring as a rewarding experience and stated they would recommend mentoring interns to other professionals.

Project. In this section, we examined the experience of supervising intern research projects, focusing on the time commitment, challenges, and successes. Nearly all respondents (approximately 95%) reported that interns had a positive impact on team projects. Most respondents observed improvements in interns’ technical skills ($n = 16$), and even more noted growth in their interpersonal skills ($n = 17$).

Respondents overwhelmingly agreed that interns demonstrated a considerable or high level of proficiency with statistical software (~95%). Additionally, 62% believed interns showed a strong understanding of data collection techniques, and 66% reported a solid grasp of research design principles.

On average, project supervisors spent 4.5 hours per week ($SD = 1.2$) meeting one-on-one with interns and another 4.5 hours per week ($SD = 2.4$) working on intern-related projects outside of those meetings. Most respondents ($n = 18$) felt that interns effectively presented their project findings and contributed meaningfully during project discussions ($n = 18$).

Application. In this section, we examined how reviewers evaluate internship applications, with a focus on identifying the skills and attributes they prioritize in the process. Nearly all respondents ($n = 20$) reported reviewing cover letters or statements of interest, while only seven reviewed letters of recommendation or contacted references.

Reviewers most frequently prioritized problem-solving ability ($n = 8$), followed by technical and analytical skills ($n = 6$). When evaluating cover letters or statements of interest, 95% of respondents looked for a clear connection between the applicant’s skills and the specific projects they planned to work on. They also emphasized the importance of professional tone and proper formatting. Fewer reviewers prioritized alignment of the applicant’s background with company goals (70%) or the inclusion of specific achievements (65%).

When reviewing letters of recommendation or speaking with references, all respondents valued letters that highlighted the applicant’s work ethic and communication abilities. They placed less emphasis on applicant’s punctuality or whether the letter was tailored to the specific position.

Interview. In this section, we examined how interviewers evaluate internship candidates and what they prioritize during the interview process. Most respondents (approximately 84%) said they do not share interview questions with candidates in advance. They most often asked about candidates’ prior research experience and proficiency with statistical software. In contrast, they rarely inquired about

leadership experience or how candidates navigate ethical dilemmas.

When discussing organizational fit, most interviewers ($n = 16$) frequently or always asked why candidates were interested in the organization. However, only a few ($n = 6$) asked about how candidates aligned with the company's culture and work environment. Questions related to bias or equity in assessment appeared in less than 20% of interviews. Additionally, only a small number of respondents reported using case studies (i.e., scenarios presented to applicants to examine how they would handle various situations; $n = 5$) or requiring applicants to complete live data analyses during the interview ($n = 2$).

Qualities. In this section, we identified the specific qualities and skills that supervisors prioritize when selecting and working with interns. Respondents most frequently selected the ability to analyze and interpret data ($n = 19$), verbal and written communication skills ($n = 11$), and critical thinking skills ($n = 11$) as essential for a successful intern. None of the respondents ranked client interaction skills or leadership skills among the top five most important attributes.

Respondents consistently emphasized technical competence. Most considered experience with statistical software essential ($n = 20$) and many ($n = 15$) highlighted the importance of data management and cleaning. While many viewed experience with machine learning or simulation studies as desirable, they did not consider these experiences essential.

When evaluating psychometric knowledge, respondents overwhelmingly identified classical test theory ($n = 18$) and item response theory ($n = 17$) as essential. They viewed generalizability theory and structural equation modeling as desirable but not required. All respondents valued descriptive statistics, while fewer prioritized abilities to conduct more advanced methods such as hierarchical linear modeling ($n = 7$) or non-parametric tests ($n = 3$).

In the area of test development and validation, respondents most often selected reliability ($n = 14$) and validity ($n = 12$) as essential. They placed less emphasis on equating ($n = 2$) and techniques for classroom assessment ($n = 1$). Although no respondents identified skills related to equitable assessment, social justice, or culturally responsive assessment as essential, many considered them desirable, particularly bias identification and mitigation.

Finally, 90% of respondents said they offered an internship to help students enter the profession, while 19% said they limited consideration to applicants pursuing careers directly aligned to their organization's work.

