# APA GUIDELINES for Psychological Assessment and Evaluation

**APA TASK FORCE ON PSYCHOLOGICAL ASSESSMENT AND EVALUATION GUIDELINES**

APPROVED BY APA COUNCIL OF REPRESENTATIVES
**MARCH 2020**

AMERICAN PSYCHOLOGICAL ASSOCIATION

This document will expire as APA policy in 10 years (2030). Correspondence regarding the *APA Guidelines for Psychological Assessment and Evaluation* should be addressed to the American Psychological Association, 750 First Street, NE, Washington, 20002-4242.

**Suggested Citation**

American Psychological Association, APA Task Force on Psychological Assessment and Evaluation Guidelines. (2020). *APA Guidelines for Psychological Assessment and Evaluation*. Retrieved from https://www.apa.org/about/policy/guidelines-psychological-assessment-evaluation.pdf

AMERICAN
PSYCHOLOGICAL
ASSOCIATION

# APA GUIDELINES for Psychological Assessment and Evaluation

**APA TASK FORCE ON PSYCHOLOGICAL ASSESSMENT AND EVALUATION GUIDELINES**

APPROVED BY APA COUNCIL OF REPRESENTATIVES
**MARCH 2020**


**Task Force Workgroup**

***Linda F. Campbell, PhD (co-chair)***
University of Georgia

***Lisa D. Stanford, PhD, ABPP (co-chair)***
Akron Children's Hospital

***Vincent C. Alfonso, PhD***
Gonzaga University, School of Education

***Michael J. Cuttler, PhD, ABPP***
Law Enforcement Services, Inc.

***Stephen T. DeMers, EdD***
University of Kentucky

***Giselle A. Hass, PsyD, ABAP***
Independent Practice

**APA Staff**

***Catherine L. Grus, PhD***
Chief Education Officer
Education Directorate

***C. Vaile Wright, PhD***
Senior Director, Health Care Innovation
Practice Directorate

# AUTHOR'S NOTE

This document was developed by representatives from the Board of Professional Affairs, the Committee on Psychological Tests and Assessment, the Committee for the Advancement of Professional Psychology, the Committee on Professional Practice and Standards, the Board of Educational Affairs, and the Association of State and Provincial Psychology Boards in collaboration with APA staff. The work group included Vincent C. Alfonso, PhD; Linda F. Campbell, PhD (co-chair); Michael J. Cuttler, PhD, ABPP; Stephen T. DeMers, EdD; Giselle A. Hass, PsyD, ABAP; Lisa D. Stanford, PhD, ABPP (co-chair); and APA staff members Catherine L. Grus, PhD, and C. Vaile Wright, PhD.

The work group also acknowledges the earlier contributions from April Harris-Britt, PhD, representing the Board for the Advancement of Psychology in the Public Interest, and several members of CPTA who made substantive contributions to Guideline 5. The work group extends its appreciation to the APA staff members who facilitated the work of the guideline, including Marianne Ernesto, Mary G. Hardiman, MS, and Sarah A. Rose..

# TABLE OF CONTENTS


| | |
|---|---|
| Introduction | 3 |
| Scope | 4 |
| Audience and Stakeholders | 4 |
| Statement Distinguishing Between Guidelines and Standards | 4 |
| Compatibility With APA Ethics Code | 4 |
| Definition of Terms | 5 |
| Documentation of Need/Public Benefit | 5 |
| Development of the Guidelines | 6 |
| Selection of Evidence | 7 |
| Competence | 9 |
| Psychometric and Measurement Knowledge | 14 |
| Selection, Administration, and Scoring of Tests | 16 |
| Diverse, Underrepresented, and Vulnerable Populations | 18 |
| Training and Supervisory Qualifications and Role | 21 |
| Technology | 22 |
| References | 25 |

# INTRODUCTION

The purpose of the *American Psychological Association (APA) Guidelines for Psychological Assessment and Evaluation (PAE)* is to assist and inform psychologists of best practice when psychological instruments, including psychometric tests and collateral information, are used within the practice of psychological assessment and/or evaluation. As the discipline of psychology has expanded, the application of psychological assessment has also developed in response to new areas of practice. Integrated medical and primary care, online assessment and scoring, and global initiatives are examples of these new areas. Since the last publication of test user qualifications guidelines (APA, 2001), neuropsychology, forensic psychology, cognitive science, consulting, industrial/organizational, integrated health, and other fields have evolved into more defined and recognized specific areas of practice with developing professional practice guidelines, standards of practice, and identified consistency with the APA Ethics Code (APA, 2017a). Commensurate with the growth and varied expertise in the assessment field is the need for (a) core knowledge and skills that are essential for all psychologists whose assessment services impact individuals, groups, organizations, and the public and (b) specific domains of expertise within the field requiring additional knowledge, skills, and clinical experience to address the expansion of assessment service delivery across new populations and settings. These guidelines focus on the core knowledge and skills that psychologists, including those in specialty areas of practice, should strive to possess to deliver assessment services competently. Additionally, the purpose of these guidelines is to inform clients/patients, the public, other professionals collaborating with psychologists, regulatory bodies, and training programs of the qualification expected to acquire professional competency as defined by the APA Ethics Code and to maintain high professional standards of practice. The guidelines apply to standardized tests of ability, aptitude, achievement, attitudes, interests, personality, cognitive functioning, mental health, and other construct domains.

APA's (2001) "Guidelines for Test User Qualifications" were responsive to the identified problem of misuse of tests within the United States and internationally. The competence of the test users was the primary reason for concern, and the International Test Commission (2001) and several other countries have issued guidelines to address this concern. Competence of test users continues to be a growing problem despite the development of multiple guidelines, procedures, and standards. The current guidelines focus on two areas identified as deficits in the 2001 guidelines: the need for greater specificity of assessment constructs and foundational knowledge, and the need to address the expanded and more specialized roles of psychologists using assessment and evaluation tools in various clinical and other professional settings.

## Scope

The scope of the *APA PAE Guidelines* addresses the use of psychological instruments, including psychometric tests and collateral data when used within the context of psychological assessment, and focuses on (a) assessment procedures, (b) professional competencies in psychological assessment and evaluation, and (c) assessment knowledge competencies. There are several topics that are outside the scope of these guidelines. These include, but are not limited to, an in-depth treatment of technology-based assessments, test security, and the use of assessments for certification and licensure. These PAE guidelines apply to all assessment procedures whether or not the tests are referenced by psychological terminology (e.g., psychological testing) and apply to any assessment and evaluation procedures (e.g., job selection, performance appraisals, clinical assessment of mental health status, psychoeducational/multifactored evaluation of learning issues) that could result in psychological distress for the individual being assessed, result in the diagnosis of a deficit, or impact the individual's well-being. The guidelines apply broadly to professional competencies of psychologists conducting assessments and evaluations, including selection and evaluation of tests used, test administration, scoring, interpretation, report writing and feedback, and work within the limits of documented training and experience with specific tests used. Psychologists strive to meet the knowledge and skill-based competencies established to set a high standard in psychological practice such as professional and ethical decision-making, rights of test takers, justification for use of tests, sensitivity to diversity, and technical knowledge of tests used. The individual guidelines address clinical and knowledge competencies that are recommended to psychologists, and seek to inform psychologists' understanding of the scope of assessment practice in psychological testing.

## Audience and Stakeholders

The APA PAE guidelines are important for those directly involved in the process of testing, assessment, and evaluation, including the following:

- Psychologists who directly conduct assessments, such as administer, score, and accurately interpret tests, compose reports, or give feedback or any component thereof.
- Psychology students pursuing a higher degree.
- Those responsible for selecting psychological tests, assessing the need, determining the use of tests, and making recommendations.
- Those who purchase and/or oversee the acquisition and maintenance of psychological assessment tools and procedures.
- Those who supervise others in conducting aspects of assessment and those who work under the oversight of a higher credentialed person.
- Those who generate, use, interpret, and/or give feedback to those who have been evaluated.
- Test developers, who generate test content and are involved in collecting evidence for psychometric reliability, validity in predicting key outcomes, and equity across different subgroups of interest.
- Educators and trainers who are responsible for instruction on psychological assessment, testing, and evaluation.
- Employers who use assessments to make personnel decisions.
- The public, test takers, and those impacted by test results given to others.
- Legislative and regulatory bodies that make decisions about or monitor psychological assessment and evaluation.
- Relevant professional psychology associations.

## Statement Distinguishing Between Guidelines and Standards

The term *guidelines* refers to statements that suggest or recommend specific professional behavior, endeavor, or conduct for psychologists. Guidelines differ from standards. Standards are mandatory and thus may be accompanied by an enforcement mechanism (e.g., the *Ethical Principles of Psychologists and Code of Conduct*; APA, 2017a). By contrast, guidelines are aspirational in intent; they are not mandatory, definitive, or exhaustive. They aim to facilitate the continued systematic development of the professional and to promote a high level of professional practice by psychologists. A set of guidelines may not apply to every professional and clinical situation within the scope of that set. As a result, guidelines are not intended to take precedence over the professional judgments of psychologists that are based on the scientific and professional knowledge of the field and the ethics code. Further, federal and state laws also supersede guidelines.

