

Current State of Standardization Adherence: A Reflection of Competency in Psychological Assessment

Cortney Wolfe-Christensen and Jennifer L. Callahan
Oklahoma State University

Identification of competencies to be mastered during training and maintained through lifelong learning has recently received considerable attention. This study investigated a single, commonly identified core assessment skill competency, adherence to standardized administration procedures, to establish a current appreciation of baseline competency in the profession. Standardization adherence among 150 professionals averaged 70% and was not significantly related to level of education or the time elapsed since degree attainment. Contrary to expectations, adherence was significantly higher for less frequently administered tests than for those more commonly administered. Accessibility to procedural instructions (e.g., embedded within stimuli) was unrelated to levels of adherence.

Keywords: psychological assessment, standardization procedures, adherence, professional competencies, lifelong learning, competency

Consistent with other professions, psychology is committed to developing and maintaining professional competencies necessary for independent practice. Although varying definitions of professional competence exist, most include reference to competency being essential for the protection of the public (e.g., National Council of State Boards of Nursing, 1996), which requires the accurate mastery of specific skills (e.g., Durley, 2005; Epstein & Hundert, 2002; Roberts, Borden, Christiansen, & Lopez, 2005; Rodolfa et al., 2005).

The Examination for Professional Practice in Psychology (EPPP) is thought to represent one method of protecting the public by ensuring that licensed psychologists evidence a minimal level of factual knowledge. However, a sizable gap may exist between performances on examinations and evaluations of clinical competence (Dawes, 1996), a problem which is not unique to the profession of psychology (e.g., in pediatrics clinical care; Joorabchi & Devries, 1996). Furthermore, passing the EPPP is, most likely, a one-time early career requirement that is not necessarily sufficient for ensuring that competency is maintained via lifelong learning throughout one's career. In fact, some have speculated that some competencies may actually diminish over a career of professional practice (Slate, Jones, & Murray, 1991),

perhaps partly owing to observer drift and a lack of feedback on performance (Kratochwill & Sheridan, 1990).

Little empirical attention has been devoted to this topic, but it is needed given the recent efforts at defining specific competencies that should be expected for independent practice and maintained through lifelong learning. In particular, the 2002 "Competencies Conference: Future Directions in Education and Credentialing in Professional Psychology" advanced the defining of specific competencies by convening workgroups of highly respected psychologists, with several of the workgroups subsequently publishing consensus statements (the workgroup consensus statements can be obtained online at http://www.appic.org/news/3_1_news_Competencies.htm). The workgroups were highly productive in identifying core competencies, facilitating the possibility of exploring how such competencies are developed and maintained.

A promising model conceptualizing this process is the "cube" model (Rodolfa et al., 2005), which articulates a three-dimensional structure that places foundation competency domains (i.e., knowledge, skills, attitudes, and values) in relationship to the varying functional activities of psychologists in different stages of professional development. In this regard, professional competency is not viewed as something that is achieved during training and before licensure, but rather as a career-long endeavor that necessitates ongoing assessment of competencies. This viewpoint is shared by the Assessment of Competence Workgroup that met during the 2002 Competencies Conference, which concluded that assessment of competencies should occur throughout the career life span (Roberts et al., 2005).

A separate workgroup focusing specifically on competency in psychological assessment therefore formulated a set of core competencies that are meant to span one's career (Krishnamurthy et al., 2004). In fact, Krishnamurthy et al. (2004) described assessment competence as "a defining aspect of psychological expertise" (p. 726) that is "essential to all health-service practices in psychology, transcending specialties" (p. 732). This group specifically identified eight core markers of competency in psychological assessment, the most elab-

CORTNEY WOLFE-CHRISTENSEN, MS, is a 4th-year graduate student in the clinical psychology program at Oklahoma State University. Her research interests include combining neuropsychology and pediatric psychology to examine neuropsychological, emotional, behavioral, and social adjustment outcomes in children diagnosed with pediatric cancer.

JENNIFER L. CALLAHAN, PhD, ABPP, is now an assistant professor in the Department of Psychology, University of North Texas, Denton, Texas, and also holds a visiting appointment at the Department of Psychology, Yale School of Medicine. Her research interests span the therapy and assessment domains, with particular emphasis on training and the development of professional competencies.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Jennifer L. Callahan, Department of Psychology, University of North Texas, P.O. Box 311280, Denton, TX 76203-1280. E-mail: jennifercallahan@unt.edu

orated of which is an area termed “technical assessment skills” (p. 732). Within this single area of competency, the group further identified seven technical assessment skills, including “systemic data gathering” (p. 732).