Logistics. In this final section, we examined the logistical aspects of internship experiences. Self-report data from mentors/supervisors were used to corroborate information from internship calls reviewed in Study 1 and to supplement missing logistical details.

Most internships lasted 8 weeks ($n = 15$) and required full-time commitment at 40 hours per week ($n = 16$). All respondents confirmed internships were paid, though compensation varied widely. Total pay ranged from \$5,000 to \$15,000, while hourly rates ranged from \$30 to \$45 per hour, typically paid in biweekly increments or as a lump sum at the end of the internship.

Remote internships were most common ($n = 12$), followed by in-person ($n = 6$) and hybrid ($n = 3$) formats, with hybrid positions often involving in-person attendance during the first or last week of the experience. Some organizations offered housing-related financial support, including travel and hotel costs. Support came in the form of monthly payments (\$700–\$1,200), or as a one-time payment ranging from \$2,000 to \$5,000.

Most internships did not include conference funding ($n = 12$), though some offered conference support unconditionally ($n = 2$) or conditionally ($n = 5$) based on proposal acceptance. Interns typically worked on a single project ($n = 13$) or on multiple research projects ($n = 8$), often balancing independent work ($n = 16$) with team collaboration ($n = 10$).

Nearly all respondents (~95%) expected a final deliverable, which included internal presentations ($n = 18$), conference proposals ($n = 14$), internal reports ($n = 13$), and/or research papers ($n = 5$). All respondents reported that interns had a very high degree of autonomy, though only a few ($n = 6$) allow interns to retain ownership of produced work (e.g., code). Many companies also supported international interns, typically contingent on an existing visa relationship with the intern's institution.

Finally, several respondents described unique experiences offered by the internship, such as attending a standard setting, technical advisory committee meetings, conferences, staff retreats, and happy hours, with one respondent even noting a whale-watching experience.

Study 3: Internship Survey—Students

After gathering perspectives from internship mentors and organizers, we sought to incorporate student viewpoints to provide a more holistic understanding of internship experiences in educational measurement. Student responses allowed us to capture how students navigate the internship process and what value they associate with participation, thereby adding depth to the broader picture of internship experiences, complementing the insights gained from Study 1 and Study 2.

Methods. We developed a survey targeting graduate students in educational measurement and related fields and administered it through Qualtrics. We designed the survey to take no more than 5 minutes to complete. To capture the experiences of students at different stages, we incorporated branching logic that directed respondents to items based on whether they were considering applying for an internship, had already completed an internship, or were currently completing one (see Appendix D).

Following IRB approval, we distributed the survey link via email to professional contacts, including members of the NCME Graduate Student Issues Committee (GSIC) and the Educators of Measurement SIGIMIE, in addition to sharing it on LinkedIn. We also encouraged recipients to forward the link to graduate students who were planning to apply for or had completed an internship, thereby relying on a snowball sampling approach consistent with earlier phases of the project.

Data collection ran from January 6, 2026 to January 23, 2026 with reminders distributed during that period. Of the 42 initial responses, we removed eight because they did not consent to participate. We then removed an additional six because respondents provided consent but did not complete any

items. The final analytic sample included 28 responses, 27 of which were complete.

Sample demographics. The sample consisted primarily of doctoral students ($n = 25$), with two master's students. The two part-time and 25 full-time students represented 15 unique institutions among those that chose to disclose their affiliation (6 opted not to report their institution). Most were in their third year of graduate study ($n = 9$), with program tenure ranging from 1 to 6 years. Eleven students indicated All But Dissertation (ABD) status, although two students were unsure of whether they were ABD.

Regarding internship participation, 5 students had completed 2 or more summer internships, 9 had completed 1 summer internship, 2 had completed 1 or more academic-year internships, 2 were currently participating in an internship, and 11 had never completed an internship. Sixteen respondents provided their internship site(s), with eight organizations from Study 1 represented in addition to seven internship sites submitted via open response. Sixteen students planned to or were currently applying for a summer 2026 internship; three were unsure, and nine did not intend to apply.

Results.

Pre-experience. To locate internship opportunities, students most frequently used LinkedIn ($n = 21$), followed by the NCME website ($n = 18$), faculty recommendations ($n = 17$), email communications ($n = 13$), other graduate students ($n = 12$), employer-specific websites ($n = 8$), and other sources ($n = 2$). Sixteen students reported searching for opportunities outside educational measurement, including positions in data science, policy, machine learning, and public health, among others.