These guidelines are professional practice guidelines that are applied to the practice of psychology and not to specific treatments that may be used in the practice of psychology. Professional practice guidelines are directed to practitioners and are intended to provide aspirational guidance in ethical and professional decision-making in professional practice.

## Compatibility With APA Ethics Code

These guidelines are consistent with the current APA Ethics Code (APA, 2017a). The guidelines are also consistent with and acknowledge other relevant APA policy such as the *Standards of Accreditation for Health Service Psychology* (APA, Commission on Accreditation, 2015), the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA] et al., 2014), and the *Professional Practice Guidelines: Guidance for Developers and Users* (APA, 2015b).

## Definition of Terms

### ASSESSMENT QUALIFICATIONS

The term *assessment qualifications* refers to the combination of knowledge, skills, abilities, training, experience, and practice credentials that are deemed desirable for the use of psychological tests and assessment materials. However, the term describes two types of qualifications: (a) generic assessment knowledge and skills necessary for typical uses of tests and (b) specific qualification for the responsible use of tests in specific settings and for specific purposes (APA, 2001).

### ASSESSMENT

The term *assessment* refers to a complex activity integrating knowledge, clinical judgment, reliable collateral information (e.g., observation, semistructured or structured interviews, third-party report), and psychometric constructs with expertise in an area of professional practice or application. Psychological assessment is a problem-solving process of identifying and using relevant information about individuals, groups, or institutions for the purpose of decision-making and recommendations (APA, 2001). This includes sensitivity toward the inclusion of diverse and underserved populations.

### EVALUATION

Although the terms *evaluation* and *assessment* are often used interchangeably, evaluation is a component and often end product of the assessment process.

### PSYCHOLOGICAL TESTING

Psychological tests are defined as any psychometrically derived measurement instrument that assesses the psychological constructs in which a structured sample of an examinee's behavior in a specified domain is obtained and subsequently quantified, scored, interpreted, and synthesized using a standardized process for the purpose of evaluative conclusion or recommendation (AERA et al., 2014).

According to these AERA, APA, and National Council on Measurement in Education (NCME) standards, although tests may differ vastly in terms of item and presentation design, response format, and purpose, the commonality across all test instruments is that the process by which the subject's responses are evaluated and scored is standardized. In addition, other instruments that can be included in psychological assessment protocols where subject responses are aggregated and scored in a standardized fashion, such as structured diagnostic interviewing systems (e.g., Structured Clinical Interview for DSM-5 Disorders, Clinician Version; First et al., 2016) and/or structured life history (i.e., biodata; Oswald et al., 2004) inventories and other job performance prediction instruments (Schmidt & Hunter, 1998) developed using psychological assessment principles and used with concurrent declarations of generalized predictive utility in similar settings, also fall within the purview of the AERA/APA/NCME standards.

## Documentation of Need/Public Benefit

### AVOIDANCE OF HARM

Psychological testing, assessment, and evaluation is a core component of psychological practice, treatment planning, and subsequent decisions regarding those served. Psychologists are guided by professional standards of practice in engaging in psychological testing, assessment, and evaluation to be compliant with competency expectations and to avoid harm to clients. Guidelines for psychological assessment and evaluation are important for use by psychologists to aspire to fair, respectful, and competent service delivery and treatment of examinees.

### EMERGING UNDERSERVED OR VULNERABLE CLIENT POPULATIONS

The population served by psychologists continues to reflect the shift of focus to underserved, rural, older adult, immigrant and refugee, and other vulnerable populations. Ongoing research and empirical evidence support the need to recognize culture, language, ethnicity, socioeconomic status, gender, disability, and other forms of human diversity in decisions regarding service delivery. Several factors contribute to one's ability to conduct an assessment, formulate differential diagnoses, and develop contextually and culturally relevant treatment recommendations: choice of test instruments, relative weight of multiple data points, and contextual factors relevant to decision-making all must be considered when assessing the increasingly diverse population that psychologists serve. Guidelines are important in identifying expectations for competent service to varied populations.

### PUBLIC INFORMATION

Adoption of assessment qualifications and competencies for psychologists provides the public with identifiable information regarding training, evaluation, supervision, and competence of those from whom they seek services. In turn, competency in assessment protects the recipients of assessment services by clarifying appropriate assessment processes and supporting quality in assessment service delivery, validity and accuracy of psychological reports, and appropriate use of assessment results. Further, professional practice guidelines are not legal or regulatory documents and are not enforceable by law or through professional association codes of ethics. Guidelines rather serve as an aspirational template from which psychologists can identify decision-making factors for assessment determinations for protection of the public and in conformance with the standard of within the profession of psychology.

### PROFESSIONAL GUIDANCE

*Advances in Knowledge and Practice.* Research and ongoing studies continue to expansively and specifically increase our understanding of human behavior, cognition, and affect. As a means of measurement and evaluation of these factors, existing psychological tests are revised, and newly developed tests are designed to improve accuracy and validity of testing. Professional practice guidelines are

necessary to assist psychologists in their professional development and ongoing commitment to be knowledgeable in current and improved means of assessment.

*Specialized Areas of Practice*. Psychology has continued to grow into new areas of practice and into multidisciplinary roles across different settings and specialties (e.g., forensics, neuropsychology, geropsychology, technology, integrated and primary care). Specialized areas of practice call upon psychologists to be competent in specific aspects of testing and evaluation that may extend beyond traditional and general areas of assessment. Guidelines apply to all psychologists and therefore provide a common set of practice expectations that include those engaging in specialized practice. Guidelines offer information and education not only for those engaging in specialized assessment but also for all psychologists to have an informed understanding of practice in multiple areas of specialization.

*Specialized Populations*. Professional practice guidelines express broad and general expectations for psychologists conducting assessments and evaluations, increase awareness of individual differences that may affect assessment, and assist psychologists in differentiating general assessment principles and practices from those specific to specialized settings. Psychologists, by virtue of core training, are expected to demonstrate general assessment competency. However, those who work with populations with distinct characteristics or in a service delivery context that may require specialized knowledge and skills will benefit from guidelines that specify the fields of knowledges and skills appropriate for specialized populations and settings.

*Professional Risk-Management*. Professional practice guidelines lend direction to psychologists in those areas not governed by federal or state regulations. Therefore, these guidelines assist psychologists in understanding the parameters of practice for assessment and evaluation in those circumstances when regulation have not been developed. Guidelines may also provide documentation for best practice for protection of psychologists in litigious situations.

*Development of New Technology*. The APA (2013) *Guidelines for the Practice of Telepsychology* identify ethical factors in telepsychology practice (e.g., competence, informed consent, standards of care, testing and assessment). The *Guidelines for PAE* are a critical companion document to the telepsychology guidelines, as they will assist and inform psychologists who are contemplating and/or actively adapting existing assessment protocols to online and other electronically enhanced technology platforms.

*Changing Social Norms*. Provision of psychological services has been affected by increasingly diverse populations (e.g., gender, socioeconomic status, race, ethnicity), electronic communication, increased access to information via social media, and other factors that affect the context of assessment and qualifications for assessments. Publishers are preparing online versions of traditional assessment instruments and are encouraging practitioners to use automated applications such as online administration, scoring, and interpretation. Professional practice guidelines addressing assessment competency will support these evolving changes from traditional practice in a competent and effective manner in response to the expanding scope of assessment in integrated health, services through teleconferencing, services to diverse populations, and forensic settings.

### LEGAL AND REGULATORY MATTERS

*Laws*. Changes in federal and state laws or absence of applicable laws often generate the need for guidelines. Assessment user credentials and scope of practice in testing, assessment, and evaluation have been the subject of new legislation, particularly at the state, provincial, and territorial jurisdictional levels. Guidelines for psychologists conducting assessment will provide an important resource for those who practice in jurisdictions that have revised or created new regulations as well as for psychologists in those jurisdictions that have not specifically addressed testing, assessment, and evaluation.

*Court Decisions and Case Law*. Periodically, federal laws make changes that affect the practice of psychology (e.g., qualification as an expert witness). Psychological assessment is often an integral part of forensic evaluation. Guidelines for assessment and evaluation can serve as an important deliberative component for the determination of expert status (e.g., *Daubert v. Dow Pharmaceuticals*), special education eligibility (e.g., Individuals with Disabilities Education Act), and disability determination (e.g., Social Security Administration). Psychologists who have used test instruments and other means of assessment related to a court case or eligibility decision are supported by the profession's endorsement of assessment guidelines.

## Development of the Guidelines

The *APA 2001 Guidelines for Test User Qualifications* were scheduled to expire in 2016. During the 2016 Spring Consolidated meeting, the Board of Professional Affairs (BPA) and the Committee on Psychological Tests and Assessment (CPTA), which are the two oversight groups for these guidelines, proposed the formation of a working group to identify competencies necessary for test users that would address the misuse of psychological tests. The item was discussed during the 2016 Spring and Fall Consolidated meetings, and a recommendation was made by BPA, CPTA, and the Committee for the Advancement of Professional Psychology (CAPP) to (a) endorse a working group with the purpose of developing professional practice guidelines and (b) to expand the working group to include representatives from the Board for the Advancement of Psychology in the Public Interest and the Board of Educational Affairs. Additionally, a representative from the Association of State and Provincial Psychology Boards was included.