Although systematic data gathering is only a single skill within psychological assessment competency, this skill is easily operationalized for research investigations and was therefore selected for the current study examining professional competence across the career life span. Although this study is meant as an exploration of a discrete skill, it is not meant to convey that assessment competence consists solely of this skill. Competency in psychological assessment requires a combination of knowledge and technical skills that interface with refined, nuanced critical thinking ability.

Although easily operationalized for empirical study, systematic data collection can be operationalized in a number of different, but equally important, manners. Previous research in this area has primarily operationalized systematic data collection by measuring the prevalence and types of scoring errors, particularly those associated with the commonly administered Wechsler intelligence scales (see, e.g., Franklin, Stillman, Burpeau, & Sabers, 1982; Ryan & Schnakenberg-Ott, 2003; Slate & Jones, 1990). The results from these studies have consistently indicated that scoring errors are more common than not and that greater numbers of test administrations do not improve scoring accuracy (Franklin et al., 1982; Slate & Jones, 1990). For example, Ryan and Schnakenberg-Ott (2003) found that both graduate students and doctoral-level psychologists were significantly inaccurate in their scoring of Wechsler Adult Intelligence Scale, third edition (WAIS-III), protocols, even though they consistently reported high confidence in their test administration abilities. Similarly, Slate and colleagues (1991) documented that graduate students’ scoring accuracy was unrelated to the number of practice administrations of the test. In fact, the researchers concluded that 88% of the Full Scale IQs reported by the examiners were affected by inaccurate scoring. Findings from these studies have suggested that the scoring accuracy of the Wechsler scales is indeed problematic across a range of students and practitioners and that increased experience with the scales was not sufficient to facilitate competency.

For this investigation, systematic data collection was operationalized in a novel manner to explore the relationship between experience and competency and provide a snapshot view of current baseline competency in the profession. In particular, this study examines the specific skill of adherence to standardized procedures that are associated with specific assessment instruments. This study therefore builds on earlier investigations of scoring accuracy by examining standardization adherence as another component of systematic data collection and, secondarily, exploring whether lifelong learning might be evidenced by those who engage in frequently assessing clients, repeatedly using assessment instruments, and/or using instruments with prominently placed standardized procedures. To our knowledge, this is the first study to look at global data collection procedures across a variety of commonly administered psychological tests.

Method

Participants

The participants in this study included 150 students and professionals (49 men and 101 women) engaged in administering com-

mon cognitive tests who responded to personal requests to complete an online survey. Participants spanned a range of educational levels, with 20 currently holding a bachelor’s degree (13 BA, 7 BS), 59 with a master’s degree (38 MA, 21 MS), and 71 with a doctoral degree (15 PsyD, 56 PhD). The amount of time elapsed since the participant’s degree was earned ranged from 0 to 38 years ($M = 5.90$, $SD = 7.17$). The majority (84%) of the participants reported that they have, or will have, graduated from an American Psychological Association (APA)-accredited program, and slightly more than half (53%) indicated that they had/have had some specialty training in neuropsychological assessment during graduate school. Participants were drawn from 36 U.S. states and three Canadian provinces and worked in a range of settings including independent practice (17.3%), hospital (57.3%), academic setting (16.0%), or other (9.3%).

Materials

We composed a novel survey to assess adherence to standardized procedures on several commonly administered cognitive and neuropsychological tests. The survey included questions to elicit demographic and training background information, but focused on questions regarding specific procedures associated with both common (e.g., Block Design subtest of the WAIS-III) and less common (e.g., Boston Naming Test) assessment instruments. To combat possible response-set formation, questions about routine clinical practice (e.g., whether brief feedback is given at the conclusion of testing) were integrated into the questions examining standardization adherence. In addition, a few questions were composed that asked about the usage of nonexistent, but plausible, standard procedures (e.g., asking whether the participant instructs clients “not to cross any lines” on the Trail Making Test). For examples of survey items, see Table 1.

