Spring 4-2014

Survey on the Prevalence of Measurement in Undergraduate Psychology Curricula across the United States

Kevin A. Dahlman
University of Nebraska-Lincoln, kevindahlman85@gmail.com

SURVEY ON THE PREVALENCE OF MEASUREMENT IN UNDERGRADUATE PSYCHOLOGY CURRICULA ACROSS THE UNITED STATES

by

Kevin A. Dahlman

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Arts

Major: Educational Psychology

Under the Supervision of Professor Kurt F. Geisinger

Lincoln, Nebraska

May, 2014

SURVEY ON THE PREVALENCE OF MEASUREMENT IN UNDERGRADUATE PSYCHOLOGY CURRICULA ACROSS THE UNITED STATES

Kevin A. Dahlman. M.A.

University of Nebraska, 2014

Adviser: Kurt F. Geisinger

For several years, especially with the increasing use of high-stakes assessments, the high demand for qualified testing specialists has far exceeded the low supply, naturally resulting in a severe shortage of quantitative psychologists. One possible solution for combatting this imbalance is to increase the exposure of psychometric methods within national curricula for educating undergraduate students of psychology. Therefore, the current study seeks to assess the prevalence of psychological measurement across national undergraduate psychology curricula, since it has not been evaluated for roughly 15 years. Six hundred fifty psychology department chairpersons across the United States were emailed invitations to complete a 22-item survey pertaining to the prevalence of measurement within their own undergraduate psychology curriculum. Results indicated that psychometrics is becoming slightly more prevalent as compared to 15 years ago; however, measurement coverage is still failing to reach the recommended guidelines set by national conventions for the undergraduate education of psychology students.

ACKNOWLEDGEMENTS

This thesis project was conducted as a partial requirement for a master's degree and completed under the supervision of the Buros Center for Testing staff. The author would like to thank his advisor, Kurt F. Geisinger, whose assistance made the successful completion of this project possible. Gratitude must also be given to Jesica L. Jonson for providing assistance during the study's developmental stages, as well as Greg W. Welch in accepting the invitation to become an independent reader. Moreover, acknowledgement should be extended to the invaluable input provided by Carina M. McCormick, Kate N. Sirota, and Jonathan Sikorski during the manuscript preparation stages.

TABLE OF CONTENTS

The field of psychological measurement, commonly known as psychometrics, is currently in need of attracting prospective students, as evidenced by the imbalance between the high demand for psychometricians and far smaller supply of testing specialists. Clay (2005) attributed this imbalance to a number of reasons, including: (a) the general unawareness of this field, (b) the lack of quantitative ability in interested applicants, and (c) the negatively skewed age distribution of currently qualified instructors. In fact, according to the American Psychological Association (APA, 2007), "50% of all members of APA Division 5 (Evaluation, Measurement, and Statistics) are 55 years of age or older" (p. 5). Nationally, only about 30 testing specialists graduate from doctoral programs each year (APA, 2007), and many of these graduates are hired by large testing companies that often offer higher salaries than those typically found within academia-career paths (Herszenhorn, 2006). This severe shortage resulted in the development of an APA task force with the primary objective being to increase the number of quantitative psychologists, specifically measurement specialists.

In 2006, Leona Aiken chaired this specific APA task force with the assistance of eight other members (APA, 2007). This committee was assembled with the mission of three interwoven intentions: (a) addressing the deficient supply of qualified students, (b) increasing exposure of the field to potential students, and (c) evaluating the current state of training opportunities available in the field of quantitative psychology. They attributed the small supply of qualified applicants to the general lack of understanding of quantitative psychology, resulting from a deficiency of well-organized information as well as a limited availability of training resources. In fact, according to the report of this APA task force (2007), less than 25 PhD programs in quantitative psychology existed

throughout North America, and collectively summed across all these programs, only awarded approximately 30 doctoral degrees per year. Other domains within psychology (e.g., clinical, developmental, social, and cognitive) produce hundreds of doctoral degrees every year. For example, according to the report from the APA task force, there were 1221 PhD degrees awarded in clinical psychology across the country in 2003 alone. To complicate matters, most undergraduate psychology majors in the United States are not well-oriented in mathematical reasoning through classic undergraduate curricula (APA, 2007; APA, 2011; Halpern, 2010). Understanding of this context can guide the exploration of possible solutions to increase the number of qualified applicants seeking graduate degrees related to quantitative psychology, most notably in psychometrics.

The members of the aforementioned APA task force (2007) highlighted two typical ways in which students generally gain awareness of quantitative psychology: (a) through the small social network of individuals currently exposed to the field, or (b) by randomly stumbling upon the discipline due to mere chance. Because these two avenues alone inevitably lead to a limited exposure of quantitative psychology, the task force deemed it necessary to target the undergraduate student population through a variety of ways. They explicitly recommended reaching out to students by providing information about the field through the following opportunities: introductory-level psychology courses, department chairs, Psi Chi publications, conferences for psychology undergraduates, and regional psychological associations. Furthermore, the APA task force also suggested increasing psychometric exposure in undergraduate curricula in psychology by placing more emphasis on quantitative courses and/or topics, such as

statistics and measurement. This last recommendation is the primary focus for the current project.

Before progressing, it is necessary to clarify the definition of psychological measurement, or psychometrics as it is sometimes known. Hubley and Zumbo (2013) defined psychometrics as "a field of study that focuses on the theory and techniques associated primarily with the measurement of constructs as well as the development, interpretation, and evaluation of tests and measures" (p. 5). Through the general process of measurement, one is basically quantifying an observation. However, psychological studies typically involve unobservable characteristics, or *constructs* as they are commonly referred, making it quite difficult to obtain numerically accurate representations of the construct under investigation. This natural complexity for measuring the directly unobservable provides the foundation of psychometric theory, which underlies all disciplines of psychology, basic and applied.

Broadly speaking, psychologists can be categorized as practitioners, researchers, and teachers of psychology, or some combination of these. Through the use of sound psychometric theory, researchers typically develop assessments that are designed to be used by the practitioners of psychology for diagnostic or other purposes. Furthermore, these instruments may also be applied clinically to evaluate the effectiveness of a specific intervention before and after a treatment plan has been introduced to the client. For example, scores on the Beck Depression Inventory – II (BDI-II; Beck, Steer, & Brown, 1996) could assist in the initial diagnosis of depression. Moreover, once this diagnosis has been made, subsequent scores on the BDI-II may also be utilized to assess the effectiveness of a particular treatment for alleviating the client's depression. Aside from

these clinical settings, the practice of psychometrics can also be applied to a vast array of other contexts as well, especially in the fields of education and business, which are discussed below.

Psychological measurement is an increasingly essential component for high-stakes decision making within nearly every imaginable domain. For instance, many businesses are primarily concerned with maximizing utility, which essentially results in the selection of those applicants who will be successful while rejecting those that would otherwise fail. Thorndike (1982) elaborated more fully on this idea through the use of a utility matrix in his textbook, *Applied Psychometrics*, an edition that followed his well-known volume, *Personnel Selection* (Thorndike, 1949). This utility matrix maximizes the two conditions listed above, ultimately leading to profitability for businesses using this model. The model also seeks to minimize financial losses through the other two conditions: (a) rejecting those who would have succeeded, and (b) selecting those who end up failing. Therefore, the continued use of psychological measures that facilitate utility-based decisions, is vital from a fiscal perspective.

In keeping with Thorndike's (1982) notion of maximizing utility in corporate settings, the use of psychometrics has perennially been common practice for nearly all academic domains, from preschool through the highest levels of education. Ultimately, every teacher needs some way to assess the extent of their students' achievement with respect to attaining curricular objectives. Although these objective-driven achievement levels are unarguably constructs, many teachers do not identify themselves as practitioners of psychometrics. Nevertheless, many teachers use a criterion-referenced approach for assigning grades, consistently developing and administering classroom

assessments designed to measure student achievement. In addition, standardized assessment is another commonly used form of psychometrics in education, especially for normative purposes in K-12 settings to compare students within grade levels. Moreover, albeit controversially, results from these standardized measures have also been used to make high-stakes decisions during teacher evaluation processes, affecting teachers' professional lives and development (e.g., salary, promotion, tenure). For more information regarding the use of standardized assessments, such as accountability measures in education, please see Linn (2000).

The far-reaching influence of psychometrics is also evident in higher education. For example, the Graduate Record Examination (GRE) is commonly used by most graduate programs to assist in their selection of graduate school applicants. Analogous to corporate hiring decisions that seek maximum utility, selection committees are eager to accept individuals who will complete their degree in a timely manner, as well as become engaged and successful in the program. Furthermore, even if such individuals do successfully attain their degree objectives, many disciplines require additional testing as part of the certification process for various advanced practitioners, such as doctors, psychologists, lawyers, and financial analysts. Licensure testing of these professionals ensures the general public that persons managing their health, legal, and fiscal matters are qualified to do so, thus providing our community with a sense of security.

However, despite the diversity of psychometric applications, the demand for quantitative psychologists is much higher than the current supply. Reasons for this shortage, as noted above, may be attributable to: (a) an underqualified applicant pool not meeting the intellectual demands of the discipline, (b) the small number of training

programs available, and (c) students' limited exposure as undergraduates to the field of measurement. Therefore, it might be argued to increase awareness of this area of study at the undergraduate level in order to recruit prospective students. However, prior to suggesting an increase of this exposure, a baseline of the current measurement coverage within the nationwide curriculum of undergraduate studies in psychology is needed. Explicitly, how much of this curriculum is currently devoted to psychometrics? This general research question is the primary motive for the present study.

Very little research has focused on the extent of measurement coverage for undergraduate education in psychology, although two previous studies have examined the prevalence of overall quantitative exposure for the doctoral training of psychology students (Aiken, West, Sechrest, & Reno, 1990; Aiken, West, & Millsap, 2008).