When asked about their primary motivation for pursuing an internship, students most often selected gaining practical experience ($n = 24$), exploring potential career pathways ($n = 23$), and building a professional network ($n = 21$). Two respondents indicated financial considerations as a primary motivator.

Most respondents agreed that their graduate coursework adequately prepared them to obtain an internship ($n = 16$). However, a smaller subset expressed neutrality ($n = 7$) with three respondents disagreeing to some extent with this statement.

During and post-experience. Among students with internship experience, most characterized their work as research-focused ($n = 8$), while others described it as operational ($n = 3$) or a blend of both ($n = 3$). Overall, students reported substantial professional and personal benefits, with 79% indicating that their internship contributed to their professional development. Similarly, 71% reported increased confidence in their career goals, although 43% noted that the experience also prompted them to reconsider their intended career path.

Students also identified growth in skill development, with 79% reporting learning skills during their internship that they would not have acquired through their graduate coursework; a similar percentage felt that their graduate training had adequately prepared them to succeed in the internship environment. Students generally rated mentorship positively, with

71% describing the mentorship they received as valuable, although two students reported otherwise.

Despite variation in the nature and perceived quality of their experiences, all students stated that they would encourage other graduate students to pursue internship opportunities. However, about 64% believed that the benefits they gained were specific to their particular internship site rather than broadly representative of internships within educational measurement. When asked about potential employment with their internship organization, most students expressed interest ($n = 8$) while two said they would not consider it and three were unsure.

Discussion

In this three-part investigation, we examined graduate internships in educational measurement and assessment from the perspective of internship calls (Study 1), internship mentors and organizers (Study 2), and graduate students (Study 3). Together, these studies provide an account of the competencies sought by organizations, insights into internship supervision, and an understanding of the motivations and experiences of students navigating these opportunities. The integration of preliminary evidence across studies allows for a clearer understanding of expectations, barriers, and opportunities within the current landscape of educational measurement internships. Below, we first note connections between our findings and the foundational competencies outlined by Ackerman et al. (2024). We then synthesize findings in relation to our guiding research questions and discuss implications for graduate students, graduate programs, internship organizers and mentors, and the broader field. We view this work as a valuable first step toward developing criteria for evaluating internships within the field of educational measurement.

Connections and Misalignment with the Foundational Competencies

Domain 1: Communication and Collaboration. Findings from Studies 1 and 2 demonstrate strong alignment between internship expectations and foundational competencies within Domain 1. In Study 1, internship calls emphasized communication at multiple stages of the application and internship process. Application materials required students to communicate effectively by articulating their skills and interest in cover letters and presenting qualifications in CVs, and postings frequently identified written and oral communication skills as preferred or required qualifications. Beyond the application stage, many internship experiences culminated in a final product (e.g., internal report, conference presentation), providing opportunities to communicate technical work to varied stakeholders. This pattern is consistent with Ackerman et al.'s (2024) emphasis on presenting to, "...both technical and general audiences" (p. 9).

Study 2 findings further reinforced the importance of these competencies. Internships were described as collaborative, often involving close work with mentors or project supervisors. Mentors consistently emphasized the importance of communication skills during the application process, with all respondents reporting that they valued letters of recommendation or reference conversations that spoke to applicants' communication abilities, and 11 explicitly identifying written and verbal communication as essential for success. Once

selected, interns were expected to contribute to discussions and effectively present their work, with mentors highlighting these as key aspects of successful performance. These findings suggest that communication competencies are both expected of applicants and reinforced within internship experiences.

Domain 2: Technical, Statistical, and Computational Competencies. Findings also show strong alignment with Domain 2 competencies. In Study 1, internship postings frequently emphasized technical skills, particularly proficiency in statistical software, as required or preferred qualifications. Consistent with this, Study 2 respondents reported that interview processes often included questions about applicant's software proficiency, suggesting that these competencies are assessed during selection.