A search of the literature did not yield any recent studies detailing the frequency of administration associated with a broad range of assessment instruments. As a result, the survey was also designed to assess how often specific tests were administered. In part, this was done to provide additional information on the current status of the field. More specific to this study, it provides a means of determining whether the novel survey adequately included standardization adherence questions for the most frequently administered tests.

Procedures

An online survey was created and an E-mail with a link to the survey was posted on select psychologist-verified member electronic mailing lists to describe the study and solicit participation. To recruit student and faculty participants, an e-mail was sent to professors in various psychology graduate programs across the United States and Canada; after generating a list of programs included in *Graduate Study in Psychology* (APA, 2006), all programs that reported having an emphasis in neuropsychology were contacted. The survey was not registered with any Internet search engines, no key words were programmed into the survey page to assist search engines, and random checks to determine whether major search engines were locating the survey yielded no hits during the course of the study. This study was approved by the Institutional Review Board at Oklahoma State University.

Table 1
Examples of Questions from the Adherence Survey

Question	Answer choices ^a
When administering Grooved Pegboard, do you have your patient move from “right to left” when using their left hand (i.e. opposite direction of the right hand)?	Yes/no
When administering Block Design from the WAIS-III, do you use a standard presentation for the blocks (i.e. specific numbers of red blocks, white blocks, and combination blocks)?	Yes/no
Do you ever convert any non-standard scores to Standard or T-scores for comparisons with other tests?	Yes/no
When administering Symbol Search from the WAIS-III, how many pages of stimuli do you present at once?	1 page (booklet folded along spine)/2 pages
During the administration of the Wisconsin Card Sorting Test, how often do you give feedback other than the “right” (correct) or “wrong” (incorrect) responses?	Never/occasionally/often/almost always
Do you give a practice trial prior to administering the Purdue Pegboard?	Yes/no

Note. WAIS-III = Wechsler Adult Intelligence Scale, third edition.

^a “Do not administer” was also a response option for all items pertaining to standardization.

Results

Preliminary Analyses

Before running analyses, the sample was examined with regard to the number of patients seen each year, and three outliers were removed. Specifically, 2 participants reported seeing 0 patients and 1 participant reported seeing 850 patients annually. Thus, the final sample consisted of 147 participants. Given that assessment batteries are not identical across practitioners, many of the participants reported that they did not administer a particular test throughout the survey. Therefore, to account for these data, each participant was assigned a mean assessment score based on their individual responses. These scores were calculated on the basis of correct and incorrect answers to standardization procedures as detailed in instructions associated with each test. Correct answers were given a score of 1, incorrect answers were given a score of 0, and “do not administer” was coded as missing, thereby eliminating those questions from the calculation of mean assessment scores. Consequently, possible values ranged from 0 to 1, but the obtained individual mean assessment scores actually ranged from .38 to .92 (aggregate $M = .69$, $SD = .11$).

Primary Analyses

To determine whether differences in adherence exist among educational levels, a one-way analysis of variance (bachelor’s degree vs. master’s degree vs. doctoral degree), with the assessment score as the dependent variable, was conducted. The results indicated no significant differences between the groups (BA and BS, $M = .68$, $SD = .13$; MA and MS, $M = .69$, $SD = .11$; and PhD and PsyD, $M = .70$, $SD = .10$). Similarly, a bivariate correlation revealed no significant relationship between the length of time since the participants’ highest degree was earned and their mean adherence score ($r = -.02$, *ns*).

Next, to determine whether differences existed between professionals who see few patients and those who see many patients, the sample was separated into quartiles based on the number of adult patients seen each year (range = 3–450, $M = 101.72$, $SD = 87.50$). An independent t test to investigate whether differences existed between the highest and lowest quartile in terms of

their assessment score was nonsignificant, $t(65) = -.95$, *ns*. Stated another way, accuracy did not significantly differ between those professionals who administer many batteries annually ($M = 247.52$, $SD = 64.95$) and those who administer relatively fewer batteries annually ($M = 19.78$, $SD = 8.94$).

To examine whether adherence was higher on tests that are commonly used and lower on tests that are not commonly used, chi-square analyses were conducted. As evident in Table 2, the most commonly administered tests were the Trail Making Test, WAIS-III, and Wisconsin Card Sorting Task, and the least commonly administered tests were the Booklet Category Test, Hooper Visual Integration Test, and Purdue Pegboard. Mean accuracy scores were calculated for these questions, and scores above .70 were coded as “competent” and scores below .70 were coded as “not competent.” Although this cutting score is somewhat arbitrary, it is consistent with the history in educational settings of requiring a score of 70% (minimum score for achieving a grade of C) to indicate performance that is satisfactory and not below expectations. A chi-square test of association was conducted, and the results indicated that, contrary to expectations, the participants reported significantly higher adherence to standardization procedures on the most uncommonly administered tests, $\chi^2(1, N = 268) = 9.56$, $p < .01$.