Aiken, West, Sechrest, and Reno (1990) mailed surveys to 222 APA accredited psychology programs throughout North America. These surveys sought information on the extent and quality of each school's graduate-level statistical, methodological, and measurement instruction and training during the 1985/86 academic year. Aiken and her colleagues (1990) obtained a relatively high response rate of 84% (186/222) from the chairpersons of these departments. Among these responses, the researchers observed that approximately one-third of graduate programs in psychology had no faculty members trained primarily in statistics or measurement. Furthermore, less than half of all programs even offered measurement courses within their own department, often requiring PhD students to take psychometric-related courses from other departments. Additionally, only one fourth of these programs judged their students to be competent in basic methods of reliability and validity, while even fewer felt their students were skilled in classical test

theory or item analysis. Aiken, West, Sechrest, et al. (1990) concluded by these findings that the reported lack of competencies in measurement training at the graduate level may certainly be detrimental to future psychological research. Negative consequences could include selections of inappropriate designs to test theories, failure to gather the most salient data for specific research inquiries, and the reaching of incorrect conclusions due to misanalyses of data, as well as impacts on clients that are potentially equally problematic. Aiken and her colleagues (1990) thus found it paramount to revamp national curricula for graduate education in psychology, and called for a reevaluation of the progress at a future date.

Aiken, West, and Millsap (2008) completed this reevaluation twelve years later in 1998 through a replication study. Questionnaires were mailed to all psychology department (234 total) offering at least one APA accredited Ph.D. program. The methods employed were virtually identical to the previous Aiken, West, Sechrest, et al. (1990) study and received a slightly better response rate of 86% (201/234). Aiken, West, and Millsap (2008) observed that students' exposure to psychological measurement had indeed increased, with nearly two-thirds of programs offering at least one graduate-level course in measurement as compared to the less-than half from the 1986 evaluation. However collectively, these courses still had very low coverage of key topics such as classical test theory, item response theory, and test construction. Moreover, reported measurement competencies of graduate students also improved slightly, yet still remained relatively low (e.g., fewer than half could assess an instrument's reliability and only one fourth could use effective validation methods, construct tests, or even analyze items). Aiken and colleagues (2008) also noted an increase in shortened or abbreviated modules

on measurement topics within the national curriculum, a factor that resulted in an increase of departments offering at least some measurement coverage. Nevertheless, the authors maintained the same conclusion from their previous study; both the curriculum coverage and student competencies in psychometrics were still inadequate. They also noted that many of the current faculty teaching measurement, as well as other areas of quantitative psychology, were approaching the age of retirement, even furthering the gap between the low supply and high demand for quantitative psychologists. In addressing this widening gap, Aiken and colleagues (2008), as well as her colleagues from the 2007 APA task force, suggested that one possible approach was to increase measurement exposure at the undergraduate level with the intent to recruit future students.

One of the best ways for attracting potential measurement specialists may be through the infusion of basic psychometric coursework into undergraduate psychology curricula. According to standard conventions of the discipline, when assessing and summarizing nationwide curricula, dialogue at a conference specifically convened to provide this discussion takes place, followed by public dissemination reports of the meeting to the entire psychological community. Brewer (1997) provided a chronological review of these assemblies regarding undergraduate education in psychology, starting with the Cornell Conference in 1951 (Buxton et al., 1952) and progressing through to the St. Mary's Conference in 1991 (Lloyd & McGovern, 1992).

In fact, all of the national conventions on undergraduate education in psychology, including the most recent one in 2008 (Halpern, 2010), supported an instructional model taught through the paradigm of a scientific discipline housed within a liberal arts college. However, the St. Mary's Conference report (Lloyd & McGovern, 1992) stressed that one

nationwide curriculum for undergraduate psychology students would not be ideal for every school, but rather departments should have some autonomy for constructing their own curriculum. Nevertheless, the committee stressed that all psychology departments, however, should follow these six universal goals: (a) give attention to human diversity, (b) provide a breadth and depth of knowledge, (c) make students methodologically competent, (d) ensure students have practical experience to apply their knowledge, (e) teach effective communication skills, and (f) make students sensitive to ethical issues (Brewer, 1997). In addition to these six broadly defined goals, the St. Mary's committee also recommended all psychology curricula to have four basic levels of classes, beginning with Introductory Psychology, progressing through methodology and content courses, and culminating in an integrated, applied experience (e.g., internship, capstone course, and/or undergraduate research project). If departments address these six aforementioned goals and provide students with adequate coursework on the basic class types, then institutions should have some freedom in constructing their own, unique curriculum. Additionally, and in alignment with the current study, the St. Mary's committee recommended that a course in psychometric methods be a part of every curriculum's methodology course sequence.

Subsequently, Perlman and McCann (1999) sought to evaluate undergraduate psychology curricula through an extensive review of 400 course catalogs from the 1996-97 academic year and compared their observations to the suggested model from the St. Mary's Conference in 1991. Using catalogs in place of surveys ensured a 100% response rate, which eliminated the possibility of nonresponse bias, achieving consistency in the data so as to not rely on a respondent's memory or one's own subjective interpretation.

Perlman and McCann (1999) observed that only nine percent of four-year institutions actually required that psychology majors take a course in psychometrics, and roughly 60% of the 400 institutions assessed even offered such a course. Therefore, the overwhelming majority of psychology departments were not conforming to the curriculum guidelines from the St. Mary's Conference, namely failing to include an undergraduate course in psychometrics, as was highlighted under the umbrella of the methodology course recommendations (Lloyd & McGovern, 1992). Messer, Griggs, and Jackson (1999) also found comparably low results pertaining to measurement coverage in undergraduate-psychology curricula.

As in the Perlman and McCann (1999) study, Messer et al. (1999) also used course catalogs, however sampled from the 1993-94 academic year with respect to the number of *degree* offerings. Therefore, if an institution offered more than one type of psychology degree, it would be counted more than once. This methodological approach could account for the slightly discrepant findings because departments that offered multiple psychology degrees would need to have more course offerings, as well as varying degree requirements, to accommodate these different specializations. Messer et al. (1999) documented that although 82% of psychology *degrees* offered a course in measurement, just 19% of degree options actually required a psychometrics course to be taken. However, when studying the general psychology degree option, the researchers observed that only 13% made the psychometric methods course a requirement, which is much closer to the 9% from the Perlman and McCann (1999) study.

Similar to the St. Mary's Conference in 1991, another conference convened at the University of Puget Sound in 2008 to investigate the current status of undergraduate

education in psychology (Halpern, 2010). Collectively, 80 psychologists participated in this meeting. These individuals were selected primarily based on their excellence and expertise for instructing undergraduate students of psychology. The 80 participants were divided into nine "working groups" with each group receiving a specific topic to have an elaborate discussion with regards to undergraduate education in psychology.

Halpern (2010) noted that roughly 88,000 students graduated with a baccalaureate degree in psychology in the United States, coupled with over one million undergraduate students that enroll in psychology courses every year. In fact, Introduction to Psychology was the second most commonly taken undergraduate course in the nation, behind only Basic English Composition (McGovern et al., 2010). Due to these voluminous numbers, the entire discipline needs to be accountable for the students that are being produced, thereby generating the need for the 2008 National Conference on Undergraduate Education in Psychology. Throughout this conference, the working groups discussed the notion of *psychologically literate citizens* with the development of these citizens being the primary concern for educating undergraduate students of psychology.

Psychologically literate citizenship, as defined by McGovern et al. (2010), "describes a way of being, a type of problem solving, and a sustained ethical and socially responsive stance towards others" (p. 21). Upon achievement of this desired outcome, individuals graduating with a psychology degree should be able to use their acquired skills to enhance the communities in which they live. However, in order to ensure that all psychology students in the United States become *psychologically literate*, the 2008 National Conference called for the development of a more unified, core curriculum for educating undergraduates in psychology.

Dunn and colleagues (2010) were charged with the task of evaluating the content of nationwide curricula experienced by students pursuing a baccalaureate degree in psychology. Just as previous conferences asserted, Dunn et al.'s (2010) working group reaffirmed that undergraduate education in psychology continues to be taught via the scientific method and housed within the liberal arts tradition. Through this paradigm of scientific psychology, a common core for all curricula would be essential in order to ensure that psychology graduates meet certain standards. Any implementation of a single, unified curriculum must be broad, balanced, and flexible to allow for the competing beliefs amongst educators so they can cater to the distinctive needs of each student as well as their own institution.

Dunn and colleagues (2010) listed a number of recommendations for the implementation of a loose, core curriculum, specifically: having all students take Introductory Psychology before any other psychology coursework, requiring students take classes pertaining to the scientific method, integrating diversity and ethics into the program, providing students with opportunities to apply their newly-acquired knowledge, and requiring students to partake in a capstone project. The utilization of this unified, yet flexible, approach for educating psychology undergraduates not only offers quality assurance to future employers, but also increases the "likelihood of potential success and personal satisfaction no matter what paths students follow after earning a baccalaureate degree" (p. 60). Therefore, educators must cater to all undergraduate students of psychology, whether or not they have aspirations of attending graduate school.

Appleby (2000) produced a list of five basic skills that employers desire from those earning a baccalaureate degree in psychology: (a) social skills; (b) personal skills;

(c) communication skills; (d) information gathering and processing skills; and (e) numerical, computer, and psychometric skills. As a result, Appleby noted that even the general workforce expects some degree of competency in psychometric knowledge when hiring college graduates who majored in psychology, providing more credence to the need for the current study to evaluate measurement coverage in the undergraduate curricula. Furthermore, Landrum et al. (2010) elaborated on the personal skills expected of these students, and stated that in becoming *psychologically literate citizens*, an individual should be able to critically evaluate the source and validity of any piece of information. Acquiring this learned, information-evaluation skill has become even more critical through the progression of the internet age and may be enhanced through increased exposure to measurement-related issues. However, with respect to knowledge of psychological measurement, one needs to determine what skills constitute competency in measurement concepts.

Analogous to what McGovern (2010) termed *psychological literacy*, Lambert (1991) coined the idea of *measurement literacy*. She stated that individuals achieve *measurement literacy* through the acquisition of four types of knowledge: (a) knowledge of the assumptions regarding the quantification and recording of observations; (b) knowledge of general statistics (e.g., frequency, probability, scales of measurement, central tendency, variability, etc.); (c) knowledge pertaining to the concepts of reliability and validity; and (d) knowledge about measurement error. The shared responsibility of ensuring that *measurement literacy* is achieved ultimately lies in the hands of students, instructors, and administrators. However, in order for students to receive quality

instruction in psychological measurement, the teachers themselves must first be well-rounded in psychometric theory.