The BPA, CPTA, CAPP, and APA staff identified the importance of a consensus document that (a) includes competencies, (b) incorporates training and experience criteria, and (c) addresses the setting-specific nature of psychological assessment as it relates to these qualifications and competencies. These criteria are central to the mission and scope of the document and the guidelines.

The reasonable options for going forward were to develop a new document or to revise the 2001 Guidelines. The Committee on Professional Practice and Standards reviewed a 2009 proposal to revise the 2001 guidelines and determined that, given the prescribed format and structure of the 2015 document, *Professional Practice Guidelines: Guidance for Developers and Users,* revision would be a more extensive project than commissioning a newly developed document. A decision was made by the work group to develop a new consensus

document of guidelines in the current 2015 format and structure. A significant identified need for new guidelines was the importance of application to clinical and health practice areas, which was not achieved in the 2001 guidelines.

The working group was composed of seven members who represented and were appointed by the aforementioned groups. The funding for the project was provided by the APA Practice Directorate.

## Selection of Evidence

The working group relied on prior documentation relevant to guideline development, including the *Guidelines for Test User Qualifications* (APA, 2001) and peer-reviewed publications from a diverse range of related topics on assessment, testing, competencies, and relevant policies and procedures. Given the setting-specific nature, purpose and scope, population targets, diversity factors, and desired education and training experience of psychological assessment, other approved guidelines and association policies were used.

# APA Guidelines for Psychological Assessment and Evaluation

# COMPETENCE

GUIDELINE 1

**Psychologists who conduct psychological testing, assessment, and evaluation strive to develop and maintain their own competence. This includes competence with selection, use, interpretation, integration of findings, communication of results, and application of measures.**

### Rationale

*Competence* is defined as "demonstrable elements or components of performance (knowledge, skills, and attitudes in their integration)" (Kaslow et al., 2009, p. 34). Competence can be diminished through not only failure of adequate initial training but also failure to self-monitor adaptation to revisions, new instruments and methods, and general advancements in assessment. The competency movement referenced as the "culture of competency" additionally specifies a "culture of assessment" outlining the importance of self-assessed competence (Roberts et al., 2005). Continual monitoring and self-assessment of competency boundaries are important in meeting standards of practice defined elsewhere. Rapid and ongoing development of instruments, procedures, norming advancements, technology, and evolving evidence-based practices can render a once-competent psychologist examiner to unethical practice through habituation of patterns and personal preferences in assessment procedure and application.

The complexity, breadth, and diversity of psychological testing, assessment, and evaluation necessitate a distinct delineation of areas of expertise. That is, psychologists consider their boundaries of expertise and practice within the legal, ethical, and professional scope of practice and competence of those boundaries. Psychologists strive to understand the limits of their expertise when the same instruments may be used for different purposes. Psychologists may be competent to administer measures of cognitive ability for the purpose of psychoeducational determinations of a learning disability but not competent to use the same tests to determine competency to stand trial. Competency is determined by both technical mastery over a particular test and the appropriately identified need for the test in the overall purpose of the assessment. (Illustrations of these diverse areas of expertise that share testing elements are noted in Guideline 4.)

Assessments are typically accompanied by referral questions. Psychologists seek to acquire the competency to determine the need and the purpose for assessment, the characteristics of the examinee, and the context and setting for the assessment typically through clinical interviews, psychometric data (e.g., cognitive, personality, performance, learning, memory, executive functioning) and collateral or supplemental materials (e.g., socioemotional measures). Without complete understanding of the need and purpose for the assessment, the characteristics of the examinee, the appropriateness of the instruments chosen, and the context and setting in which assessment occurs, interpretation and application of the results of the assessment are more likely to be limited and/or inaccurate.

In addition to technical and clinical competence, aspired-to professional competence encompasses (a) skilled communication with the examinee or client that promotes an effective working relationship; (b) the commitment to explain the risks, benefits, and possible outcomes of assessment, including in high-stakes scenarios, to the best of the examiner's knowledge and understanding; and (c) demonstration of respect for the recipients of services and the commitment to nondiscrimination and equity in professional practice. The need, purpose, and referral question are core elements in assessment decision-making; however, an environmental scan of the context in which the examinee or client is functioning related to the reason for assessment is typically considered a critical component of psychologists' competence.

Psychologists attempt to identify the most effective means by which they may remain competent in continued areas of expertise as well as in the acquisition of new skills for the purpose of expansion of scope of practice. These means may include, but are not limited to, postdoctoral courses, targeted continuing education (CE), supervision, and consultation. Engagement in assessment and evaluation often has limitations based on licensure, professional education, and training. Psychologists are encouraged to seek appropriate proficiency and/or board-level certifications through a peer-review process when such certifications are available and related specifically to the psychologists' area(s) of specialized assessment practice(s). Section 9 of the APA (2017a) *Ethical Principles of Psychologists and Code of Conduct* delineates standards of practice when performed by psychologists but does not directly address assessment competency.

### Application

Profession-wide and specialty-specific competencies are recognized and referenced by quality assurance documents and entities in psychology (e.g., Ethical Principles and Code of Conduct: APA, 2017a; Standards of Accreditation: APA, COA, 2015; Association of State and Provincial Psychology Boards, 2014) and in specific areas of practice (e.g., Hessen et al., 2018). Assessment is identified as a profession-wide competency in these and other quality assurance measures. Profession-wide competencies are evaluated by the criteria of whether they are observable, measurable, and quantifiable. This consistency is necessary to maintain continuity and objectivity across and within competencies. Assessment competency entails several functional competencies that include but are not limited to selection, use, interpretation, report of results, and use of results in response to the purpose of the assessment.

*Selection of tests or evaluation measures.* Psychologists seek to become knowledgeable of the psychometric characteristics of test instruments as well as other factors likely to impact the applicability of specific test instruments and evaluation measures to the assessment question at hand (e.g., reading levels, physical requirements, cultural background, characteristics of the

standardization/normative group). Similarly, psychologists strive to maintain appropriate knowledge of the context and environment in which the assessment is to be performed, as well as any additional relevant factors that might affect outcome and recommendations. In addition, psychologists are encouraged to obtain and to review carefully available collateral and supplemental information such as treatment histories, previous assessments, third-party observations, and/or performance reviews for consideration and incorporation when reaching conclusions and/or making recommendations. Other factors to be considered in selection of measurement instruments may include characteristics of the examinee, appropriateness of the normative group, cultural relevance, and alignment of the test instrument construct with the purpose of the testing.

*Use.* Fidelity to the purpose and use of a given instrument are often critical to the validity of a test instrument as well as the accuracy of the assessment results. As such, psychologists consider the reason for testing as well as the anticipated contribution of the selected test or battery to the operative referral question and/or process to which the results will be applied. In this regard, specific examples of the utility of assessment measures are often prescribed in the technical manual of an instrument, and as such, psychologists strive to thoroughly and critically evaluate the applicability and/or supporting scientific evidence for use of a given instrument in settings that differ substantially from those specified by the publisher.

Psychologists understand that test instruments are not typically used as singular measures but rather are integrated with other standard measures as well as nonstandardized yet valuable data points (e.g., collateral interviews, behavioral checklists, paper review of prior documents). Tests and other measurement instruments can be cited in the technical manual for multiple uses. Psychologists remain aware that although their area of expertise may support use of a test for a particular purpose, other uses of the test may fall outside the psychologists' scope of competence. Similarly, psychologists are advised that assessment is most comprehensive and accurate when multiple data points are used to arrive at a determination (e.g., diagnosis, recommendation, disposition), and as such they are encouraged to include all additive data sources in drawing conclusions, inferences, and decisions.

*Integration of data.* Psychologists seek the competency to integrate all data points and other form of findings in the writing or oral reporting of results. Multiple data points include but are not limited to standardized tests, clinical interview, collateral reporting, behavioral checklists, environmental context, and client variables. Integration of data points is a distinct skill from interpretation but results in interpretation and the formation of case conceptualization, which advances decision-making and initial formation of recommendations. The weighting of data points in the integration of data is a process that considers the cultural, ethnic, and other diversity variables that influence the context and interpretation of data points.

*Interpretation.* To accurately interpret findings, psychologists strive to understand the conceptual meaning of scores and the technical range of interpretation of any given set of individual scores (see Section 2 of these guidelines). Accurate interpretation is dependent on the psychologist's ability to integrate multiple sources of data points. Insofar as primary sources of data can be inconsistent rendering a clear determination difficult, psychologists seek to develop the knowledge and skills to critically evaluate these apparent data inconsistencies and arrive at the most viable interpretation of the data that serves the purpose of the assessment accurately (Hopwood & Bornstein, 2014).