Next, we examined whether those tests with standardized procedures embedded within materials that are necessary at the time of testing (e.g., score sheets, query guides, etc.) would have high adherence and those with less preferentially placed standardized procedures would have low rates of adherence. The tests determined to have easily accessible procedural instructions were the WAIS-III, Booklet Category Test, and Wechsler Memory Scale, third edition (WMS-III), and those deemed to have less accessible instructions were the Grooved Pegboard, Purdue Pegboard, Hooper Visual Integration Test, and Boston Naming Test. As previously described, mean scores were calculated for each category (i.e., accessible and less accessible), and scores above .70 were coded as “competent,” and scores below .70 were coded as “not competent.” A chi-square test of association indicated that, interestingly, adherence was not related to the accessibility of the standardized procedures.

Table 2
Percentage of Adult Patients Receiving Specific Neuropsychological Tests

Test	0%–10%	11%–20%	21%–30%	31%–40%	41%–50%	51%–60%	61%–70%	71%–80%	81%–90%	91%–100%
Booklet Category Test	110	12	7	3	4	0	2	2	2	8
Boston Naming Test	31	8	6	9	7	7	5	10	28	39
Clock Drawing	57	13	8	4	10	5	7	16	15	15
CVLT-II	43	4	4	6	9	9	9	13	19	34
Finger Tapping	39	14	9	7	15	6	10	13	11	26
Hooper Visual Integration Test	91	15	9	3	10	3	4	5	4	6
Lafayette Grooved Pegboard	48	12	5	7	12	2	9	15	16	24
Line Bisection	106	20	7	3	5	0	1	4	3	1
Purdue Pegboard	118	11	2	5	0	1	0	2	3	8
RAVLT	94	19	9	4	4	1	3	4	4	8
Rey-Ostreith Complex Figure	34	11	12	11	9	5	11	8	19	30
TOMM	42	30	15	9	10	8	7	4	10	15
Trail-Making Tests	5	3	2	2	1	2	4	6	27	98
Verbal Fluency Tasks (e.g., COWAT, animal naming)	9	2	2	1	3	5	5	3	28	92
WAIS-III	12	4	8	3	7	6	14	16	16	64
Wisconsin Card Sorting Task	12	13	5	8	10	6	13	11	19	53
WRAT-III	35	18	10	10	9	7	7	10	18	26

Note. Total N responding = 150. CVLT-II = California Verbal Learning Test-II; RAVLT = Rey Auditory-Verbal Learning Test; TOMM = Test of Memory Malingering; COWAT = Controlled Oral Word Association Test; WAIS-III = Wechsler Adult Intelligence Scale, third edition; WRAT-III = Wide Range Achievement Test-III.

Secondary Analyses

Item-level examinations suggest that several standardization procedures may be at a relatively increased risk for nonadherence. In most cases, this potentially results in a lower score for the patient. For example, although 81.3% of the sample reported administering Finger Tapping, only 51.6% of those respondents reported that they begin timing “on the first tap,” as described in standardized instructions, with the remainder indicating that they begin timing on telling the patient to begin. Similarly, although 89.3% of respondents reported administering the Symbol Search subtest of the WAIS-III, only 45.5% of them are following standardization procedures by presenting two pages of stimuli at once.

Adherence is also relatively poor for the Wisconsin Card Sorting Test (administered by 86.7% of the sample); however, on this test it is difficult to determine the potential impact of nonstandard procedures on patient performances and scores. More specifically, 33.6% of the sample are offering nonstandard feedback (i.e., giving feedback other than “right” and “wrong”), which is likely facilitative but could be problematic for some patients. Probably less salient to patient performances and scores, we also found that 50.8% of the respondents are incorrectly presenting the cards one at a time, instead of giving the deck to the patient, and 36.4% are incorrectly presenting the cards with the stimulus hidden (i.e., number side up).