Mentors also highlighted the importance of technical skills for success within the internship itself. Nineteen respondents identified the ability to analyze and interpret data as essential, while twenty reported that experience with statistical software was critical for intern success. Skills related to data management and cleaning were also frequently emphasized, further underscoring the central role of applied data skills in internship work.

More advanced technical skills, such as machine learning, were described as desirable but not required. This distinction is consistent with Ackerman et al. (2024), who characterize these skills as more advanced and typically developed beyond foundational training. Taken together, these findings suggest that many technical competencies are expected at entry, while internship experiences provide opportunities to further refine and extend these skills through applied work.

Domain 3: Educational Measurement Competencies. Ackerman et al. (2024) identify five subdomains within educational measurement: (a) context (social, cultural, historical, and political); (b) validity, validation, and fairness; (c) theory and instrumentation; (d) precision and generalization; and (e) psychometric modeling. Subdomain A is described as overarching, while subdomain B undergirds the remaining areas.

Alignment with Domain 3 was uneven, reflecting both areas of strong overlap and notable gaps. In Studies 1 and 2, there was clear emphasis on competencies related to psychometric modeling (subdomain E) and validity and reliability (subdomain B). Internship postings frequently required knowledge of CTT and IRT, with general psychometrics and measurement theory listed as preferred qualifications. Study 2 respondents similarly identified these areas as essential for intern success, and skills related to test development and validation were commonly emphasized.

In contrast, other subdomains were less explicitly represented. Competencies related to context (subdomain A) including equitable assessment, social justice, or culturally responsive assessment were largely absent from internship expectations. Subdomain C (theory and instrumentation) was also not explicitly emphasized, although this may reflect the scope of the data collection, as survey items did not directly probe these areas. Similarly, competencies related to precision and generalization (subdomain D) were not prominently identified, which may also be attributable to the framing of survey questions.

Synthesis across Domains. Taken together, these findings suggest that many of the competencies identified across domains function not only as targets for development but also as prerequisites for obtaining internship opportunities. That is, competencies in communication, collaboration, and technical skills are often expected prior to the start of the experience, while internships themselves provide opportunities to further refine and extend these abilities in applied contexts. Findings from Study 3 further reinforce this duality, as students reported both that their graduate training prepared them to obtain and succeed in internships and that internships provided opportunities to develop skills not otherwise acquired through their coursework. This pattern positions internships as selective, semi-advanced training experiences rather than primary points of skill acquisition, while underscoring their unique contribution to professional development.

In relation to the foundational competencies outlined by Ackerman et al. (2024), this raises important considerations regarding the timing of competency development. If foundational competencies are expected before internship participation, rather than developed through it (as suggested by Ackerman et al., 2024), this may indicate a misalignment between how competencies are conceptualized in training frameworks and how they are operationalized in applied settings.

Implications for Graduate Students

Across all three studies, internships emerged as meaningful extensions of graduate training rather than peripheral professional experiences. Students overwhelmingly reported that internships contributed to their professional development, strengthened their understanding of potential career pathways, and expanded their professional networks.

Competitiveness and Preparation. Study 1 revealed that organizations frequently require or prefer proficiency in statistical software (particularly R and Python) in addition to knowledge of CTT and IRT, comfort with data analysis, and strong communications skills. Study 2 corroborated these expectations: mentors consistently reported that interns demonstrated considerable technical proficiency and that these skills were essential for productive engagement in project work. Importantly, results from Study 3 indicated that students generally felt prepared by their graduate coursework but benefited from acquiring additional skills during their internship; skills not easily gained through coursework alone.

Based on our initial investigation, we encourage students to do the following to be competitive applicants:

- Develop strong foundational competencies early in their training.
- Tailor application materials to clearly link their experiences to the internship responsibilities described in calls.
- Prepare to articulate their motivations, professional goals, and prior applied experiences, in addition to the skills they used or developed through these experiences, for cover letters and interviews.

Navigating Logistical and Structural Barriers. Study 1 identified considerable variation in stipend levels, work modalities (remote, hybrid, in-person), expected hours, and

overall compensation practices across internships. Study 3 suggested that students weigh these logistical and financial details when deciding where to apply. Two students explicitly described financial need as the primary driver of their decisions, suggesting that internships may serve as essential supplemental income for some. This dynamic underscores a potential access concern: inconsistent compensation structures may unevenly shape who can realistically participate.