Psychologists aspire to reflect accuracy in their interpretation of test and assessment instrument results and to carefully consider and control potential sources of error and/or bias, particularly when these errors may contribute to a diagnosis, recommendation, disposition, or other high-stakes decisions (e.g., custody, employment, guardianship determination, competence and decisional capacity, disability compensation, incarceration). In this regard, errors in reporting assessment results can include overinterpretation, inconsistent interpretation, selective interpretation, and/or other misinterpretations of results. Although sources of these errors can be attributable to lack of technical knowledge, the most common sources of bias effecting interpretation of psychological assessment data include distortions and subjective weighting errors based on preconceived beliefs, and/or other intervening factors such as anchoring effects (i.e., overweighting initial data), attribution effects (i.e., favoring data from one source over another), and/or confirmation effects (i.e., selectively weighting data based on personally held beliefs). An excellent and thorough discussion of these and other biases effecting psychological assessment that affect accurate interpretability of multiple data can be found in Reynolds and Suzuki (2013).

*Reporting results.* Psychologists typically deliver assessment results through a written report often enhanced by oral explanation, depending on context and purpose. Because the functional definition of scores often employs terminology that may not be in the typical vernacular of examinees and recipients of assessment feedback, psychologists are encouraged to become aware of the meaning of scores that underlie the interpretation and strive to describe findings in a report in a suitably interpretable manner. Psychologists seek to become aware of the preferred language of the report recipient(s), reading level, and general ability to comprehend the information.

Psychological reports typically include a summary section that integrates key elements from the findings in relation to the purpose for the assessment and cohesively presents the information in an organized, comprehensible, and interpretable manner. Psychologists attempt to develop the skills to create summaries of the key findings and to identify recommendations based on the summaries that are recognized in the field as effective treatment components. Psychologists aspire to develop the competency to construct comprehensive evidence-based recommendations, recognized in the field, that address possible treatments or accommodations resulting from the findings. Psychologists working in proficiency areas that result in disposition, selection, third-party decisions, and other high-stakes assessment purposes are encouraged to develop competencies commensurate with their areas of specialty practice and expertise.

*Explanation of use and implementation.* Psychologists seek to develop the competency to explain to the recipients of the report or findings how the recommendations derived from the summaries can or

will be used to accomplish the purpose of the assessment. The various areas of specialty or proficiency in psychology can render significantly different findings and recommendations or can render similar findings and summaries that result in very different recommendations and are implemented in significantly different ways. The scores and summaries of cognitive tests of intelligence, executive functioning, memory, achievement, and personality could be the same for two people, yet the recommendations and implementation could differ depending on the purpose of the assessment (e.g., psychoeducational, child custody, competency to stand trial, preemployment screening, fitness for duty evaluations management selection, diagnostic classification, developmental disability, workers' compensation, and immigration hardship waivers). Psychologists are best advised to self-monitor their professional boundaries to appropriately conduct the elements of assessment or evaluation within their scope of practice.

GUIDELINE 2

**Psychologists who conduct psychological testing, assessment, and evaluation seek appropriate training and supervised experience in relevant aspects of testing, assessment, and psychological evaluation.**

### Rationale

Training programs of recent years incorporate competency expectations for assessment and conducting psychological testing that reflect accreditation standards, ethical standards, jurisdictional laws, and regulations (e.g., Hessen et al., 2018). Psychologists in practice have experienced variable quality and content in initial graduate training and subsequent CE, proficiency acquisition, or self-directed access to education and training. That is, a trajectory has not been established for the acquisition of competence, the maintenance of competence, supervised work experience, or criterion-based expectations for practitioners who expand their scope of practice post graduate training. The APA Ethics Code (APA, 2017a), the AERA et al. (2014), and other professional associations delineate standards of practice in assessment, measurement, and evaluation. The APA Ethics Code requires that psychologists practice within their boundaries of competence (APA, 2017a) and that psychologists in practice who seek to expand their scope of practice in assessment undertake relevant education, training, supervised experience, consultation, or study (APA, 2017a).

Assessment, among other specialty and proficiency areas in psychology, has evolved over recent years given the expansion of specialty practice in assessment and evaluation (e.g., psychoeducational, forensic, child custody, geropsychology, personality, neuropsychological, development). New and revised testing instruments and materials continue to develop given psychometric improvements (e.g., norming methods), research findings on learning, motivation, memory, and other factors. Knowledge, skills, and attitude that would meet standards of practice at an earlier time would likely be inadequate in contemporary practice. Not only the acquisition of competence but the maintenance of competence is typically required by the aforementioned standards of practice. That is, failure to gain initial competency and failure to maintain competency may both result in unethical practice (APA, 2017a).

The means by which psychologists competently expand scope of practice post formal training may occur through CE, seminars, supervised experience, and consultation. Further, the recent and developing application of technology to knowledge and skills acquisition (e.g., webinars, online CE) increases opportunities for self-directed study. Psychologists attempt to identify the most effective means through which to gain the desired set of competencies. Psychologists who want to expand their scope of practice in psychoeducational assessment are likely to pursue a somewhat different set of knowledge and skills than those who wish to conduct forensic evaluations, acceding, however, to the point of common foundational principles and standards.

Assessment is a fluid and dynamic activity that calls for focused and ongoing attention to maintenance of competence. Psychologists will be mindful that competency evolves as subject matter matures over time and that acquisition and maintenance of competency is an ongoing process that requires self-assessment and awareness of contemporary standards of practice. The ongoing process of self-assessment includes attention to the potential emergence of complex interpersonal dynamics between client and assessor (e.g., Bram & Peebles, 2014; Yalof, 2019).

### Application

Psychologists who wish to acquire or maintain competence in assessment recognize the importance of foundational competencies and special focus competencies. With the exception of specialists in neuropsychology (Hessen et al., 2018; Roper et al., 2018 Smith, 2019), a trajectory has not yet been established for the acquisition of competency. These foundational areas are described as including measurement theory and psychometrics, the components of conducting assessments (e.g., selection, administration, and scoring), integration of data points, interpretation of scores, conceptualization, and communication of results and recommendations. Psychologists may acquire foundational knowledge through coursework, webinars, self-study, CE, and other sources. These areas of foundational assessment factors are well described in several documents: *PAE Guidelines* (this document), the AERA Code of Ethics, the Recommended Competencies for Users of Psychologist Tests, the NCME, and the Standards for Educational and Psychological Testing.

Psychologists consider the decision-making factors that meet the standards of practice in their select subject matter area. Psychologists who identify an area of expanded scope of practice strive to determine their existing level of competency as well as the knowledge and skills to be attained. This determination can be made through consultation with those who are recognized for their practice in that area of expertise. Multiple methods of acquiring knowledge and skills can be adopted to include coursework, workshops, webinars, and other CE that is specifically focused on the target area of practice. Psychologists are encouraged to develop their specialty skills through subject matter instruction and

education (e.g., forensics, child custody) coupled with consultation. The subject matter component of competency frequently includes supervised experience. Knowledge of the subject matter is integrated with skill development, typically under consultation or supervision of an expert in the area.

Psychologists comply with the APA (2017a) *Ethical Principles of Psychologists and Code of Conduct*. Section 2, Competence, specifically delineates the importance of relevant education, training, supervised experience, and consultation when expanding scope of practice into areas new to psychologists. Further, the maintenance of competence requires ongoing effort to meet the standard of practice. Failure to maintain competence can result in inadvertent practice beyond the data necessary to substantiate findings, inappropriate use of instruments, and inaccurate interpretation. Psychologists recognize the importance of fair and equitable treatment in conducting assessments with diverse and underrepresented populations. They strive to use methods that are age and language appropriate and, when needed, employ the services of professional interpreters to achieve more accurate assessment results.

GUIDELINE 3

## Psychologists who conduct psychological testing, assessment, and evaluation strive to be mindful of the potential negative impact and subsequent outcome of those measures on clients/patients/examinees/employees, supervisees, other professionals, and the general public.

### Rationale

Psychologists recognize their ethical and, in some situations, legal obligations regarding the prevention of negative impact that could result from their selection, administration, interpretation, or reporting of results from psychological tests and measurements. The exception to this may be forensic evaluations, which often require an independent opinion that may negatively impact the examinee. Existing ethical and legal obligations to prevent negative impact stems from the recognition that all tests and measurement procedures have both appropriate and inappropriate uses depending on purpose and setting of the assessment, the limitations of the test or assessment procedure to address that purpose in that setting, the characteristics of the test taker in relation to the normative sample that supports the valid interpretation of the test results, and the human confidence that can be placed in the interpretation of the results obtained for that client in that setting for that purpose. In addition, psychologists remain alert to any ethical and, in some cases, legal obligation to protect the client from misuse or misrepresentation of the data. Further, psychologists may find themselves called upon to guard against reports generated by untrained individuals who could misunderstand or misuse this information in such a way that could harm a client, a recipient of test results such as an employer, a court, or a health care system or test publisher. Copyrights on proprietary assessment material are designed to be unavailable to individuals without proper training and ethical obligations to maintain security.