Exploratory Analyses

Although not a focus of this investigation, we felt it useful information for the larger field to document a few current practice behaviors. Fully 77% of participants reported using a flexible battery approach (rather than a fixed battery) to answer referral questions. Eighty-five percent reported that they routinely convert nonstandard scores to standard scores to facilitate comparisons between tests. In contrast, only 21% of the respondents indicated that they give brief, preliminary feedback to clients at the conclusion of testing.

In addition, given that it has been reported that it takes an average of nearly 17 years for clinical trial research to become incorporated into accepted practice standards (Balas & Boren, 2000), questions were posed about practice habits reflecting incorporation of the comparatively recent research on the WAIS-III and WMS-III. Such research has included a number of psychometric examinations of subtests and factor structures. The data indicate that although 92.6% of respondents reported administering the WAIS-III, only 34.6% of the sample administers the Object Assembly subtest and fewer than 3% administer it to every patient. In addition, although 80% of the participants administer the WMS-III, 52% of the sample omits the Faces subtest and 56% of the sample reported administering the Visual Reproduction subtest in place of Faces.

Implications

The development of professional competencies in psychological assessment has received increased attention in the past few years, although a consensus on the most appropriate means to evaluate mastery of these skills has yet to be determined. The present study sought to investigate the self-reported adherence of psychologists

to standardized assessment procedures. Similar to findings from other studies that assessed scoring errors on the Wechsler intelligence tests (Franklin et al., 1982; Ryan & Schnakenberg-Ott, 2003; Slate & Jones, 1990), overall adherence to standardization procedures was unrelated to either educational degree or the time elapsed since the degree was earned.

Aggregate results revealed that psychologists obtain a mean accuracy of 70% in adherence to standardized assessment procedures on measures they administer. However, whether the figure of 70% is an acceptable level of adherence is difficult to determine. Within educational settings, 70% accuracy is typically considered the minimal score necessary for satisfactory performance (i.e., a grade of C). More relevant to psychology, this nationally accepted convention of 70% was adopted as the necessary criterion for a passing score on the EPPP, the nationwide psychology licensure exam. However, passing the EPPP national psychology licensure exam is intended only to establish that a respondent has obtained a minimal level of competency necessary before independent practice (Association of State & Provincial Psychology Boards, 2006). Examination of the data (see Table 3) indicates that unless the cutting threshold is lowered to 60%, more participants are classified as incompetent than competent. Yet, this threshold is likely below what would be considered acceptable to clients, funding sources, and the public at large. Thus, an important question for the field's consideration is how typical competency can be enhanced to an acceptable level. One possibility is that measures such as this one could be used at different levels of training and experience to determine whether fundamental benchmarks of competencies are present. For example, such an examination might be incorporated into comprehensive examinations at the doctoral level, used by internship sites screening applicants, state licensing boards, or even board certification panels.

On the basis of the published report that it takes nearly 17 years before clinical trial research becomes incorporated into conventional practice (Balas & Boren, 2000), it is suggested that assessment professionals who are not engaged in further developing their skills would also not incorporate relatively recent research findings into their work. Yet, it appears that incorporation of recent research is occurring. For example, a significant proportion of the sample reported administering the Visual Reproduction subtest in place of the Faces subtest on the WMS-III, which is consistent with current literature on the increased sensitivity of Visual Reproduction to detect deficits in visual memory (Palmer, Taylor, & Heaton, 2003). Similarly, the findings indicate that Object Assembly is frequently omitted from WAIS-III administrations, which may reflect incorporation of research noting relatively low reliability for this subtest (Tulsky & Ledbetter, 2000). Thus, although the findings on competency in the practice skill of adhering to

standardized procedures may be discouraging, it is notable that participants do appear to be continuing to consume and integrate relevant research findings that affect clinical care.

In contrast to expectations, it was found that adherence to standardization procedures was significantly better for tests that are more uncommonly used in practice than those most commonly administered. Perhaps respondents consulted manuals for more unfamiliar measures while filling out the survey. Alternatively, perhaps respondents have a subjective threshold of estimated competency and, when faced with a measure they feel less competent with, they are more attentive in adhering to standardized procedures.

The participants recruited for this study were individuals who identified themselves as having an assessment specialty, as it was felt that such participants would assume greater identification of responsibility and be more frequently exposed to assessment materials. In considering the potential risks associated with expanding recruitment and the risk of including participants who rarely engage in such activities, it was concluded that error rates might be overestimated and adversely affect the perception of the field. However, this recruitment focus is not without limitations.