Regardless of teachers' theoretical orientations and academic disciplines, measurement is an integral component for every level of education and must be used to assess student performance. In fact, Wise, Lukin, and Roos (1991) postulated that roughly one third of teachers' time was spent on assessment-related activities. These researchers mailed surveys to 825 primary and secondary teachers from two different school districts, a rural district as well as an urban one, regarding instructor training in testing and measurement. Of these 825 delivered surveys, 397 teachers responded with 47% indicating that their formal undergraduate, as well as postgraduate, training in testing and psychometrics was inadequate. Most of these teachers reported self-educating themselves in measurement through trial-and-error processes from their own teaching experiences. The vast majority of respondents also felt testing and measurement to be essential for any instructor, and nearly all self-rated their measurement abilities to be fairly high. Wise, Lukin, and Roos (1991) concluded their article by urging administrators to increase measurement training in teacher-education programs by including at least one measurement course in the curriculum. If future teachers, especially those in psychology, are properly educated in psychometric theory, it allows their own students to acquire similar measurement skills.

The extent of measurement in the current undergraduate curriculum for educating psychology students in the United States has not been evaluated exclusively since the Messer et al. (1999) study. However, McKelvie (2000) provided a slightly more recent assessment of the coverage of psychological testing, albeit within the Canadian

curriculum. He was able to obtain data from the entire population of Canadian undergraduate psychology programs (N = 51) through an examination of course catalogs as well as writing to psychology department chairpersons. McKelvie (2000) observed that 76% of psychology departments offered at least one course in testing; however, only 14% listed the class as a requirement in their curriculum. These results were somewhat similar to those mentioned earlier by Messer et al. (1999), who found that 82% of psychology degree options in the United States offered at least one psychometric class, while only 19% required it.

McKelvie (2000) provided two possible explanations for the general inadequacy of testing-related courses. The first reason stemmed from a lack of coverage in measurement training at the graduate level of education (e.g., Aiken, West, Sechrest, et al., 1990; Aiken, West, & Millsap, 2008). They believed that administrative decision makers may have operated on the erroneous assumption that these topics would be taught in graduate school, and therefore, may not have deemed it necessary to provide this type of instruction for undergraduate students. Since current students inevitably become the future face of any field of study, this general lack of exposure would likely lead to a reverse-snowball effect, resulting in less and less coverage through the coming years. A second explanation for this inadequacy of measurement may have resulted from the artificial divide between experimental and correlational methods.

Typically, correlational designs are generally perceived as the most appropriate methodology in measurement-related research. In fact, Guilford (1946) gave testimony to the importance of these kinds of designs in a famously simple quote stating that "a test is valid for anything with which it correlates" (p. 429). However, true experimental

designs incorporating randomization are generally preferred in psychological research as they may establish causal relations; this is why they are typically referred to as the "gold standard" when employing a scientific approach to psychological studies. Alternatively, a correlational study does not try to formally establish causality, but simply observes the extent to which variables are related to one another. From a psychometric standpoint, these two broadly defined types of research differ in how they handle individual differences; experiments try to control for these differences, while correlational studies attempt to measure them (McKelvie, 2000). Cronbach (1957) as well as Shadish, Cook, and Campbell (2002) along with many other publications explain the similarities and differences between correlational and experimental research designs. After providing two plausible explanations for the lack of measurement exposure to undergraduate students of psychology, McKelvie (2000) also listed several arguments in favor of supplying such psychometric knowledge.

McKelvie (2000) first contended that the study of psychometrics allows for increased attention to be given to constructing and/or selecting quality dependent variables that are measured with minimal error. Incorporating well-designed, dependent measures into a study allow for greater accuracy in estimating the true effect of any independent variable. Secondly, McKelvie (2000) noted that random assignment (a key component of the experimental design) is all-too-often impossible to implement in social science research, and as a result, correlational (e.g., psychometric) methods are frequently the only feasible option for many research situations. Therefore, these methods should be taught more thoroughly to undergraduate psychology students. The third argument focused on the growing prevalence of meta-analytic and estimation techniques where the

test validation processes of psychometric theory frequently employ these procedures. For example, meta-analyses were used to provide validity evidence for various standardized assessments, such as the GMAT (Talento-Miller & Rudner, 2008), the MCAT (Julian, 2005), as well as the GRE (Kuncel, Hezlett, & Ones, 2001), just to name a few such instances. McKelvie (2000) further argued that increased exposure to psychological measurement would allow students to better distinguish between four scales of variables (i.e., nominal, ordinal, interval, and ratio) since many commonly employed psychometric methods do not rely heavily on the assumption that a dependent variable be a scale-based measure. Finally, the last two intertwined arguments presented by McKelvie (2000) dealt with applications and issues associated with using test scores as a basis for executive-decision making. Exposing students to even the most basic psychometric properties should help psychology graduates intelligibly partake in future debates regarding social and ethical implications for the growing use of high-stakes testing.

From McKelvie's (2000) arguments, one can reasonably conclude that if students have acquired a well-informed knowledge base in psychological measurement, then they are far more likely to become *psychologically literate citizens*. The 2008 National Conference on Undergraduate Education in Psychology concluded that the ultimate goal of any undergraduate program in psychology should be that students achieve *psychological literacy* (Halpern, 2010; McGovern et al., 2010; APA, 2011). Since an education in psychometrics would most definitely increase these chances, the psychological community should consider that measurement courses be included in the curricula for obtaining a baccalaureate degree in psychology. However, psychometric principles are needed to evaluate the extent to which psychology departments provide

students with proper measurement instruction. For a guide to evaluating undergraduate psychology programs, please see the eight standards highlighted by Dunn, McCarthy, Baker, Halonen, and Hill's (2007) publication, "Quality Benchmarks in Undergraduate Psychology Programs."

The current study's overall objective was to evaluate the extent of psychometric coverage currently offered through undergraduate psychology curricula. In addressing this purpose, an electronically-based web survey was utilized to assess the quantity, as well as the quality, of the measurement-related exposure experienced by undergraduate students of psychology across the United States via the following five research questions:

1. How much of the national undergraduate curriculum in psychology is devoted to psychological measurement?
2. What are the credentials and ranks of instructors who typically teach psychometric courses?
3. Do psychology departments often outsource their responsibility for educating undergraduate students in psychometrics? If so, is there an available market for providing measurement instruction via online classes or online learning modules?
4. What measurement topics do psychology departments find most important for undergraduates to learn?
5. How competent in various psychometric topics are recent graduates who obtain a baccalaureate degree in psychology?

Method

Instrument

The cross-sectional survey administered in the current study was delivered to psychology department chairpersons through the Qualtrics web-based software (Qualtrics, 2012). This instrument was derived directly from the five research questions regarding the prevalence of measurement in undergraduate psychology curricula across the United States. The finalized, response-adaptive measure required participants to provide answers to a minimum of 14 items with a maximum of 22, contingent upon an individual's responses to the questions presented. A copy of all 22 items can be found in Appendix A.

The first four questions on the instrument were demographic in nature, inquiring the name of the respondent's academic institution and department, as well as his/her affiliation within the psychology department and whether or not he/she had ever taught a course in psychological measurement. Following these demographic questions, respondents were also asked for the average number of individuals per year (over the last five years) graduating with a major in psychology. The next two questions were more directly aligned with the research questions under investigation.

Specifically, these two subsequent items inquired if the respondent's psychology curriculum required that a measurement class be taken, and whether or not his/her department offered a class in psychometrics. If one's answer to the latter question was "no," then respondents skipped the next four items. On the other hand, if his/her response was "yes," then the survey inquired about the measurement course(s) provided through the department. These contingent four items addressed the number of

measurement courses offered, how often they are taught, the rank of the instructor who taught these course(s) most often, and the field of his/her graduate degree. The survey then redirected participants back to the universally administered questions.

Items 12 and 13 inquired whether other departments within the same university system also offered measurement-related courses for which undergraduate students were encouraged to enroll. If no measurement classes were recommended to be taken outside the department, then the next question was skipped; this subsequent item simply requested the name(s) of other departments offering measurement course(s). All remaining items on the instrument dealt with students' understanding of specific psychometric principles, as well as the curriculum's course offerings of other content domains within the field of psychology that provide cross coverage for measurement-related topics.

To gauge the overall degree of measurement competency for recent psychology graduates, all respondents were supplied a list of 12 general topics pertaining to psychological measurement and requested to select those areas he/she believed to be important for undergraduate students of psychology to understand. Specifically, the following 12 psychometric topics were presented as well as a thirteenth respondent-provided "other" option: reliability, test validity, legal/ethical issues in testing, test construction, item writing/analysis, classical test theory, test bias, use of tests in selection, ability testing, achievement testing, personality testing, and attitude/survey measurement.

After revealing one's subjective views of important measurement aspects, each respondent was also queried on their interest for out-sourcing his/her institution's psychometric education via distance learning. First, respondents were requested to list

the names of courses within their own department that cover measurement-related topics, but are not specifically a psychometric course. Furthermore, participants were also asked if their departments would consider offering online measurement courses through a third party institution, and/or if they had interest in providing their students with supplemental, online learning modules pertaining to psychometric topics. If a respondent revealed an interest in offering distance education, then he/she was prompted to provide those particular measurement-related topics (the same 12 topics from the preceding paragraph) for which he/she would find most beneficial for offering online learning modules. Additionally, participants were requested to select, from a different list of 12 options, the content courses (i.e., introduction to psychology, research methods, statistics, personality, development, cognition, social psychology, history and systems, abnormal psychology, counseling, educational psychology, and industrial/organizational, as well as an "other" response option) for which the supplemental materials would be desired.

After one's interest of distance education for psychometrics was revealed, or possible lack thereof, all respondents were finally taken to the instrument's last substantive item. Ultimately, the questionnaire concluded with an inquiry on psychology major graduates' understandings of measurement-related topics. Specifically, this Likert-based item asked participants to rate, on a four-point scale (unprepared, somewhat unprepared, somewhat prepared, and prepared), their undergraduate students' average degree of understanding on each of the 12 previously mentioned measurement-related topics. Once this 12-part item was completed, the survey reached completion and respondents were once again thanked for their time and participation.