### Application

Psychologists strive to understand and maintain their competence in the selection, administration, and interpretation of psychological tests as well as their client's ability to engage in and understand fully the assessment process including the potential risks and negative outcomes that could result in addition to any positive benefits they may be seeking. Psychological tests are continually changing through test revisions, research findings about appropriate or inappropriate applications or unanticipated complications, or threats to validity in using the test with a particular client population or for a particular purpose or in a particular setting. Psychologists assume an appropriate degree of responsibility to understand the strengths and weaknesses of all assessment procedures they employ, to use the latest versions or forms of all tests and procedures, and to seek information and understanding about any limitations or concerns in using the tests they have selected with the client in the setting and for the purpose of their assessment. Psychologists also strive to prevent or minimize misuse of assessment results that can be anticipated and to correct misuse or misinterpretation of assessment findings that come to their attention following the release of this material.

Psychological assessments are typically used in the hope of gaining information that can direct an intervention, ameliorate a problem or difficulty, gain insight into abilities and skills, or inform decision-making. Sometimes the results of psychological assessment reveal findings that could be seen as detrimental to the interests of the examinee achieving their goals or problematic in some way unrelated to the original purpose. Where appropriate, psychologists strive to obtain informed consent, which may include capacity to consent and freedom to withdraw, such that the examinee understands that the results of assessment may not provide the desired outcome and, depending on the context of the assessment, may not be protected from unwanted disclosure. There may be such instances in which safeguards should be taken if the disclosure of the purpose of testing will spoil or influence the results. These safeguards might include debriefing after the assessment rather than full disclosure prior to beginning the assessment. For example, an individual who agrees to undergo psychological assessment to demonstrate appropriateness for advancement in their job may not receive the individual's desired result, and results would be shared with the potential employer. The examinee agrees to release the results of the assessment to their employer as a condition of employment. The results of the assessment reveal significant difficulties in job-related abilities that might not only preclude promotion but even threaten their continued employment. This highlights the need to recognize that there are many consumers to assessment results, including an employer or organizational client. The psychologist is typically expected to seek the client's informed consent and full understanding of the range of outcomes that could result from the assessment and the potential recipients who might gain access to the results before consent is given.

Psychologists strive to use tests appropriately and to understand the strengths and weaknesses of the assessment procedures they use and to correct any misuse they

discover not only for the benefit of the client being assessed or the client or agency or sponsor requesting the evaluation but also to maintain the reputation of the profession of psychology and the public's confidence that psychologists are competent, current, and responsible in their selection, administration, and interpretation of assessment procedures.

GUIDELINE 4

## Psychologists strive to consider the multiple and global settings (e.g., forensic, education, integrated care) in which services are being provided.

### Rationale

Psychological tests are used in a variety of settings for a variety of purposes. Validity is not a unitary property of the test instrument. Rather, validity evidence is evaluated within the context of these multiple settings and purposes. In addition, most tests are developed within a cultural or regional context where the test developer has an intended target population of test takers for a specific purpose. The psychologist, however, strives to recognize when the selection and use of this instrument deviates from the expected or intended purpose and recognize the significance and implications of such deviations. The psychologist considers these deviations when interpreting test results for a particular client and reporting results for a specific purpose and within a specific context. These deviations can include not only using a test in a setting for which it was not designed (e.g., a broad assessment of academic achievement intended as a screening tool for adults used for diagnostic assessment of a child's learning difficulties) but also using a test designed, developed, and normed in one country or region of the world in a different country or region. The psychologist strives to understand the significance of such cultural and linguistic deviations and to acknowledge possible influences, including limitations and potential errors, in their use and interpretation of tests.

### Application

Competence in psychological assessment is typically situation specific or setting specific. A competent psychologist maintains a foundational fund of knowledge and skill about tests and test procedures in general (e.g., validity, reliability, normative population) that supports selection and use of a test instrument or assessment procedure in a given situation. However, a specific individual is rarely, if ever, a competent user of psychological assessment in all the settings and contexts where assessment tools are used. Psychologists strive to develop the contextual knowledge and skill to be competent in one or perhaps two of the primary arenas where psychological tests are employed. These primary arenas may change as the profession evolves and/or as the professional in question develops additional skills and experience. At present, the following areas of professional practice tend to be the primary areas in the use of psychological assessment: clinical, forensic, neuropsychology, police and public safety psychology, educational/school psychology, geropsychology, industrial and organizational psychology, employment selection/coaching, and integrated health care.

Each of these practice settings calls for the psychologist to develop adequate knowledge of practice standards, legal standards, collaborative professional roles and desired purposes, methods and outcomes that call for the use of psychological tests and procedures. For example, a psychologist working in a forensic setting is expected to be not only suitably competent in their knowledge of psychological testing, individual differences, psychopathology, and other aspects of psychological theory and research but also appropriately knowledgeable about the legal context in which the results of their psychological testing will be reported and used. Depending on their area of practice, a competent forensic psychologist might be expected to know courtroom procedure; rules of evidence; rules of jury selection; or legal definition or elements of terms such as *insanity*, *dangerousness*, and *consciousness of guilt* (Weiner & Otto, 2014). The work of neuropsychologists may, at times, overlap with the work of forensic psychology (Demakis, 2012), calling for attention to both legal and clinical issues associated with this type of specialized assessment.

Similarly, a psychologist working in a school environment with a task of identifying children in need of special educational services not only strives to be competent in knowing how to select, administer, and interpret a psychological test of cognitive ability, academic achievement, or emotional adjustment and functioning but also seeks to know and understand special education law and requirements around eligibility for services (Wright & Wright, 2015), as well as the student's cultural context, the classroom context, and how it affects manifestation of learning and adjustment difficulties. In addition, a suitable level of knowledge about best practices in classroom methods is important to make helpful and appropriate recommendations of educational interventions based on test data gleaned from the use of psychological tests.

In the employment/coaching area of practice, the competent user of psychological tests strives to understand the purpose of any evaluation, who is the client, who has a legal right to access the results of an assessment, and whether the procedures selected to be used provide sufficient reliability and validity for this purpose and this client in this context. Psychologists working in this area seek to understand employment law and the legal standards for what constitutes employment discrimination versus assessment of job appropriate aptitudes and skills.

With clinical service providers and those working in health care delivery systems, the competent psychologist strives to understand how the results of any psychological assessment will be used so that the appropriateness of that use can be evaluated and any cautions or limitations in use of the findings can be noted in any report. In systems of integrated care where psychologists work as part of a multidisciplinary team, the psychologist strives to make sure that the presentation of assessment results is understandable for the other team members as well as the client and presented in such a way that any possible complications or limitations in the interpretation of the findings is made known and addressed in the report.

Psychologists also strive to be mindful of the problems associated with the increasing use of the psychological tests globally, including delivering tests across legal boundaries of states, provinces, territories, or countries. In some instances, using psychological tests across legal boundaries, such as states in the same country, may provoke a concern about intentionally or unintentionally circumventing

mechanisms of public protection such as state licensing laws but not necessarily issues about validity or appropriateness of the test use. In some situations where tests are given to test taker populations that were not represented in the normative sample, or where no validity studies are available to demonstrate that the same constructs are being assessed adequately in this population, then such use of psychological tests across legal boundaries may be not only inconsistent with regulatory standards but may also wind up being deemed invalid, incompetent, and unethical.

Psychologists seek to use psychological tests only in contexts and with populations and for purposes that are valid and appropriate based on empirical evidence that the normative sampling, the language or translation used, the administration procedures employed, and the clarity and accuracy of the results reported are legal, reliable, valid, and appropriate.

# PSYCHOMETRIC AND MEASUREMENT KNOWLEDGE

GUIDELINE 5

**Psychologists who provide psychological testing, assessment, and evaluation demonstrate knowledge in and seek to appropriately apply psychometric principles and measurement science as well as the effects of external sources of variability such as context, setting, purpose, and population.**

### Rationale

The organization, as well as some text and references that appear in this section, has been sourced from the *Recommended Competencies for Users of Psychological Tests*, originated by the APA CPTA in 2015. To effectively choose, administer, interpret, and evaluate psychometric instruments, practitioners are encouraged to maintain thorough and current working knowledge of the psychometric principles that underlie the design and utility of the test instruments they use. The primary components of this knowledge are described under the following headings.

*Descriptive Statistics*. Descriptive statistics are the foundational components of test construction and interpretation. Psychologists should be familiar with the basic descriptive functions defining the composition and distribution of standardization samples upon which instruments are based and apply that knowledge when choosing instruments and/or interpreting individual results. Similarly, suitable knowledge of the characteristics of means and standard deviations is critical when comparing individual performance on various test scales, especially those that are norm referenced. Common descriptive statistics relevant in this regard include measures of central tendency (e.g., mean, median, and mode) and measures of variation (e.g., variance and standard deviation). Likewise, correlations and other indices of association (e.g., chi-square) are commonly used for examining the degree of convergence or divergence between two or more test score scales, whereas frequency distributions of scores describe the varying levels of the construct or other predicted criterion outcome found in groups of test takers.