First, this sample may not be representative of the larger population of psychologists, as the male-to-female ratio and practice setting distributions differ somewhat from those reported by Sweet, Peck, Abramowitz, and Etzweiler (2002). More specifically, in the current sample there were twice as many women as men, as opposed to the expected 1.6:1 male-to-female ratio (Sweet et al., 2002), and the practice setting was more likely to be a hospital than independent practice, which is also in contrast to the findings reported in Sweet et al. (2002). In addition, the sample varied greatly with regard to the number of patients assessed each year. However, the results indicated that adherence to standardization procedures was not significantly related to the annual number of patients seen. Although some respondents reported seeing patients infrequently, this finding may reflect that such individuals could still be heavily involved with the assessment measures via research, supervision, or teaching activities. For example, some respondents may teach graduate coursework in assessment, supervise the work of trainees and/or psychometricians, or use the identified measures in research. Such individuals may therefore report a low number of annual patients seen while still considering themselves to have expertise in assessment procedures.

An additional limitation of the current study is that no measure of social desirability was included, and it is likely that persons who consider themselves to have expertise in an area would be invested in appearing in a professionally desirable manner. However, a search of the literature did not locate any norm-referenced measures of professional desirability applicable to the population of interest in this study. For the purposes of managing an impression of professional desirability, it is believed that some participants may have consulted manuals or other sources of information about standard procedures to assist them in completing the survey. Readers are therefore cautioned that the findings here may somewhat underestimate the rates of deviation from standard procedures. Future studies in this area may be enhanced by direct observation of participant responding or actual practice habits.

Furthermore, this study focused on participants working with adult patients. Although adherence to standardization is also certainly im-

Table 3
Cumulative Totals of Participants Categorized as Competent Versus Incompetent at Varying Cut Percentage Scores

Accuracy level	Competent (<i>N</i>)	Incompetent (<i>N</i>)
60% threshold	115	32
70% threshold	67	80
80% threshold	15	132
90% threshold	1	146

portant when working with more youthful populations, standardized procedures often vary within a given test as a function of the patient's age. Additionally, thoughtful flexibility in procedures is perhaps more frequently required when working with very young populations. Therefore, the findings from this study may underestimate error rates of psychologists as a whole. This assumption, however, may be a limitation to this study. It is likely that nonstandard procedures are sometimes necessary while working with adult clients for therapeutic reasons or clinical hypothesis testing. Although including survey questions asking participants why they deviate from a standard was considered, it was determined that doing so would require a mechanism for first drawing attention to the appropriate standard. Future research that allows for direct observation would benefit from such follow-up interviewing.

Although the current study had several limitations, the results suggest that, as a group, psychologists are truly addressing the individual needs of their patients, as evidenced by the flexibility in test batteries. Although some tests are administered consistently across practitioners (e.g., Trail Making Test, WAIS-III, and Wisconsin Card Sorting Test; see Table 2), it appears that psychologists are tailoring their batteries to adequately address the referral question for each patient (Ardila, 1992).

In addition, the majority of psychologists are converting raw scores to standard scores to allow for comparison across measures. These procedures allow the results to be interpreted in conjunction with other tests as well as supplementary information gathered through clinical interview and behavioral observation (APA, 2006). The alteration to batteries and conversion of raw scores to standard scores suggest that the participants are staying up to date with research findings in the field via continuing education, be it formal (e.g., workshops) or informal (e.g., reading journals, colleague consultation).

The results from the current study should serve as a baseline measure of adherence to standardization procedures in psychological assessment. Because assessment has been identified as one of the eight core competencies of professional psychologists, it is recommended that repeat measurement be conducted periodically in the future to assess the impact of the establishment of the competencies document and the push for a focus on the development of competencies in training (Association of Directors of Psychology Training Programs, 2002).