Although natural variation in internship experiences may warrant differences in logistics such as stipend amounts, we recommend that students critically evaluate:

- Stipend adequacy and alignment with personal financial needs, especially for remote positions where compensation may not account for students' local cost-of-living realities.
- Access to conference funding and other professional development resources, recognizing that some internships offer financial support that reduces future expenses, while others require students to self-fund these opportunities.
- The timing and competitiveness of internship calls by monitoring postings early (e.g., late fall) and considering how variable deadlines, rolling review processes, and applicant volume may affect access to preferred opportunities. Although most application deadlines fall between January and March, students should remain diligent in identifying valued opportunities.

Meaningful Components of the Internship Experience.

Across our three studies, several features of internship experiences emerged as consistently influential for students' professional growth and satisfaction. Study 3 demonstrated that students valued internships not only for technical skill development but also for mentorship and professional networking. Study 2 showed that mentors perceived interns as contributing meaningfully to research and operational work, highlighting the applied and collaborative nature of internship tasks. Study 1 revealed that internship calls varied in the type of work offered and the extent to which organizations articulated mentorship structures or cohort models. Taken together, these findings suggest that the structure of internships, including the nature of the work, the quality of mentorship opportunities, opportunities for networking, and the presence of peer support, plays a fundamental role in shaping the value of the experience.

Based on our investigation, students should consider:

- The nature of project work, as internships vary in their emphasis on research, operational tasks, or a blend of both, as aligning the project focus with their skill-development needs and career goals can enhance the value of the experience.
- The opportunities for professional networking and socialization, including chances to engage with team members, participate in meetings, contribute to collaborative projects, or present findings, as these activities support long-term professional identity and connection to the field.
- The cohort structure and peer support available, as internships with multiple interns or those that encourage interaction among interns can enhance learning, reduce isolation (especially in a remote setting) and strengthen peer-based professional development.

Internship experiences by their very nature can differ in various ways, often due to factors such as location and size of the company/organization. Students should strive to be proactive and informed consumers of internship opportunities, recognizing that access can sometimes be shaped by structural, financial, and logistical constraints.

Implications for Graduate Programs and Faculty

Graduate programs are critical intermediaries in preparing students for internships, shaping both skill development and awareness of opportunities. Students in Study 3 indicated their coursework prepared them for internship participation, and mentors in Study 2 affirmed that interns typically arrived with adequate foundational competencies. However, the results across all three parts of our study show that early and intentional preparation remains essential. Programs may find it valuable to:

- Encourage enrollment in key methodological courses (e.g., CTT, IRT) prior to application cycles. Programs should transparently convey that the absence of these foundational courses often limits students' preparedness and substantially reduces their competitiveness as applicants.
- Integrate hands-on analytical or measurement tasks using real or authentic datasets so students can discuss these experiences in cover letters or interviews.
- Support students in articulating the relevance of graduate coursework to applied settings.
- Provide structured opportunities, such as colloquia or workshops, focused on preparing application materials, including interview practice sessions, cover letter writing feedback, and CV development.

Graduate faculty can also help students interpret the substantial differences across internship experiences revealed in Study 1, which ranged in duration, stipend amounts, other forms of compensation (e.g., conference support), work modalities, and deliverable expectations.

Addressing Information Asymmetries. Across Studies 1 and 3, a central theme was the absence of a comprehensive, centralized resource for measurement-related internships. Students relied heavily on LinkedIn, the NCME website, faculty emails, peer networks, or chance encounters to learn about internship opportunities. Graduate programs can play an important role in mitigating information gaps by:

- Maintaining updated lists of past internship placements or opportunities shared by alumni.
- Sharing internship repositories with students (e.g., American Statistical Association, NCME, Psychometricians Discord Community).
- Actively disseminating internship calls to current students.
- Encouraging engagement with professional networks (e.g., NCME, conferences, LinkedIn).
- Familiarizing students with opportunities outside traditional educational measurement settings (e.g., data science, policy, public health), which some students are already pursuing, per Study 3.