Sampling Procedures

The target population for this study consisted of psychology departments within academic institutions in the United States offering a bachelor's degree in psychology. The sampling frame was constructed through the College Navigator (nces.ed.gov/collegenavigator; U. S. Department of Education (ED), Institute of Education Statistics (IES), & National Center for Education Statistics (NCES), 2012), which is an internet instrument typically used for individuals desiring to pursue higher education, providing potential students with the ability to screen the vast array of academic institutions through various search criteria. The College Navigator's data are obtained directly from the NCES's Integrated Postsecondary Education Data System (IPEDS), and in accordance with the Higher Education Act of 1965, every academic institution that participates in the federal student financial aid program is required to submit data to IPEDS. As a result of this federal statute, the coverage error between the sampling frame obtained from the College Navigator and the target population should be negligible.

Being a web survey, this project relied on the acquisition of email addresses for psychology department chairpersons, which were collected through each individual's associated university website. The first sample of email invitations yielded a relatively small, undesirable response rate, thus requiring a second, similar recruitment phase in order to reach an acceptable number of completed surveys. In creating the sampling frame from which both samples were selected, the search capabilities of the College Navigator identified the target population of four-year institutions offering a bachelor's degree in psychology. To create this listing, the current study used the following three

search filters, with the selected option in parentheses: “Institution Type” (4-year), “Level of Award” (bachelor’s), and “Browse for Programs” (psychology). The use of these search criteria resulted in a collective list of 1387 institutions.

To ensure that colleges of all sizes from every U. S. were represented, a combined explicit and implicit stratification design was incorporated (Piazza, 2010; Kish, 1965). In the explicitly stratifying stage, the aforementioned search criteria were used to yield a listing of schools within each U. S. state, including the District of Columbia, resulting in 51 total strata. After the study’s first sampling iteration, 435 institutions of higher education (based on the number of members in the House of Representatives) across the United States were chosen to receive survey invitations. Specifically, the number from each state was derived directly from the number of federal districts apportioned to that particular state according to the 2010 U. S. Census (U. S. Census Bureau, 2010). For example, Illinois had 18 representatives in 2010 and Nebraska had three; therefore, 18 schools were selected from Illinois and three from Nebraska.

After explicitly creating all 51 files, every data set was then separately exported into Microsoft Excel, where each spreadsheet was implicitly sorted in descending order with respect to total student population size. Once these schools were ordered, a systematic sample was obtained using a fixed selection interval approach (Piazza, 2010) always starting with the selection of the state’s most populous bachelor-granting institution. These interval sizes varied across states and were determined by dividing the total number of schools within each state by the number of federal districts partitioned to that state. This simple division resulted in a numerical index (usually with decimal values), for which the remainder was dropped leaving a whole number, *k*. Once this *k*

value was determined for each state, every k^{th} institution, starting with the most populous school first, was chosen from the Excel spreadsheet. However, only public and private (not-for-profit) colleges and universities were selected, intentionally omitting all private, for-profit institutions. An abridged example for one of these 51 spreadsheets, only revealing information about the school names from the state of Minnesota, can be viewed in Appendix B; invitations were sent to those institutions that are highlighted.

After all 435 four-year institutions were selected, these colleges and universities were then combined into another Excel spreadsheet containing the following five variables: state, school name, size of total student body, institution type (public or private), and the school's internet address. All web addresses were then visited and individual searches of these webpages yielded the inclusion of three more variables: a link to the psychology department's website, name of the department chairperson, and his/her email address. Assuming the school websites were navigational friendly, the department chair was typically found through a search conducted from the institution's main website for the keywords "psychology department chair." Usually, email addresses for the chairpersons were obtained through this search process as well, which were subsequently recorded into the spreadsheet. However, if the desired information was unattainable through this generalized process, then department webpages were investigated further, using methods specific to the features of the individual website. Once this finalized list of email addresses was obtained, survey invitations could then be delivered.

Through the use of Qualtrics (Qualtrics, 2012), the original email invitations, which were designed so that informed consent was given by clicking the survey link,

were distributed on the morning of October 26, 2012. A copy of this document is found in Appendix C. If a recipient failed to respond to this initial request, then a reminder email was also delivered three weeks later during mid-November (Appendix D), followed by a final reminder at the end of November (Appendix E) if still no response had been received after the first two requests. Since the first sample of 435 email recipients failed to yield an acceptable number of responses, all the aforementioned procedures were duplicated at the beginning of the following spring semester. Specifically, the entire College Navigator sequence was performed again, this time for attaining 215 more email addresses. For this second iteration, the same original email invitations were distributed on the morning of February 21, 2013, with a reminder email to participate during mid-March, followed by a final reminder notice in mid-April. Once all responses were received, the data were then ready to be analyzed.

Results

Collectively, across both samples, a total of 226 surveys were initiated with 181 being fully completed, resulting in an overall response rate of 27.8% (181 surveys completed/650 delivered invitations). Consequently, this yielded a sample size of N = 181 completed cases to be used for all subsequent analyses conducted through IBM SPSS Statistics 21. However, the College Navigator identified a finite target population of 1387 academic institutions offering a bachelor's degree in psychology across the United States, thus constituted 13% of the entire target population (181/1387).

In addition to the survey responses, data were also collected regarding demographic information pertaining to institutional type (public or private) and presence of a graduate program within the department. Two different, two-leveled independent variables were created to compare the presence of a graduate program. The first variable compared psychology departments with a graduate program to those without one; the second variable only compared those offering graduate degrees, examining the difference between those that are and those that are not APA accredited.

Across the entire study, 39 independent-samples *t* tests and 15 chi-square tests for independence were analyzed, where the former was used to evaluate scale variables and the latter for categorical ones. Explicitly, these 39 *t* tests were used on 13 scale measures (i.e., each of the 12 perceived student competencies and a composite score summing these 12 student ratings) for each of the three demographic variables (i.e., institution type, yes/no grad school, and APA/not-APA accredited programs). Collectively, using the same three independent variables, the 15 chi-square tests evaluated five categorical variables pertaining to psychometric courses: instructor ranking, requirement of a course,

offering of a course, regularity of course offerings, and number of courses offered. Although, after a Bonferroni correction was applied to control for the accumulation of experiment-wise type I error rates, none of the 54 inferential comparisons revealed any significant differences between the three aforementioned demographic variables. A few of these comparisons achieved marginal significance, but failed to make sense intuitively, so therefore are not included in the analyses.

How much of the nationwide, undergraduate curriculum in psychology is devoted to measurement? Descriptive results from the current survey indicated that 25.4% of psychology curricula require undergraduate students to take a course in measurement. An additional 21.5% of departments simply recommended a psychometrics class be taken without requiring it, resulting in most undergraduate curricula (53.0%) having no clear preference with respect to students enrolling in a measurement course. However, the vast majority of departments surveyed (80.1%) offer at least one course in psychometrics, leaving approximately 20% of psychology curricula (36/181) not offering a measurement course. Specifically, 9.9% of departments surveyed used to offer a measurement course but no longer do, 5.0% have not considered offering such a class, and another 5.0% of departments do not, but would like to offer a measurement course.

Of the 145 psychology curricula that provided undergraduate students with the opportunity to enroll in a measurement course within the department, 115 (63.5%) offered one psychometric class, 21 (11.6%) provided two such courses, and nine (5.0%) offered three or more measurement-related classes. Results also showed that when measurement courses are available within departments, they are typically offered less

frequently than every semester. Of the 145 psychology departments providing at least one psychometric course, these classes are offered by 35 departments (19.3%) every semester, 70 departments (38.7%) every academic year, 26 departments (14.4%) every other year, and less than every other year (7.7%) by 14 psychology departments. Ten institutions reported that measurement classes were offered through other departments, with educational psychology and statistics departments being the most commonly referenced (seven and three, respectively). Departments other than psychology that were cited by two or less respondents as providing undergraduates the opportunity to enroll in a measurement class included education, human development, industrial/labor relations, biometry, community mental health, social work, and criminology.

What are the rankings and credentials of instructors who typically teach psychometric courses? From the 145 psychology departments offering measurement classes, 122 (84.1%) departments typically employed tenured or tenure-track faculty as course instructors, which was nearly even across all three ranks: assistant professor (n = 41; 28.3%), associate professor (n = 40; 27.6%), and full professor (n = 41; 28.3%). The remaining 15.8% of psychology departments with measurement-course offerings for undergraduate students primarily employed adjunct faculty (n = 13; 9.0%) or instructors/lecturers (n = 10; 6.9%) with this teaching responsibility.

Even though a specialized graduate degree in psychometrics would arguably be most relevant for a measurement instructor, it was only the ninth most referenced field of study for the 145 departments that offer a course in psychological measurement. Specifically, only six (about 4.1%) of these institutions actually employed instructors with a graduate degree in psychometrics. By far the degree earned by most measurement

instructors was clinical psychology, representing employment by 50 of the 145 institutions (34.5%). Counseling and industrial/organizational psychology were both tied for second in the ranking of most commonly attained graduate degrees of measurement instructors, constituting employment from 16 departments (11.0%) for each domain.

Do psychology departments often outsource their undergraduates' education in measurement? The survey items that addressed this research question differentiated between those departments outsourcing their psychometric education within the same university and those departments that outsourcing via distance learning from an institution not affiliated with their own. With respect to the former type of outsourcing, the majority of respondents (n = 152; 84.0%) reported that measurement courses are not offered through any other department within their institution; 17 (9.4%) departments were not sure and 12 (6.6%) stated that such classes were offered through other departments. Of the institutions that offered a measurement course outside of psychology, only three departments were cited multiple times, namely educational psychology (n = 7), statistics (n = 3), and education (n = 3).

Across the 47 departments not offering a psychometric class, only one department said they would consider offering an entire course devoted to psychological measurement via distance learning from an institution other than their own; meanwhile, 19 departments were unsure if they would offer such an online course. Greater interest was revealed, however, for offering smaller online modules pertaining to specific psychometric topics designed to supplement other psychology courses within one's own curriculum. Specifically, 14.9% (n = 27) of the 180 institutions providing a response to this item displayed no interest at all for offering supplemental, online measurement modules, while

78 respondents (43.3%) provided some interest and 75 participants (41.7%) were not sure if their departments would consider offering these supplemental materials.