*Test Theory*. Critical evaluation of the efficacy and applicability of individual test instruments to the assessment question at hand, as well as the confidence with which results may be interpreted and conclusions drawn, requires working knowledge of the fundamental theories and techniques of test construction. Competency in this regard typically includes knowledge of the conceptual foundations, assumptions, and extensions of the basic premises of classical test theory (Kline, 2000), such as item difficulty, item discrimination, item and test information functions, latent trait or ability parameters, generalizability theory (Brennan, 2001), and/or item response theory when appropriate (Embretson & Reise, 2000). In this regard, psychologists strive to understand the advantages and disadvantages of these test theories in operationalizing the construct being measured to ensure appropriate inferences are made.

*Scaled Scores and Transformations.* Individual results of most tests are derived from item responses, which are grouped together in some manner to form scales and then subsequently either reported as raw scores or transformed mathematically in some manner and presented as normative comparative or standardized scores. As such, knowledge of the process and assumptions through which these groupings and transformations are created is typically considered essential for proficient test use.

*Reliability/Precision and Measurement Error.* According to the AERA et al. *Standards*, the reliability/precision of scores depends on how much scores vary across replications of a testing procedure, and analyses of reliability/precision depend on the kinds of variability allowed in the testing procedure (e.g., over tasks, contexts, raters) and the proposed interpretation of the test scores. Several approaches to the estimation of reliability/precision of test scores (e.g., Haertel, 2006) vary in their applicability and appropriateness across measurement situations. Psychologists are encouraged to become familiar with various approaches to reliability/precision estimation, factors that influence the index (or set of indexes) of reliability/precision that is appropriate for their given situation, factors that can influence the magnitude of those indexes, and professional standards pertinent to assessing the reliability/precision of test scores (see Chapter 2 in AERA et al., 2014).

*Validity and Meaning of Test Scores*. According to the AERA et al. (2014) *Standards*, validity refers to "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (p. 11). Thus, psychologists strive to understand that validity is not an inherent property of a test but rather refers to the degree to which evidence and theory support the use of a test for a particular purpose. In evaluating tests for a particular purpose, psychologists should become suitably aware of the five sources of validity

evidence described by the AERA et al. *Standards*: validity evidence based on test content, response processes, internal structure, relations to other variables, and testing consequences. For a valid inference to be drawn based on a test score, psychologists are encouraged to demonstrate that the scores generated are directly, demonstrably, and consistently related to the outcome or purpose for which the test instrument is used. Conclusions and/or recommendations resulting from use of instruments are expected to be fair; minimize bias; and are consistent with applicable standards of practice, policies, and laws.

In addition to seeking an understanding the different sources of validity evidence, psychologists strive to develop an understanding of strategies for obtaining, evaluating, and establishing each source of validity evidence; the limits of any one source of validity evidence; and the implications of how different sources of evidence can be integrated into a comprehensive validity argument to support the use of a test for a particular purpose (Kane, 1992, 2013).

The test developer and the test users may be considered jointly responsible for development and evaluation of validity evidence to support the use of a test for a particular purpose, especially when the test is used in a specialized setting and/or with a specific population whose characteristics may differ from the original population or setting upon which the test was originally developed. Similarly, when interpreting test results in settings and with populations that differ from those studied by the test developer, psychologists strive to account for the measurement effects likely to be associated with these differences and report these differences (e.g., restrictions in range and base rate of outcome criterion; Finn, 2009).

### Application

Psychologists may familiarize themselves with basic statistical principles related to test validation by reviewing available knowledge sources and/or CE opportunities. In this regard, although certainly not the only source available, an excellent practical review of descriptive statistics, including measures of central tendency, variation characterization of normal curves and frequency distributions, correlations and other indices of association degree of convergence or divergence found in groups of test takers, as well as other basic statistical principles, may be found in Keller (2006). In addition, a comprehensive review of fundamental measures of association as well as various models of prediction and test construction prediction may be found in Bandalos (2018).

Psychologists strive to understand and know when to apply the various methods for representing test information (e.g., achievement/mastery levels, diagnostic classifications, raw scores, standard scores, percentiles). Relevant concepts include types of scales, types of scores (e.g., raw, transformed, percentile, standard, norm-referenced, criterion-referenced), scale score equating, and methods for establishing cut scores (e.g., Cizek & Bunch, 2007). Psychologists are also encouraged to be aware of the advantages and disadvantages of automated scoring.

There is no absolute standard for establishing that an inference one desires to make based on test scores is "valid" for a particular purpose. Thus, psychologists strive to be able to locate and evaluate available validity evidence to judge the suitability of using a test for an intended purpose. Psychologists seek to consider the positive and negative consequences of test administration and score use, to help in ensuring that the testing process will provide more good (e.g., diagnostic information that is useful for educational or clinical purposes) than harm (e.g., negative stereotyping) to examinees. In addition to the AERA et al. (2014) *Standards*, guidance for evaluating the validity of the use of a test for a particular purpose can be found in the *Principles for the Validation and Use of Personnel Selection Procedures* (Society for Industrial and Organizational Psychology, 2018) as well as in federal employment statutes such as the Civil Rights Act Title VII (1964, 1991) and Part 1607 of the Equal Employment Opportunity Commission (1978), and the Americans with Disabilities Act (1990).

When psychologists evaluate the validity of inferences drawn from test scores, it is typically for use of scores in a given setting or with a given group of individuals. Validation evidence presented by test developers and publishers primarily addresses the use of the test scores for specific stated proposes and with subject cohorts similar to the standardization sample(s) on whom the test was originally constructed. Thus, psychologists strive to remain cognizant of the specific setting in which a test will be used, particularly if this setting differs from the original setting or purpose for which the test was designed. Similarly, when a test is used for purposes that differ from those investigated by the test developer, and/or when the characteristics of the population tested differ from the characteristics of the publisher's standardization sample, psychologists are encouraged to evaluate the validity evidence for this specific use.

# SELECTION, ADMINISTRATION, AND SCORING OF TESTS

GUIDELINE 6

**Psychologists who conduct psychological testing, assessment, and evaluation endeavor to select (a) assessment tools that demonstrate sufficient validity evidence for their uses, sufficient score reliability, and sound psychometric properties and (b) measures that are fair and appropriate for the evaluation purpose, population, setting, and context at hand.**

### Rationale

Psychologists seek to provide assessment services only within the boundaries of their competence that is based on their education, training, supervised experience, consultation, study, and professional experience. Psychological testing and other assessment procedures are areas of professional practice in which psychologists have been trained and are uniquely qualified to conduct. Psychologists are thus encouraged to be knowledgeable about and account for the impact of test results in diverse populations and across different settings, as well as the limitations of measures even when guidelines are followed for test selection. When conducting assessments and evaluations, psychologists are aware of the responsibility that may be ascribed to them for the appropriate choice of measures that reflect evidence-based practice, sound psychometric properties, and awareness of the context, including patient characteristics, that can impact test results. Careful and informed measurement selection benefits the intended examinees by ensuring valid assessment, fair utility of results, and generation of recommendations that are applicable to the intended context. These guidelines are consistent with the standards articulated in the most recent edition of *Standards for Educational and Psychological Testing* (AERA et al., 2014).

### Application

To choose the best test or instrument for the testing, assessment, or evaluation purpose for the population, setting, and context at hand, psychologists strive to determine who they are evaluating, for what reason, where, and under what conditions, as well as what domains are to be addressed. The more information gathered prior to the testing, assessment, or evaluation and an understanding of these characteristics or variables should assist in selecting a reliable and valid test or instrument to use among the myriad ones available commercially or in research. Choosing a test or instrument may seem daunting, but there are ways to make decisions efficiently. For example, knowing the question(s) to be answered ahead of time shortens the selection time. Consulting with colleagues who conduct testing, assessments, or evaluations or who are engaged in test or instrument development or research in a particular area may prove very beneficial.

Psychologists are encouraged to become familiar with the psychometric properties of any test or instrument they use to make screening, diagnostic, or intervention decisions and/or status recommendations. Psychometric characteristics of tests and other instruments include but are not limited to standardization, reliability, and validity. These psychometric characteristics are foundational to decision-making and form the basis of interpretation. The following steps, adapted from Alfonso (2004), may be considered by psychologists to assist in selecting the best test or instrument to use:

- Read and evaluate reviews of the tests or instruments and research studies that used them, including meta-analyses, systematic reviews, and specialized practice parameters, in addition to edited handbooks and volumes on assessment.
- Read the test or instrument publisher's technical manual(s) and review carefully the design, standardization sample(s) characteristics, descriptive statistics of standardization sample(s), composition and independence of scales, and supporting evidence of validity and reliability.
- Make sure the tool is accessible, appropriate for the purpose, and in the language that one would need to use, within the context of evidence-based practice.
- Become familiar with the administration instructions, scoring, and accurate completion of test protocols and record forms. A thorough review of the items also provides psychologists with an opportunity to evaluate the content validity of the instrument with particular reference to the referral question and/or the purpose to which the results will be applied.
- Practice and observe the administration of an instrument with an appropriate colleague. There is no better way to learn a test or instrument than to be the examinee and know what it is like to be asked to solve problems, rate behavior, or disclose personal information.
- Attend training seminars or workshops with the test or instrument's author(s) or experts in the field.
- Explore publishers' websites for up-to-date information on their tests and instruments.
- Join a professional listserv where intellectual discussions take place.
- Attend a graduate class or several sessions of a course to learn about new developments in evaluation or to learn about specific tests or instruments.