Implications for Internship Organizers and Mentors

Organizations offering internships make important contributions to students' career development and the broader measurement workforce. The findings from our three-part study highlight several areas where internship organizers and mentors can strengthen the clarity, structure, and equity of internship opportunities. Study 1 showed substantial variability in how organizations communicate internship expectations, including inconsistencies in required skills, stipend amounts, work modality, project descriptions, and mentorship structures. Study 2 revealed similar variability in mentorship preparation, interview practices, and expectations for intern contributions, despite mentors overwhelmingly describing their roles as rewarding and professionally beneficial. Study 3 further underscored the importance of mentorship quality for students' experiences. Together, these investigations show that alignment, or misalignment, between internship calls, interview practices, and on-the-ground experiences has meaningful implications for transparency, equity and professional development in the field. Despite natural variation across internship experiences due to inherent differences, it remains important to critically examine this variability and its implications.

Enhancing Transparency and Clarity. Organizations can improve applicants' fit and support more equitable access by offering clear and complete internship descriptions. Based on our findings, organizations should:

- More clearly distinguish required from preferred qualifications in internship calls.
- State stipend amounts, work modality expectations, anticipated deliverables, and project structure clearly and explicitly.
- Provide more specific descriptions of technical expectations, as common phrases such as "knowledge of," "familiarity with" or "experience with" encompass a wide range of possible competencies.
 - Additionally, offering explicit detail on broad topic areas such as "IRT," "CTT," "R," and "Python" would further increase transparency (e.g., equating, scoring, specific software packages).
- Describe mentorship formats, including expected frequency of meetings.
- Highlight unique professional development opportunities (e.g., conference funding, training opportunities, committee engagement).

Improving the clarity of internship materials can help students make informed decisions, supporting stronger alignment between applicant skills and organizational needs.

Strengthening Mentorship Structure. Mentors consistently reported that supervising interns was rewarding (Study 2) and students consistently rated mentorship as a central component of positive experiences (Study 3). However, only a small proportion of mentors received formal preparation for their role. To enhance the consistency and quality of mentorship, organizations should consider:

- Providing mentors with structured training or guidance outlining expectations.
- Offering orientation materials for both mentors and interns to support shared understanding of goals.

- Ensuring that mentors have adequate time allocated for supervision, as indicated by the average 9 hours per week reported in Study 2.

These steps can help organizations continue to foster supportive and developmentally meaningful internship environments.

Reconsidering Interview Practices. Study 2 found that interviews frequently emphasized technical skills and organizational interest while rarely addressing equity, ethics, or organizational culture. To help ensure organizational fit and to promote a more transparent interview process, organizations may wish to consider:

- Providing interview questions (or areas of discussion) to applicants in advance.
- Integrating questions that assess broader competencies such as collaboration, professional values, or ethical reasoning, if aligned with or relevant to the work.

Addressing Compensation. Study 1 revealed variation in stipend amounts across internships, including discrepancies between on-site and remote positions. Because students' cost-of-living context differs widely, and because some rely on internships for supplemental income, organizations should consider:

- Reviewing whether stipend amounts fairly reflect expected workload and modality.
- Offering additional financial support for conference participation to promote more financial value, especially if unable to increase stipends.
- Transparently communicating compensation details to reduce uncertainty for prospective applicants.

Ensuring that compensation aligns with student needs and internship expectations supports broader participation.

Implications for the Field

Findings across our three studies carry potential implications for the broader educational measurement community and for professional organizations like NCME. Study 3 showed that many internships pursued by students were not listed on NCME's website, suggesting limited visibility of the full internship landscape. Study 1 revealed that internship calls increasingly emphasize programming, analytics, and interdisciplinary skills typically associated with adjacent fields such as data science. Study 2 indicated that access to internships may depend on informal networks, variable institutional emphasis, and uneven dissemination of information. And, although few students described funding as their primary motivator, variability in stipend amounts (Study 1) and student responses about financial considerations (Study 3) indicate that compensation can play an important role in students' ability to pursue internships. Taken together, these results suggest that NCME has an opportunity to enhance its current internship resources by better reflecting the emerging areas of measurement work and providing timely information for students.

Based on these findings, NCME and the broader field should consider:

- Expanding the scope and visibility of internship listings beyond traditional "educational measurement" roles to incorporate applied analytics, data science, or mixed methods opportunities.