Collectively, 153 psychology departments were at least open to providing online measurement modules for one or more psychometric topics. Of these 153 institutions, interest was shown for the following topics with the number of departments advocating a particular topic in parentheses: reliability (114), test validity (113), item writing and analysis (87), legal/ethical issues in testing (83), test bias (82), test construction (81), attitude and survey measurement (66), personality testing (65), achievement testing (61), ability testing (57), use of tests in selection (55), and classical test theory (41). The frequency for each measurement topic listed above is depicted in the Pareto chart in Figure 1.

Generally, online measurement modules were cited as being beneficial for the following course subjects of psychology (listed in order of decreasing frequencies): research methods (116), personality (86), industrial/organizational psychology (62), introductory statistics (60), abnormal psychology (56), educational psychology (41), counseling psychology (34), introductory psychology (30), social psychology (29), developmental psychology (18), cognitive psychology (17), as well as history and systems of psychology (16). A visual summary of these frequencies can be viewed in Figure 2. There was also some interest revealed for other psychology courses that were not explicitly stated on the instrument; interest was revealed for offering supplementary learning modules for the classes of assessment/measurement (6), clinical psychology (4), bio/neuropsychology (3), and senior research projects (1).

What measurement-related topics do psychology departments find most important for undergraduate students to learn? The Pareto chart depicted in Figure 3 displays general measurement topics listed on the horizontal axis as well as each category's corresponding frequency on the y-axis. All psychology departments, aside from one, identified the general psychometric topics of reliability and test validity (both had 180 votes out of 181 possible votes) to be most important for undergraduate students to understand. In order of importance, the frequencies and percentages, located in parentheses out of 181 total, for the other 11 measurement topics were: legal/ethical issues (n = 152; 84.0%), test bias (n = 146; 80.7%), test construction (n = 132; 72.9%), personality testing (n = 130; 71.8%), attitude and survey measurement (n = 120; 66.3%), ability testing (n = 109; 60.2%), achievement testing (n = 109; 60.2%), item writing and analysis (n = 106; 58.6%), use of tests in selection (n = 103; 56.9%), classical test theory (n = 78; 43.1%), and "other" (n = 12; 6.6%). The only psychometric topic not chosen by at least half of the respondents was classical test theory, yet was still viewed as important to 43% of psychology departments. "Other" written-in topics provided by respondents included: statistics as they relate to testing, clinical/counseling assessments, neuropsychological testing, basic testing, decision making, forensic applications, intelligence testing, introduction to item response theory, and multicultural issues. However, open to subjective interpretation, many of these responses could have also been categorized under some of the explicitly stated topics.

How competent are recent psychology graduates in measurement? Using the same response options as the previous paragraph, participants were also asked to judge the measurement competencies of their recent psychology graduates on a four-point

Likert scale, using the response options of unprepared, somewhat unprepared, somewhat prepared, and prepared. The average Likert rating across all departments (N = 181) for each psychometric topic can be found in Table 1. This table displays the descriptive statistics (means and standard deviations) for each judged competency, as well as a summary score, across all respondents.

These perceived measurement competencies of recent graduates aligned closely with the same measurement topics that departments also found to be most important for undergraduate students of psychology to understand. In fact, a Spearman's rank order correlation between these judged competencies and the number of departments deeming each psychometric topic as important was exceptionally high, $r_s(10) = .89$, $p < .05$. Overall, these data suggest that undergraduate psychology curricula are reflective of the measurement domains that psychology departments consider to be most important.

Discussion

The findings from this cross-sectional survey revealed that the current state of measurement coverage within the undergraduate curricula of psychology is not aligned with the guidelines set by the 2008 National Conference (Halpern, 2008), which explicitly recommended a measurement course as part of the undergraduate course progression. In fact, fewer than half of the departments even recommend, and considerably less require, that a psychometric class be taken. This discrepancy from the national guideline, though, does not seem to be an issue related to the understaffing of measurement specialists. If this were indeed a lack-of-supply concern, it would seem logical that more than one institution would have displayed interest in offering an online course devoted primarily to measurement. Nevertheless, appeal was substantially greater for smaller, online modules covering psychometric topics within core subjects of psychology. The enormity of this discrepancy for preferring modules over courses, coupled with the general lack of a measurement class requirement, suggests that many departments consider other psychology courses to be more essential for undergraduate students to learn. However, this contention does not necessarily implicate the widespread belief that psychometrics is not an important topic, as evident by the large amount of interest displayed for online learning modules pertaining to measurement topics.

As noted by Stoloff et al. (2010), the average number of undergraduate classes offered across all U.S. psychology departments is 39; and typically, 7 to 10 of these classes are mandatory, constituting a majority of the average 38 credit-hour requirement of psychology majors. Therefore, there may not be much perceived flexibility on behalf of administrators for including a required measurement course in the curriculum,

especially if they deem it desirable to allow students some degree of choice in their course selection. However, this survey failed to provide any such questions to address this line of reasoning for not including a measurement course, which may be one of its limitations.

Quite possibly the greatest limitation of the current study, though, resided in the relatively low response rates (27.8%) obtained from department chairpersons. As with any type of survey design, it is extremely important that one acquires a high response rate since large sample sizes, especially in terms of proportions, are always desirable for making any inferential decisions about a population. By procuring a greater proportion of responses from the target population, the likelihood of error due to nonresponse is minimized; specifically, the larger the response rate obtained, the greater the magnitude of nonresponse bias required to manifest in inaccurate inferences. Therefore, from a generalization standpoint, it is imperative that any survey research obtains as high a response rate as possible. Although, without knowledge of the true parameter values, it cannot be determined with absolute certainty if the number of completed responses (N = 181) from the current survey was indeed large enough, making the current assessment of the external validity of these results a highly subjective one.

In their *Survey Methodology* textbook, Groves, Fowler, Couper, Lepkowski, Singer, and Tourangeau (2009) state three possible reasons that nonresponse may occur at the instrument level: (a) the instrument was not delivered successfully, (b) the respondent simply refused to participate, and (c) the respondent was incapable of participating. Generally, these three sources of nonresponse may affect the quality of survey findings in distinct ways. Aligning well with the first two sources of nonresponse,

and likely contributing to the low response rate of this study, Tourangeau, Conrad, and Couper (2013) contend that failed deliveries coupled with the ease of declining participation from email invitations have been continuously growing concerns for the survey researchers. For example, a single typo in an email address will result in a failed delivery, while such a minute error in a mailing address can likely be detected by the postal worker and still delivered successfully. Furthermore, although email invitations may offer convenience and ease in accessibility through a clickable URL, it can also be just as easy to click the delete button, especially due to the tremendous amount of spam email that often pollutes an inbox. Moreover, a survey request may never even receive the recipient's attention if the invitation is filtered into his/her junk folder. Consequently, methods for maximizing response rates of web surveys via email invitations should be further investigated in future research.

Shih and Fan (2008) used meta-analytic techniques through an examination of 37 publications from 1995 to 2006 that directly compared web- and mail-based surveys. After controlling for undeliverable invitations, they found that response rates were generally higher for hard copy surveys sent through the mail (45%) than for those electronically delivered over the internet (34%). However, this overall tendency was not observed to generalize to college populations, whose response rates were roughly 3% greater on web surveys than mailed ones, although no exact percentages for each mode were provided. Compared to the results provided by Shih and Fan (2008), the current survey's 27.8% response rate was roughly six percent less than the overall average as reported by these researchers for similar web-based instruments. Although, Shih and Fan

(2008) mentioned that using follow-up reminders did lead to significantly higher rates of responding, consistent with this study's decision for including two follow-up reminders.

One of the publications that was actually part of the aforementioned meta-analysis, Shannon and Bradshaw (2002), compared response rates, response times, and costs related to the two methods of distribution. Similar to the current study, Shannon and Bradshaw's (2002) sample consisted solely of faculty members, but specifically from colleges in the southeastern United States (N = 377). Once the initial surveys were distributed, follow-up reminders were sent two weeks, as well as four weeks, after the original invitations to participate were delivered. They found that mailed surveys obtained a 47% response rate, while the emailed surveys had a lower response rate of 32%, closer to the 27.8% observed from the current study. In regards to initial survey requests, a strength of the electronic mode of delivery is that emails had significantly quicker response times than their hard-copy counterparts. In fact, before the first hard copy was returned in the mail, roughly 80% of the completed electronic surveys had already been received. Aside from response time, another desirable characteristic of emailing surveys is the cost. Total postage costs for hard-copy questionnaires were more than triple the amount of money necessary for electronically administered surveys ($394.68 and $124.41, respectively).

Putting aside the response-rate limitation, results from this survey suggest that psychometric coverage has increased slightly since the Perlman and McCann (1999) study; however, it still remains low. The current study observed that 80.1% of psychology departments offered at least one measurement course to undergraduates, and a fourth of departments required that a psychometrics course be taken. This is still an

increase from the results obtained by Perlman and McCann (1999), in which 60% of psychology curricula offered a measurement class, and only nine percent actually required it. Optimistically speaking, psychometric coverage has improved over the past sixteen years; nevertheless, more exposure is still necessary, especially since most employers of those earning a bachelor's degree in psychology expect their new hires to have at least some degree of competency in psychological measurement (Appleby, 2000).

The desire to incorporate more psychometric exposure to undergraduate students of psychology is predicated on the assumption that there actually are enough instructors qualified to teach such courses. However, as previous publications have highlighted (e.g., Aiken et al., 1990; Aiken et al., 2008; APA, 2007; Clay, 2005; Herszenhorn, 2006), the overall shortage of competent measurement teachers has consequently limited the availability of measurement courses being offered. Currently, most measurement instructors are not trained specifically in psychometrics (only 4.1%), but rather the most commonly obtained graduate degree of these teachers is clinical psychology (34.5%). This could be the result of the abundant supply of clinical psychologists, for which more than a thousand doctoral degrees are awarded per year (APA, 2007). This abundance, although, may not affect psychology departments focusing primarily on undergraduate education; however for those institutions offering graduate programs, the need for measurement specialists to teach psychometric courses becomes more pronounced. While clinical psychologists seem to be the modal course instructors, many of these individuals may indeed have sufficient measurement training in their background, and thus be qualified to teach such courses.