Several tests and measurement texts and resources are available for psychologists to update their psychometric knowledge base and test, assessment, or evaluation skills. Some examples include Bandalos (2018); Graham et al. (2013); *Psychological Testing: Principles, Applications, and Issues* (Kaplan & Saccuzzo, 2017); and *Standards for Educational and Psychological Testing* (AERA et al., 2014). Reading or even perusing these resources can be helpful when evaluating the psychometric properties of tests and other instruments and in selecting the best one to answer the referral questions and address the purpose of the testing, assessment, or evaluation.

GUIDELINE 7

**Psychologists who conduct psychological testing, assessment, and evaluation strive to use multiple sources of relevant and reliable clinical information collected according to established principles and methods of assessment.**

### Rationale

Individual performance on psychological tests is only one piece of assessment and is conceptualized in a context of presenting concerns, reason for referral, background, course of illness, influential factors, and population-specific contributions that are secured from multiple sources. These may include clinical interview with the examinee, clinical interview with sources other than the examinee, completion of valid self-report and third-party report measures, observation of behavior, and review of relevant records. Additionally, psychologists strive to recognize that more than one reliable and valid measure is appropriate to assess all relevant domains of functioning to examine convergence and deviation of findings. Convergence and divergence of information from multiple sources informs the assessment process and increases confidence in and appropriate utility of test results (AERA et al., 2014). Information from multiple resources is valuable and useful but may not be feasible in some situations.

### Application

Best practices in testing, assessment, and evaluation suggest a multisource, multimethod, and multisetting approach because human behavior and functioning are highly complex (Sattler, 2014). Typically but not always, the younger the client, the more variable behavior is with different people and across settings (Bracken & Theodore, in press). Similarly, it is not uncommon to obtain moderate to markedly different ratings from various individuals in the client's life. Thus, multiple sources including parents, extended family members, colleagues, supervisors, teachers, and sometimes community members are typically considered important in the testing, assessment, or evaluation process (Achenbach, 2017; Burns & Haynes, 2006; Stage et al., 2006). It is the psychologist's challenge to integrate data from multiple sources. This process includes drawing upon research, theory, and clinical experience in situations where different data sources are seemingly incongruent (Bram & Peebles, 2014). Likewise, psychologists strive to recognize the advisability of using multiple methods more often than not. For example, standardized, norm-referenced tests, interviews, behavior observations, performance-based measures, patient–examiner relationship, and rating scales are some methods that may assist in gathering the information needed to address the concerns of the client. It may be rare but nonetheless important to test, assess, or evaluate clients across settings such as school, home, and place of work. Psychologists should strive to be aware of situation-specific behavior and to address the possibility or variable functioning based on setting. Psychologists attempt to remain cognizant of the importance of interpreting tests and measures in a way that is consistent with their intended use, empirical literature, and other evidence-based factors.

GUIDELINE 8

**Psychologists who conduct psychological testing, assessment, and evaluation strive to be aware of the need for test selection, scoring, and administration to reflect the appropriate normative comparison, situational influences, effort, and standardized administration as indicated.**

### Rationale

In psychological assessment and testing, psychologists recognize the importance of considering age, gender, ethnicity, primary language, and individual influences when selecting and administering appropriate tests. This procedure allows for valid normative comparisons, classifications, ipsative profiles, and other interpretations to inform recommendations based on test results. Additionally, the agedness of the norms and the continued relevance and definition of the constructs being measured by a particular test can be particularly important considerations in test selection. There is also increasing literature to indicate the importance of assessing effort, both in symptom presentation and with suboptimal performance, when conducting psychological testing. Examinees may underperform for many reasons, and not adequately assessing effort limits the interpretation of test results. Without systematically assessing effort, it becomes difficult to discern if variability and patterns of test results reflect actual performance or the influence of low effort, motivation, or some other factor besides ability. Psychologists strive to conduct assessment in adherence with standardized administration procedures to support valid interpretations of norms, classification decisions, comparisons, and other test score–based inferences and applicable to the circumstances. Adhering to standardized testing conditions will minimize confounds that could lead to misinterpretations of test results. Psychologists are also encouraged to judiciously test the limits of standardized administration when doing so is necessary to answer referral questions, determine meaningful treatment implications, and/or clarify conditions under which functioning varies (see Bram & Peebles, 2014).

### Application

Psychologists strive to select the most appropriate test or instrument for a specific purpose and seek to cultivate knowledge of testing, assessment, or evaluation practice in the context area and associated norms when more than one normative set is available. Knowledge of test or instrument characteristics such as psychometric properties (presented earlier), basis in theory and research, and normative data (where appropriate) is typically expected to influence test selection. For example, normative data or decision rules may not be accurate when (a) important features of the client are not represented in the norm group, (b) administration or scoring procedures do not follow those used in standardizing the test or instrument, (c) characteristics of the test or instrument may affect its utility for the situation (e.g., ceiling and floor effects), (d) the test or instrument contains tasks

that are not culturally relevant to the client, or (e) the validity evidence does not support decisions made on the basis of the obtained data. In these cases, alternative instruments and/or other data sources might be considered. Psychologists strive to remain aware of the importance of integrating examinee motivation and standardized assessment of effort (e.g., symptom validity testing, performance validity testing) into performance interpretation.

Psychologists seek an understanding of how the construction, administration, scoring, and interpretation of tests or other instruments match the purpose of testing. Mismatches in these dimensions between the selected test or instrument and the testing purpose are important factors that psychologists strive to consider, as these may invalidate usual interpretation of a client's performance. For psychologists to select an appropriate test or instrument for a particular use, it is important that they seek to understand and consider the intended use of any scores, the method and procedures used to develop or revise the test or instrument under consideration, the definition of the construct that the test or instrument purports to measure, and the definition of the test or instrument purpose and its intended context of use.

Standardized administration is most important in that the validity of the responses leading to data points may be compromised when adherence to standardization is not met. Important points that contribute to standardization of administration include (a) a high level of knowledge of instructions given the presentational variance within instrument; (b) awareness of instructions on feedback, query, prompts, and response to questions; (c) facility with supplemental material and audio/visual components; (d) inclusion of behavioral observations; and (e) compliance with equipment and placement of materials.

Scoring of responses lends itself to error more than most other aspects of testing in that there is often little uniformity within instrument and errors in scoring are unforgiving in that the examiner cannot typically return to make corrections but rather the subtest becomes spoiled. Examiners strive to be alert to common mistakes, which include (a) miscalculation of base and ceiling, (b) simple arithmetic errors, (c) reversal and discontinuance mistakes, (d) failure to record examinee comments in addition to scores, and (e) imprecise timing on timed tests.

Knowledge about procedural requirements, confidentiality of information, communication of results, and security are important for many applications, as is knowledge of standardized administration and scoring procedures and understanding a psychologist's ethical and legal responsibilities and the legal rights of test takers. Psychologists are encouraged to understand the legal and ethical issues related to the release of test data, including issues of confidentiality, depending on the context of the testing, assessment, or evaluation and the characteristics of the client such as motivation or effort expended to perform successfully. Psychologists strive to be able to explain results and limitations to diverse audiences, and to include in the report the purpose of the test, assessment, or evaluation and the setting in which it occurred.

# DIVERSE, UNDERREPRESENTED, AND VULNERABLE POPULATIONS

GUIDELINE 9

**Psychologists who conduct psychological testing, assessment, and evaluation strive to practice with cultural competence.**

### Rationale

The foundation to the integration of individual differences and cultural identity to psychological assessment is based on the value that every person assessed should have an opportunity to demonstrate their standing on the talents, capacities, and traits being assessed without being impeded by construct irrelevant sources. Therefore, psychologists endeavor to select and use instruments, assessment techniques, interpretations, and analysis that do not introduce, perpetuate, or contribute to biased or unfair results. Providing multiculturally competent assessment services to individuals goes beyond the consideration of individual differences and includes a conceptual framework that limits inferences on the sole basis of group labels as to include the social and cultural world of the client and the intersectionality of the client's identities, as well as the influence of cultural identities of client and examiner in the therapeutic relationship or any evaluative setting. Psychologists strive to understand how culture can interact with every aspect of the assessment process and that they adapt their practices as needed.