References

- American Psychological Association. (2006). *Graduate study in psychology 2006*. Washington, DC: Author.
- Ardila, A. (1992). Luria's approach to neuropsychological assessment. *International Journal of Neuroscience*, *66*, 35–43.
- Association of Directors of Psychology Training Programs. (2002). *ADPTC practicum competencies workgroup report on practicum competencies*. Retrieved July 15, 2006, from [http://www.adptc.org/public_files/Practicum%20Competencies%20FINAL%20\(Oct%20'06%20Version\).pdf](http://www.adptc.org/public_files/Practicum%20Competencies%20FINAL%20(Oct%20'06%20Version).pdf)
- Association of State & Provincial Psychology Boards. (2006). *EPPP exam information*. Retrieved July 20, 2006, from <http://www.asppb.org/epppExam/faq/score.aspx>
- Balas, E. A., & Boren, S. A. (2000). Managing clinical knowledge for health care improvement. In J. Bommel & A. T. McCray (Eds.), *Yearbook of medical informatics* (pp. 65–70). Stuttgart, Germany: Schattauer.
- Dawes, R. (1996). Licensing: The myth of protecting the public. In R. Dawes, *House of cards* (pp. 133–177). New York: Free Press.
- Durley, C. C. (2005). *The NOCA guide to understanding credentialing concepts*. Washington, DC: National Organization for Competency Assurance. Retrieved July 27, 2006, from <http://www.noca.org/members/CredentialingConcepts.pdf>
- Epstein, R. M., & Hundert, M. (2002). Defining and assessing professional competence. *JAMA* *287*, 226–235.
- Franklin, M. R., Stillman, P. J., Burpeau, M. Y., & Sabers, D. L. (1982). Examiner error in intelligence testing: Are you a source? *Psychology in the Schools*, *19*, 563–569.
- Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, DC: Author. Retrieved July 18, 2006, from <http://www.apa.org/science/FinalCode.pdf>
- Joorabchi, B., & Devries, J. M. (1996). Evaluation of clinical competence: The gap between expectation and performance. *Pediatrics*, *97*, 179–185.
- Kratochwill, T., & Sheridan, S. (1990). Advances in behavioral assessment. In T. Gutkin & C. Reynolds (Eds.), *The handbook of school psychology* (2nd ed., pp. 328–364). New York: Wiley.
- Krishnamurthy, R., VandeCreek, L., Kaslow, N., Tzaeau, Y. N., Miville, M. L., Kerns, R., et al. (2004). Achieving competency in psychological assessment: Directions for education and training. *Journal of Clinical Psychology*, *60*, 725–739.
- National Council of State Boards of Nursing. (1996). *Assuring competence: A regulatory perspective*. Chicago: Author.
- Palmer, B. W., Taylor, M. J., & Heaton, R. K. (2003). Accuracy of WAIS-III—WMS-III joint factor scores when one or more subtests is omitted or an alternate subtest is employed. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 392–450). San Diego, CA: Elsevier Science.
- Roberts, M. C., Borden, K. A., Christiansen, M. D., & Lopez, S. J. (2005). Fostering a culture shift: Assessment of competence in the education and careers of professional psychologists. *Professional Psychology: Research and Practice*, *36*, 355–361.
- Rodolfa, E., Bent, R., Eisman, E., Nelson, P., Rehm, L., & Ritchie, P. (2005). A cube model for competency development: Implications for psychology educators and regulators. *Professional Psychology: Research and Practice*, *36*, 347–354.
- Ryan, J. J., & Schnakenberg-Ott, S. D. (2003). Scoring reliability on the Wechsler Adult Intelligence Scale—third ed. (WAIS-III). *Assessment*, *10*, 151–159.
- Slate, J. R., & Jones, C. H. (1990). Identifying students' errors in administering the WAIS-R. *Psychology in the Schools*, *27*, 83–87.
- Slate, J. R., Jones, C. H., & Murray, R. A. (1991). Teaching administration and scoring of the Wechsler Adult Intelligence Scale—Revised: An empirical evaluation of practice administrations. *Professional Psychology: Research and Practice*, *22*, 375–379.
- Sweet, J. J., Peck, E. A., Abramowitz, C., & Etzweiler, S. (2002). National Academy of Neuropsychology/Division 40 of the American Psychological Association practice survey of clinical neuropsychology in the United States, Part I: Practitioner and practice characteristics, professional activities, and time requirements. *Clinical Neuropsychologist*, *16*, 109–127.
- Tulsky, D. S., & Ledbetter, M. (2000). Updating to the WAIS-III and WMS-III: Considerations for research and clinical practice. *Psychological Assessment*, *12*, 253–262.

Received August 17, 2006

Revision received April 11, 2007

Accepted June 4, 2007 ■