Even though the general undergraduate exposure to psychometrics has gradually increased over the last several years, it has still not reached the explicit recommendations set by the 2008 National Conference on Undergraduate Education in Psychology. As stated earlier, one of the primary purposes for organizing this convention was to hold undergraduate educators accountable for the 88,000 baccalaureate degrees awarded in psychology every year through a more unified, yet flexible, national curriculum (Halpern, 2010). Specifically, within every department's methodology course sequence, this committee suggested that measurement be an integral part of every psychology curricula. Through incorporating at least one measurement class into every curriculum, it would most assuredly result in an increased production of high-quality, *psychologically literate citizens*, which should be the ultimate goal for any department providing an undergraduate education in psychology. Of course it is possible that measurement serves a component of one or more classes, but the respondents' appraisals of their students' capabilities indicate that their instruction could be enhanced.

Achieving *measurement literacy* (Lambert, 1991) through an acquisition of the four types of measurement knowledge (i.e., quantifying observations, general statistics knowledge, reliability/validity, and measurement error) should also be incorporated into the production of *psychologically literate citizens*. The successful acquisition of these measurement aspects should result in an enhanced understanding for the research methods typically used by all social sciences. Although, the only competency ratings greater than 3.0 (somewhat prepared) were reliability and validity, implying that most departments believed their students' comprehension of other measurement topics was inadequate. These other psychometric areas should not be neglected in undergraduate

psychology curricula since they also play a fundamental role in many of the methodologies used to advance the scientific study of psychology.

Parallel to Lambert's (1991) description of the four aspects underlying *measurement literacy*, McKelvie (2000) provided numerous arguments supporting the deliverance of psychometric knowledge to the undergraduate student body of psychology. For instance, enhancing one's measurement education would inevitably result in increased exposure to: (a) the development and selection of appropriate dependent measures; (b) more practical research methods since correlational (e.g., psychometric) designs are often more feasible than experiments; (c) meta-analyses that are regularly used by test validation procedures and are being found more frequently in the literature; (d) the four scales of measurement and an enhanced ability to differentiate between them; and (e) the applications and ethical/social issues associated with the growing use of high-stakes testing. These powerful arguments proposed by McKelvie (2000) provide a strong rationale supporting the inclusion of more measurement-related content to undergraduate students of psychology.

If psychology majors are to become *psychologically literate citizens*, it is paramount that more measurement-related topics be incorporated into all the undergraduate curricula of psychology. However, in order to increase measurement familiarity, there needs to be a combined, collective effort on behalf of all parties involved. Namely, these stakeholders that must take action are administrators, teachers, as well as students of psychology. Initially, the decision to increase psychometric exposure should start with departmental faculty approving a revised curriculum involving greater measurement exposure. Subsequently, responsibility then flows to the instructors

of psychology to teach these methods in an effective manner. Ultimately though, accountability resides within the students, who must make an active decision to learn the psychometric material being presented to them.

The future for the scientific study of psychology hinges on what is being taught through the undergraduate curricula, and we must not forget that "today's students are tomorrow's psychology educators, researchers, practitioners, and consumers of psychological knowledge" (Landrum et al., 2010; p. 160). Therefore, active measures must be taken to increase measurement competency among undergraduate students of psychology.

References

Aiken, L. S., West, S. G., Sechrest, L., & Reno, R. R. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of PhD programs in North America. *American Psychologist, 45*(6), 721-734.

Aiken, L. S., West, S. G., & Millsap, R. E. (2008). Doctoral training in statistics, measurement, and methodology in psychology: Replication and extension of Aiken, West, Sechrest, and Reno's (1990) survey of PhD programs in North America. *American Psychologist, 63*(1), 32-50.

American Psychological Association. (2007). *Report of the Task Force for Increasing the Number of Quantitative Psychologists.* Retrieved from http://www.apa.org/research/tools/quantitative/quant-task-force-report.pdf

American Psychological Association. (2011). Principles for quality undergraduate education in psychology. *American Psychologist, 66*(9), 850-856.

Appleby, D. C. (2000). Job skills valued by employers who interview psychology majors. *Eye on Psi Chi, 4*(3), 17.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II (BDI-II). San Antonio, TX: Pearson.

Brewer, C. L. (1997). Undergraduate education in psychology: Will the mermaids sing?. *American Psychologist, 41*(1), 65-71.

Buxton, C. E., Cofer, C. N., Gustad, J. W., MacLeod, R. B., McKeachie, W. J., & Wolfe, D. (1952). *Improving undergraduate instruction in psychology.* New York, NY: Macmillan.

Clay, R. (2005). Too few in quantitative psychology. *Monitor on Psychology, 36*(8). 26-28.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12,* 671-684.

Dunn, D. S., Brewer, C. L., Cautin, R. L., Gurung, R. A. R., Keith, K. D., McGregor, L. N., … Voigt, M. J. (2010). The undergraduate psychology curriculum: Call for a core. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (pp. 47-61). Washington, DC: American Psychological Association.

Dunn, D. S., McCarthy, M. A., Baker, S., Halonen, J. S., & Hill IV, G. W. (2007). Quality benchmarks in undergraduate psychology programs. *American Psychologist, 62*(7), 650-670.

Groves, R. M., Fowler, Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Guilford, J. P. (1946). New standards for test evaluation. *Educational and Psychological Measurement, 6,* 427-438.

Halpern, D. F. (Ed.). (2010). *Undergraduate education in psychology: A blueprint for the future of the discipline.* Washington, DC: American Psychological Association.

Halpern, D. F., Anton, B., Beins, B. C., Bernstein, D. J., Blair-Broeker, C. T., Brewer, C. L., … Rocheleau, C. A. (2010). Principles for quality undergraduate education in psychology. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (pp. 47-61). Washington, DC: American Psychological Association.

Herszenhorn, D. M. (2006, May 5). As test-taking grows, test-makers grow rarer. *The New York Times*, pp. A1, A19.

Higher Education Act of 1965, 20 U.S.C. § 1094 *et seq.* (1965). Retrieved from http://nces.ed.gov/ipeds/about/

Hubley, A. M. & Zumbo, B. D. (2013). Psychometric characteristics of assessment procedures: An overview. In Geisinger, K. F. (2013). *Theory and methods of educational measurement* (pp. 5-21). Washington, DC: American Psychological Association Books.

Julian, E. R. (2005). Validity of the Medical College Admission Test for predicting medical school performance. *Academic Medicine, 80*, 910-917.

Kish, L. (1965). *Survey sampling*. New York, NY: John Wiley & Sons, Inc.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127*(1), 162-181.

Lambert, N. M. (1991). The crisis in measurement literacy in psychology and education. *Educational Psychologist, 26*(1), 23-35.

Landrum, R. E., Beins, B. C., Bhalla, M., Brakke, K., Briihl, D. S., Curl-Langager, R. M., … Van Kirk, J. J. (2010). Desired outcomes of an undergraduate education in psychology from departmental, student, and social perspectives. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (pp. 145-160). Washington, DC: American Psychological Association.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16.

Lloyd, M. A. & Brewer, C. L. (1992). National conferences on undergraduate psychology. In A. E. Puente, J. R. Matthews, & C. L. Brewer (Eds.), *Teaching psychology in America: A history* (pp.263-284). Washington, DC: American Psychological Association.

McGovern, T. V., Corey, L., Cranney, J., Dixon Jr., W. E., Holmes, J. D., Kuebli, J. E., … Walker, S. J. (2010). Psychologically literate citizens. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (pp. 9-27). Washington, DC: American Psychological Association.

McKelvie, S. J. (2000). Psychological testing in the undergraduate curriculum. *Canadian Psychology, 41*(3), 141-148.

Messer, W. S., Griggs, R. A., & Jackson, S. L. (1999). A national survey of undergraduate psychology degree options and major requirements. *Teaching of Psychology, 26* (3), 164-171.

Perlman, B. & McCann, L. I. (1999). The structure of the psychology undergraduate curriculum. *Teaching of Psychology, 26*(3), 171-176.

Piazza, T. (2010). Fundamentals of applied sampling. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd ed., pp. 139-168). Bingley, UK: Emerald Group Publishing Ltd.

Qualtrics. (2012). Qualtrics software. Provo, UT: Qualtrics. Retrieved from: www.qualtrics.com

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin Company.

Shannon, D. M., & Bradshaw, C. C. (2002). A comparison of response rate, response time, and costs of mail and electronic surveys. *The Journal of Experimental Education, 70*(2), 179-192.

Shih, T. & Fan, X. (2008). Comparing response rates from web and mail surveys: A meta-analysis. *Field Methods, 20*(3), 249-271.

Stoloff, M., McCarthy, M., Keller, L., Varfolomeeva, V., Lynch, J., Makara, K., … Smiley, W. (2010). The undergraduate psychology major: An examination of structure and sequence. *Teaching of Psychology, 37*, 4-15.

Talento-Miller, F., & Rudner, L. M. (2008). The validity of Graduate Management Admission Test scores. *Educational and Psychological Measurement, 68,* 129-138

Thorndike, R. L. (1949). *Personnel selection*. New York, NY: John Wiley and Sons, Inc.

Thorndike, R. L. (1982). *Applied psychometrics*. Boston, MA: Houghton Mifflin.

Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys.* New York, NY: Oxford University Press.

U. S. Census Bureau. (2010). United States Census 2010: Apportionment data. Retrieved from http://2010.census.gov/2010census/data/apportionment-data-text.php

U. S. Department of Education, Institute of Education Statistics, National Center for Education Statistics. (2012). *College Navigator* (IPEDS 2011-2012, Fall Collection). Retrieved from http://nces.ed.gov/collegenavigator/

Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education, 42*(1), 37-42.

Table 1

Judged Measurement Competencies of Graduates Obtaining a Bachelor's Degree in Psychology

Topic	Valid *N*	*Mean*	*SD*
Reliability	173	3.37	0.56
Validity	173	3.24	0.61
Legal/Ethical Issues	169	2.95	0.75
Test Bias	171	2.78	0.85
Attitude/Survey Measurement	166	2.73	0.73
Personality Testing	168	2.65	0.81
Ability Testing	167	2.39	0.85
Achievement Testing	167	2.34	0.83
Test Use & Selection	164	2.33	0.77
Test Construction	167	2.30	0.83
Item Writing & Analysis	166	2.27	0.87
Classical Test Theory	168	2.02	0.89
Summary Score	158	31.27	6.21

Note. These average Likert ratings are based on a four-point scale, where *1* = unprepared, *2* = somewhat unprepared, *3* = somewhat prepared, and *4* = prepared. These ratings are listed in order from highest to lowest degree of perceived competencies of recent graduates obtaining a bachelor's degree in psychology across all sampled departments in the United States (*N* = 181).