- Updating how internships are categorized and promoted to reflect the interdisciplinary nature of measurement work.
- Maintaining a centralized, publicly accessible repository of current internship calls, past internship calls, and past internship placements of members in addition to expanding available search criteria for the current NCME repository to include options such as stipend amount and skill requirements.
- Encouraging organizations to post stipend amounts, work modality, and financial supports clearly and consistently.

Limitations and Future Directions

Our work faces several limitations that must be considered when interpreting findings. Study 1 included a modest sample of internship calls from a single application cycle, which may not reflect opportunities across different years or sectors of the field. This limited representation was evidenced by reports of student internship placements in Study 3. Additionally, Study 2 included 22 mentors representing approximately 19 internship programs. Although the number of organizations offering graduate-level internships in educational measurement is relatively small, the sample limits the generalizability of mentors' experiences and perspectives. Further, the collection of internship logistics (e.g., modality, length of internship) via self-report may introduce inaccuracies. Finally, Study 3 included 28 students from at least 15 institutions, leaving many programs unrepresented and limiting diversity of institutional context in the results. Additionally, students with multiple internship experiences were not asked to reflect on differences in the benefits of those experiences, representing an area for future research.

Institutional representation in Study 3 in particular raises questions about how sampling processes and institutional culture shape internship engagement. Programs that emphasize academic career trajectories may not promote internships as strongly as programs oriented toward applied measurement careers, leading to uneven encouragement, guidance, and access across institutions. As evident in our results, benefits to skill development, mentorship, and networking from participating in an internship go beyond what a graduate program can offer. Therefore, even if students are training for an academic position, they can benefit from internship participation. Differences across programs may mirror informal pipelines (e.g., alumni hiring interns from their alma maters, hiring of interns from the same institution as previously successful interns) which may contribute to inequities in who learns about, applies to, and secures internship positions. As such, programs should reflect on how their advising structures, communication practices, and professional development expectations influence students' awareness of and access to internship opportunities. At the same time, this presents a greater need and opportunity for organizations like NCME to step in to promote the value of internships and showcase past experiences as well as future opportunities.

Study 1 was further constrained by the absence of a centralized internship repository. Without a single curated source, some relevant opportunities, especially those positioned in adjacent domains such as data science or public health, may not have been identified, even though students

reported seeking out and completing such internships. Studies 2 and 3 also represent snapshots of experiences and expectations during specific time periods and may not reflect future conditions, particularly as the field evolves in response to rapid changes in data science and artificial intelligence.

Future work should examine internship experiences, mentorship practices, and organizational decision-making using qualitative approaches such as interviews or focus groups to deepen understanding of the relational contextual dynamics that shape internships. Longitudinal research would also be valuable for assessing the long-term influence of internships on career trajectory, skill development, and engagement in the field. As internship structures and expectations continue to change, especially in interdisciplinary directions, ongoing research will be essential to understanding how internships can remain accessible and responsive to emerging professional demands.

Conclusion

Our three studies offer a thorough first step in the examination of the graduate internship landscape in educational measurement. The findings suggest that foundational competencies function as pre-requisites for internship participation rather than outcomes of the experience. Internships continue to provide meaningful opportunities for applied skill development and career exploration. However, students' access to these experiences may be uneven due to variability in information dissemination, compensation practices, mentorship structures, and institutional supports.

Graduate programs can prepare students for internships, helping them build the foundational skills required for participation and mediating inequities in access to information and opportunities. Organizations offering internships can strengthen transparency in their calls, clarify mentorship expectations, and adopt practices that promote equitable selection processes. At the field level, NCME and similar professional associations have the opportunity to expand and modernize internship listings to reflect the increasingly interdisciplinary nature of measurement work and to support access for students across institutions.

As the field evolves, coordinated efforts among programs, organizations, mentors, and professional associations will be essential to ensure that internships remain accessible, meaningful, and aligned with the profession's future needs. Together, these studies provide a starting point for future work on how the field prepares emerging professionals and how internships can continue to contribute to the measurement workforce.

Conflicts of Interest Statement

We have no conflicts of interest to disclose.