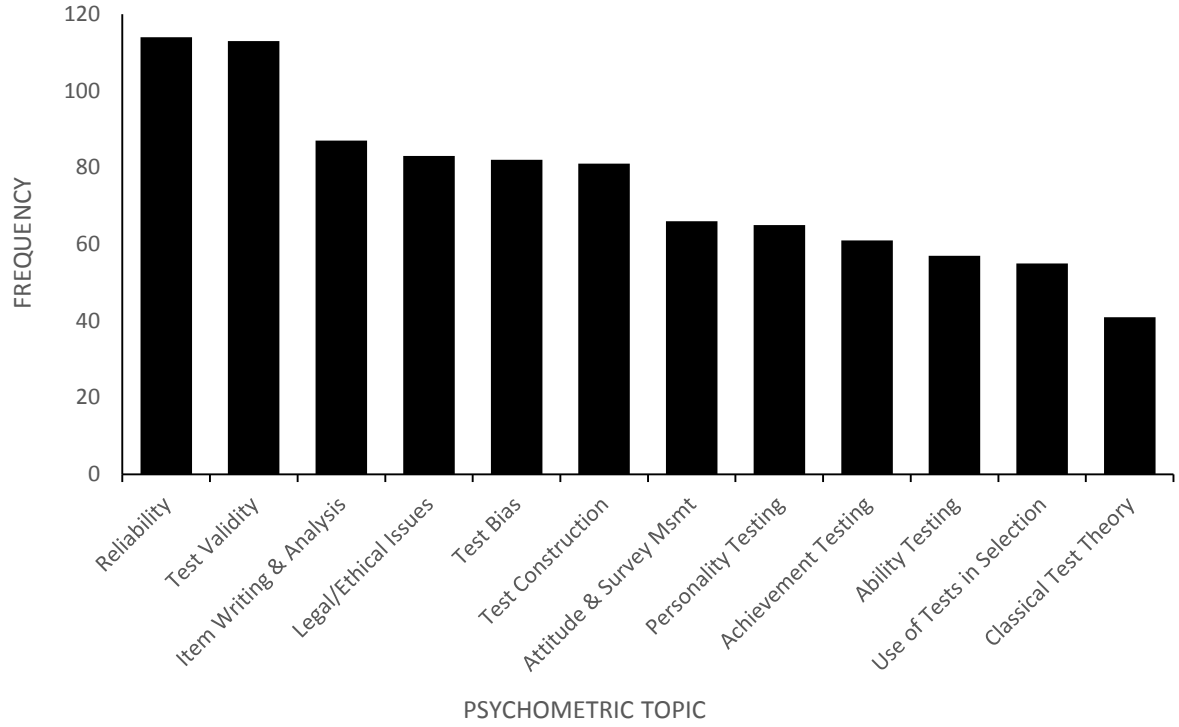

Figure 1. Pareto chart depicting the number of psychology departments (n = 153) displaying interest for offering an online learning module regarding certain psychometric topics.

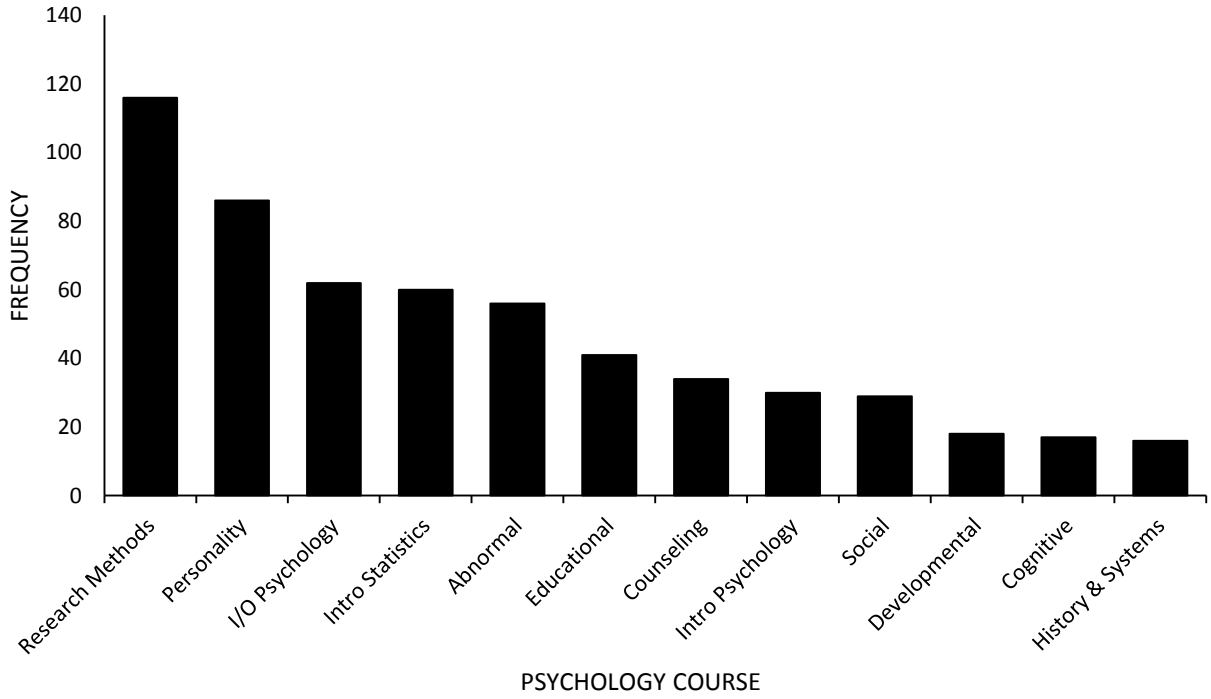

Figure 2. Pareto chart displaying the number of respondents citing online measurement modules as being beneficial within the following course subjects of psychology.

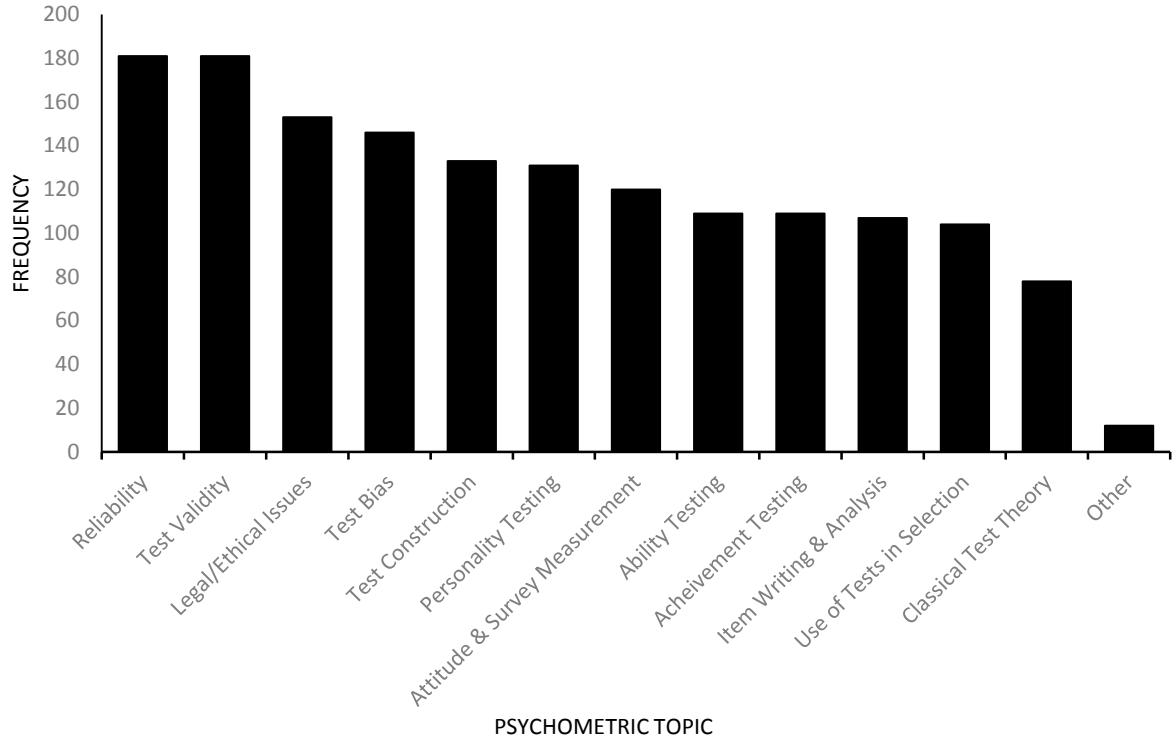

Figure 3. Pareto chart displaying the number of psychology departments (N = 181) who identified each of these psychometric topics as "important for an undergraduate psychology student to understand." The other response option (n = 12) consisted of respondent-supplied psychometric topics, specifically those pertaining to: statistics as they relate to testing (n = 2), clinical/counseling assessments (n = 2), neuropsychological testing (n = 2), basic testing, decision making, forensic applications, intelligence testing, introduction to item response theory, and multicultural issues. Open to interpretation, however, many of these written-in topics have varying degrees of overlap with some of the explicitly stated categories.

Appendix A

Survey on the Prevalence of Measurement in Undergraduate Psychology Curricula across the United States

Thank you for taking a brief break from your busy schedule to fill out our survey on psychological measurement. This assessment was created as part of a master's thesis project under the supervision of the Buros Center for Testing at the University of Nebraska-Lincoln. The questionnaire is designed to gather information on the prevalence of measurement-related topics in undergraduate psychology curricula across the United States. All identifying information will be kept confidential and results will only be presented in group format. By starting the survey it implies your informed consent to participate. We do ask that this form be filled out by either one of the following individuals: the most recent measurement instructor (preferably) or the chair of the psychology department. Your input is greatly appreciated and we ask that this survey be completed no later than May 10, 2013. Any questions/concerns that may arise should be directed to Kevin Dahlman (kdahlman@huskers.unl.edu). Thank you for your time!