ORCID

Autumn N. Wild, M.A.  <https://orcid.org/0000-0002-5975-0856>
 Brian C. Leventhal, Ph.D.  <https://orcid.org/0000-0002-6480-2016>
 Joseph M. Kush, Ph.D.  <https://orcid.org/0000-0003-0183-494X>

References

- Ackerman, T. A., Bandalos, D. L., Briggs, D. C., Everson, H. T., Ho, A. D., Lottridge, S. M., Madison, M. J., Sinharay, S., Rodriguez, M. C.,

- Russell, M., von Davier, A. A., & Wind, S. A. (2024). Foundational competencies in educational measurement. *Educational Measurement: Issues and Practice*, 43(3), pp. 7–17. <https://doi.org/10.1111/emip.12581>
- Alahmadi, S., Goodman, J., Jackson, J., Li, X., & Runyon, C. (2023, April 15). Internships in the measurement profession: A discussion among organizers, mentors, and students. [Conference session]. National Council on Measurement in Education Annual Meeting, Chicago, IL.
- Anderson-Cook, C. M., Hamada, M. S., Moore, L. M., & Wendelberger, J. R. (2017). Statistical mentoring at early training and career stages. *The American Statistician*, 71(1), pp. 6–14. <https://doi.org/10.1080/00031305.2016.1200491>
- Bay, L., Cao, Y., Kuang, H., Marion, S., & Wild, A. N. (2025, April). *Bridging graduate students and practice: Internships in educational measurement*. [Moderated discussion]. National Council on Measurement in Education Annual Meeting, Denver, CO.
- Cheng, Y., Pei, B., Filonczuk, A., & Thu Le, A. (2024). Commentary: A data-driven analysis of recent job posts to evaluate the foundational competencies. *Educational Measurement: Issues and Practice*, 43(3), pp. 39–44. <https://doi.org/10.1111/emip.12630>
- Gaikwad, H. V., Pandey, S., Divekar, R., & Patil, D. (2024). Comparative analysis of student performance in offline vs. online internships: Impacts on skill development, professional growth, and academic outcomes. *Journal of Engineering Education Transformations*, 38, pp. 188–195.
- Hora, M., Chen, Z., Parrott, E., & Her, P. (2020). Problematizing college internships: Exploring issues with access, program design and developmental outcomes. *International Journal of Work-Integrated Learning*, 21(3), pp. 235–252.
- Jenkins, Z., & Yao, C. (2025). A critique of race-neutrality discourse among internship benefits: A critical discourse analysis. *Education + Training*, 67(2), pp. 171–183.
- Kang, J., & Girouard, A. (2022). Impact of UX internships on human-computer interaction graduate students: A qualitative analysis of internship reports. *ACM Transactions on Computing Education*, 22(4), pp. 1–25. <https://doi.org/10.1145/3517132>
- Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Association of American Colleges and Universities.
- National Council on Measurement in Education. (2025, March). *Graduate student internships & fellowships*. <https://www.ncme.org/students/resources-students/internships-fellowships>
- Rowe, A., Randall, J., Banjanovic, E., & Spingler, L. (2024, November 13). *The lifecycle of an internship: Insights from the experts* [Webinar]. National Council on Measurement in Education Annual Meeting. https://youtu.be/wRJuG1QA0I?si=9A27uteH_kWpYIL3
- Schnoes, A. M., Caliendo, A., Morand, J., Dillinger, T., Naffziger-Hirsch, M., Moses, B., Gibeling, J. C., Yamamoto, K. R., Lindstaedt, B., McGee, R., & O'Brien, T. C. (2018). Internship experiences contribute to confident career decision making for doctoral students in the life sciences. *CBE—Life Sciences Education*, 17(1), pp. 1–14. <https://doi.org/10.1187/cbe.17-08-0164>
- Wilson, K., Hawk, A., & Miller, M. (2022). Implementing a hybrid internship during the COVID-19 pandemic. *New Review of Academic Librarianship*, 29(1), pp. 77–96. <https://doi.org/10.1080/13614533.2022.2082989>

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Supporting File: emip70045-sup-0001-SuppMat.docx
Supporting File: emip70045-sup-0002-SuppMat.docx
Supporting File: emip70045-sup-0003-SuppMat.docx
Supporting File: emip70045-sup-0001-SuppMat.docx