1. Please provide the name of your academic institution:
   _____

2. Please provide the name of your department:
   _____

3. What is your affiliation with the department?
   - o Full-time Faculty Member – Department Chair
   - o Full-time Faculty Member – Not Department Chair
   - o Adjunct Faculty
   - o Departmental Staff Member

4. Have you ever taught a course in psychological measurement?
   - o Yes
   - o No

5. On average, over the last five years, how many students graduated with a major in psychology *per academic year*?
   - Less than 25
   - 25-49
   - 50-74
   - 75-99
   - 100 or more

6. Does your department require undergraduate students to take a course in psychological measurement?
   - Yes
   - No
   - Not required, but it is recommended

7. Does your department offer a class in psychological measurement?
   - Yes
   - No, but we would like to offer a course
   - No, we used to offer such a course but no longer do
   - No, we have not considered offering a course

8. How many undergraduate courses pertaining specifically to psychological measurement are available through your department?
   - 1
   - 2
   - 3 or more

9. How often are psychological measurement courses taught in your department?
   - Every semester
   - Every other semester
   - Every other year
   - Less than every other year

10. What is the rank of the person who teaches these psychological measurement courses most often?
    - Professor
    - Associate Professor
    - Assistant Professor
    - Instructor/Lecturer
    - Adjunct Faculty
    - Graduate Teaching Fellow or the equivalent

11. In what field is the graduate degree of this psychological measurement instructor?
    - Clinical Psychology
    - Counseling Psychology
    - School Psychology
    - Experimental Psychology
    - Developmental Psychology
    - Educational Psychology
    - Cognitive Psychology
    - Biopsychology
    - Psychology of Personality
    - Social Psychology
    - Quantitative Psychology
    - Industrial and Organizational Psychology
    - Psychometrics
    - Statistics
    - Education
    - Other, please provide:
      _____

12. Are there any other departments within your university that offer relevant psychological measurement courses for which your students are encouraged to enroll?
    - Yes
    - No
    - Not Sure

13. Through which other departments are the psychological measurement courses offered? (Select all that apply)
    - Educational Psychology
    - Education
    - Statistics
    - Human Development
    - Other, please provide department name(s):
      _____

14. What general topics pertaining to psychological measurement do you find important for undergraduate psychology students to understand?
    - o Reliability
    - o Test Validity
    - o Legal/Ethical Issues in Testing
    - o Test Construction
    - o Item Writing/Analysis
    - o Classical Test Theory
    - o Test Bias
    - o Use of Tests in Selection
    - o Ability Testing
    - o Achievement Testing
    - o Personality Testing
    - o Attitude and Survey Measurement
    - o Other, please provide:

      _____

15. If your institution does not have a course in psychological measurement, would your department consider offering an undergraduate psychological measurement class, through an institution other than your own, via distance learning?
    - o Yes
    - o No
    - o Not sure
    - o We offer such a course, so this question is not relevant to us

16. Are there courses in your department that cover psychological measurement topics, but are not psychological measurement courses?
    - o Yes, please list one or more such courses:

      _____
    - o No
    - o Not sure

17. Would you consider offering supplemental learning materials through online modules pertaining to psychological measurement topics to assist students?
    - o Yes
    - o No
    - o Not sure

18. What psychological measurement topic(s) would you find most helpful for offering an online learning module? (Select all that apply)
    - o Reliability
    - o Test Validity
    - o Legal/Ethical Issues in Testing
    - o Test Construction
    - o Item Writing/Analysis
    - o Classical Test Theory
    - o Test Bias
    - o Use of Tests in Selection
    - o Ability Testing
    - o Achievement Testing
    - o Personality Testing
    - o Attitude and Survey Measurement
    - o Other, please provide:

      _____

19. For which course(s) would you find online psychological measurement learning modules to be most beneficial? (Select all that apply)
    - o Introduction to Psychology
    - o Research Methods
    - o Introductory Statistics
    - o Personality
    - o Developmental Psychology
    - o Cognitive Psychology
    - o Social Psychology
    - o History and Systems of Psychology
    - o Abnormal Psychology
    - o Counseling Psychology
    - o Educational Psychology
    - o Industrial and Organizational Psychology
    - o Other, please provide:

      _____

20. How well prepared are your psychology major graduates in these measurement topics?

	Unprepared	Somewhat Unprepared	Somewhat Prepared	Prepared
Reliability	○	○	○	○
Test Validity	○	○	○	○
Legal/ Ethical Issues in Testing	○	○	○	○
Test Construction	○	○	○	○
Item Writing/Analysis	○	○	○	○
Classical Test Theory	○	○	○	○
Test Bias	○	○	○	○
Use of Tests in Selection	○	○	○	○
Ability Testing	○	○	○	○
Achievement Testing	○	○	○	○
Personality Testing	○	○	○	○
Attitude and Survey Measurement	○	○	○	○
Other, please provide: _____	○	○	○	○

21. Would you like to receive the results of this survey via email?
    - ○ Yes
    - ○ No

22. Please provide your name and email address for which you would like the results to be sent.
    Name: _____
    Email Address: _____

Appendix B

Example of Excel Data File Used in the Institution Selection Process

Name
University of Minnesota-Twin Cities
Walden University
Capella University
Saint Cloud State University
Minnesota State University-Mankato
University of Minnesota-Duluth
University of St Thomas
Winona State University
Metropolitan State University
Minnesota State University Moorhead
Southwest Minnesota State University
Saint Mary's University of Minnesota
Bemidji State University
St Catherine University
Hamline University
Bethel University
The College of Saint Scholastica
Augsburg College
St Olaf College
Northwestern College
Concordia University-Saint Paul
Concordia College at Moorhead
Gustavus Adolphus College
Argosy University-Twin Cities
College of Saint Benedict
Carleton College
Saint Johns University
Macalester College
University of Minnesota-Morris
North Central University
Crown College
Bethany Lutheran College
Crossroads College
Oak Hills Christian College

Appendix C

Initial Email Invitation without UNL Letterhead

Dear [*Insert Name*]:

Hello! My name is Kevin Dahlman and I am conducting a study on the prevalence of measurement in undergraduate psychology curricula across the United States. This study is being conducted as a master's thesis project under the supervision of Kurt F. Geisinger, Ph.D. and the Buros Center for Testing staff at the University of Nebraska – Lincoln. The results of this study will help clarify to the entire psychological community the extent to which measurement-related topics are covered in undergraduate psychology curricula across the country.

Completing the survey will require approximately ten minutes of your time and can be done in a setting of your choosing. There are no known risks associated with participating in this study and your responses will be coded with an ID key, which will be kept confidential upon completion of the data collection procedures so that no identifying information will be made available to the general public.

If you would like your institution to be reflected in the data obtained from this survey, please click on the link below. Your decision to participate would be greatly appreciated and is strictly voluntary. Please know that you are free to withdraw at any time, however by completing and submitting your responses to the survey, you have given your informed consent to participate. In doing so, you may wish to print a copy of this email for your records.

Any questions or concerns that may arise regarding this study can be directed to me or my adviser, Kurt Geisinger (kgeisinger@buros.org). If you would like to speak with someone else, please feel free to contact the Research Compliance Services Office at 402-472-6965 or irb@unl.edu.

Thank you for your time!

[*Insert Survey Link*]

Kevin Dahlman

kdahlman@huskers.unl.edu

(815)693-4363

IRB#20121012944 EX

Appendix D

Email Reminder without UNL Letterhead

Dear [*Insert Name*]:

Hello! My name is Kevin Dahlman and this is a reminder email about the invitation I sent you to be included in my study on the prevalence of measurement in undergraduate psychology curricula across the United States. If you have already completed the survey, then please disregard this message.

This study is being conducted as a master's thesis project under the supervision of Kurt F. Geisinger, Ph.D. and the Buros Center for Testing staff at the University of Nebraska – Lincoln. The results of this study will help clarify to the entire psychological community the extent to which measurement-related topics are covered in undergraduate psychology curricula across the country.

Completing the survey will require approximately ten minutes of your time and can be done in a setting of your choosing. There are no known risks associated with participating in this study and your responses will be coded with an ID key, which will be kept confidential upon completion of the data collection procedures so that no identifying information will be made available to the general public.

If you would like your institution to be reflected in the data obtained from this survey, please click on the link below. Your decision to participate would be greatly appreciated and is strictly voluntary. Please know that you are free to withdraw at any time, however by completing and submitting your responses to the survey, you have given your informed consent to participate. In doing so, you may wish to print a copy of this email for your records.

Any questions or concerns that may arise regarding this study can be directed to me or my adviser, Kurt Geisinger (kgeisinger@buros.org). If you would like to speak with someone else, please feel free to contact the Research Compliance Services Office at 402-472-6965 or irb@unl.edu.

Thank you for your time!

[*Insert Survey Link*]

Kevin Dahlman

kdahlman@huskers.unl.edu

(815)693-4363

IRB#20121012944 EX

Appendix E

Final Email Reminder without UNL Letterhead

Dear [*Insert Name*]:

Hello! My name is Kevin Dahlman and this is a final reminder about the email invitation I sent you to be included in my study on the prevalence of measurement in undergraduate psychology curricula across the United States. If you have already completed the survey, then please disregard this message.

This study is being conducted as a master's thesis project under the supervision of Kurt F. Geisinger, Ph.D. and the Buros Center for Testing staff at the University of Nebraska – Lincoln. The results of this study will help clarify to the entire psychological community the extent to which measurement-related topics are covered in undergraduate psychology curricula across the country.

Completing the survey will require approximately ten minutes of your time and can be done in a setting of your choosing. There are no known risks associated with participating in this study and your responses will be coded with an ID key, which will be kept confidential upon completion of the data collection procedures so that no identifying information will be made available to the general public.

If you would like your institution to be reflected in the data obtained from this survey, please click on the link below. Your decision to participate would be greatly appreciated and is strictly voluntary. Please know that you are free to withdraw at any time, however by completing and submitting your responses to the survey, you have given your informed consent to participate. In doing so, you may wish to print a copy of this email for your records.

Any questions or concerns that may arise regarding this study can be directed to me or my adviser, Kurt Geisinger (kgeisinger@buros.org). If you would like to speak with someone else, please feel free to contact the Research Compliance Services Office at 402-472-6965 or irb@unl.edu.

Thank you for your time!

[*Insert Survey Link*]

Kevin Dahlman

kdahlman@huskers.unl.edu

(815)693-4363

IRB#20121012944 EX