### Application

Psychologists strive to gain knowledge of the cultural identities of their clients and the state-of-the-art literature regarding the best fit of the client with assessment procedures and instruments. Psychologists strive to consider how test administration, results, and interpretation may be influenced by individual diversity factors (e.g., age, disability, race, ethnicity, gender, religion/spirituality, sexual orientation and gender diversity, social class, language, acculturation, immigration/refugee status; APA, 2017b) and also to consider the intersection of cultural identities. Given the current abundance of scholarship related to race/ethnicity as well as the scholarship focused on other identity groups (e.g., age, disability, race, ethnicity, gender, religion/spirituality, sexual orientation and gender diversity, social class, language, immigration/refugee status; APA, 2017b), psychologists have strong scientific resources and frameworks to consider when conducting psychological assessments. Psychologists endeavor to learn about their own cultural identities and cultural attitudes and beliefs to minimize the negative influence of these issues in

their professional competence with diverse clients. Psychologists cultivate an understanding that achieving cultural competence is a lifelong process that involves, in addition to acquiring cultural knowledge, continuous cultural humility (Hook & Watkins, 2015), self-awareness and evaluation of the attitudes, values, interactions, and power dynamics with the clients. When the use of gender, race, and ethnicity is restricted in the testing, scoring, interpretation, or analysis by legal requirements in certain fields (e.g., employment), examiners endeavor to familiarize themselves with legal and regulatory requirements to use test information in a manner consistent with those regulatory standards.

GUIDELINE 10

## Psychologists who conduct psychological testing, assessment, and evaluation aspire to ensure awareness of individual differences, various forms of biases or potential biases, cultural attitudes, population appropriate norms, and potential misuse of data.

### Rationale

Research in many areas of specialty and proficiency has established that demographic and individual factors such as race, ethnicity, gender, age, geography, dominant language, socioeconomic status, and cultural variables of the examinee exert an important impact on the validity and reliability of results of psychological tests (Sandoval et al., 1998). These individual factors, to the extent that they make the examinee deviate from the test's normative standards, may modify the way in which personal characteristics will be manifested or are interpreted by the examiner and may provide a profile that is completely at odds with the reality of the examinee. Individuals from nondominant cultural, racial/ethnic, and other identity groups have been found to demonstrate unique presentation of symptoms, have different awareness and explanation of psychological distress, and are more likely to distrust providers and authority, factors that limit the generalizability of findings derived from most psychological instruments (Sue & Sue, 2013). Similarly, the test construct may function differently in different populations, with research accumulating to suggest that the scores of some instruments did not capture the intended psychopathology and unique cultural differences across groups (Janssen, 2011; Kim et al., 2009; Kim et al., 2011).

Cultural factors, attitudes, and demographic characteristics of the examinee also impact the applicability of the standard administration and related factors of the assessment process, such as the relationship between examiner and examinee (Butcher et al., 2016). For instance, test-taking demands that are irrelevant to the construct being measured and that impact an individual's ability to demonstrate their ability such as education, mastery of the test language, diverse values, expectations, acculturation, and psychosocial stresses derived from their minority status may compromise the test-taking performance and subsequently render the findings invalid (Diaz-Santos & Hough, 2016). Personal values, expectations, and attitudes of the examiner equally have been found to impact the judgment of examiners to the point that it may diminish the accuracy of their test selection, use, and interpretations (APA, 2003).

### Application

When choosing a test to administer, examiners endeavor to use the available information regarding the test psychometric properties, including measurement equivalence and standardization sample, to evaluate the fit with the examinee's individual and demographic characteristics. Examiners strive to become familiar with the examinee's culture and subcultural context to compare with the culture of the normative sample and the test development and decide if the test will afford a fair assessment. Examiners strive to avoid using a test in a manner or for a purpose not supported by evidence-base studies. When having to modify a test or administration procedure to fit the characteristics of an examinee, examiners aspire to consider the impact of those changes on the test results. Similarly, examiners aspire to understand the cultural idiosyncrasies in performance and response style, including cultural response bias that impacts standard validity, by consulting experts or studying research that documents the use of the test with the examinee's specific group. Examiners strive to obtain the training to competently assess individuals from diverse groups and learn about the range of testing procedures' applicability across diverse groups. Psychologists aspire to understand their own worldviews and sociocultural histories' impact on the attitudes, beliefs, and values they hold with individuals with characteristics different from their own and consider how these factors impact their approach to assessment, interpretation of psychological test data, use of labels, and clinical decisions.

GUIDELINE 11

## Psychologists who conduct psychological testing, assessment, and evaluation endeavor to recognize the nature of and relationship among individual, cohort, and group differences.

### Rationale

Assessment and interpretation of psychological tests involve an understanding of the similarities and differences within the examinee, the cohort to which the examinee belongs, and the group norm samples used in test construction. Some norm-referenced tests (i.e., using standard, scaled, percentile and T-scores) imply categories of score deviation based on comparisons of the participants' performance with normative data. Scores that are outside the broad range of average in a normal distribution may be classified as deviant from the mean. Due to the inherent limitations in the diversity and variability of a test normative sample, there is a potential that the characteristics of an individual examinee are poorly represented or are completely absent in the normative group. This means that individuals who are not appropriately represented in the normative sample (e.g., due to cultural, racial, age, and ethnic differences) have a greater chance of misinterpretation of their performance because of different individual and cultural characteristics rather than poor performance.

Even when the individual examinee can be placed within a cohort of the general normative group in characteristics such as race, ethnicity, gender, education, or geographic issues, it is inevitable that members from certain minority or vulnerable groups will still differ in some variables. For instance, it has been found that most test norms do not include information regarding the socioeconomic status of the population (Dana, 2001). Socioeconomic status has been found to be a variable that, in cases of high socioeconomic status, eliminates racial/ethnic differences in certain psychological instruments (Beiser & Gotowiec, 2000; Thakker et al., 1999). Similarly, many neuropsychological test norms were developed without sufficient numbers of older adults. Older adults are at high risk of being misclassified as having significant impairment when age, education, and race/ethnicity specific norms are not used (Miller et al., 2015; Schneider et al., 2015).

### Application

Psychologists strive to consider the examinee's individual and demographic characteristics when choosing a test to administer, vis-à-vis the test's normative sample to assess the test appropriateness to the examinee. Psychologists aspire to understand the test manual information regarding the characteristics of the sample, the procedures for examining between-groups differences in test performance, and use the meaning of this information when considering the scores of an examinee and the interpretation of tests results. Examiners strive to culturally contextualize tests results with examinees who do not closely match the characteristics of the normative group. Psychologists endeavor to consult or review scientific literature that helps them understand the potential limitations and risks of using certain normative categories on examinees from a minority or vulnerable group.

GUIDELINE 12

**Psychologists who conduct psychological testing, assessment, and evaluation seek to consider the unique issues that may arise when test instruments and assessment approaches designed for specific populations are used with diverse populations.**

### Rationale

Psychological assessment instruments and interpretive methods are culture specific because they are developed from a specific sociocultural context, most often Western. This cultural context is also explicitly or implicitly linked to the individual characteristics represented in the population used in the test standardization and may be less applicable to groups with other characteristics and cultures. Assessment techniques, data-generating procedures, and standardized instruments designed for a specific population have their validity and reliability tied to this normative group and cannot be assumed to apply to other groups unless they undergo adaptations and validity, reliability, and measurement equivalence have been tested across culturally diverse sample groups (APA, 2017b).

Psychologists endeavor to use only properly up-to-date normed, standardized, and translated measures chosen specifically based on the client's ethnic group, acculturation level, language proficiency, education level, socioeconomic status, age, disability status, and other relevant demographic factors. Recent advances in psychometrics have made significant progress examining the measurement equivalence and differential item functioning between individuals from groups different from the reference population for which the test was found to be reliable. There are many psychological tests with strong scientific basis for application with Western diverse groups with adequate levels of acculturation and English language dominance.

For many areas of assessment, including cognitive function, it is important for psychologists to strive to use age-appropriate test norms. For example, slowed processing speed is a normal part of cognitive aging. As a result, average scores for older adults on tests requiring fast performance, such as the Trail Making Test, would be interpreted as indicating impairment if norms for younger people were used. The growth in the older adult population means that many psychologists who do not have specialized training in assessment of older adults may encounter older individuals in a variety of settings. There are important issues such as the necessity for test norms specific to older adults including subgroups such as those older than age 85 and those with low educational attainment or low-quality educational experiences because of historical social problems such as segregation.

Psychologists are encouraged to recognize that a psychological test needs to have established research-based validity and reliability in the culture and language in which it was constructed before it can be considered appropriate for adaptation in other languages and cultures, which may be necessary for clients with less acculturation to the Western society and whose English is not the dominant language (Butcher et al., 2016). Tests can often be translated, adapted, and renormed in diverse cultural and demographic groups following strict guidelines for translation, construct analysis, and norming to ensure that alternative versions of the test are equivalent to the original.

When a test administration or protocol requires translation to match the language dominance of the examinee because a test version in the client's language and with appropriate cultural characteristics has not been developed, a number of strict guidelines have been recommended by the International Test Commission (2017). These translations can include trained professionals conducting forward and back translations focused on meaning including local and cultural context. Nevertheless, those translated tests or protocols, if not normed, may produce less meaningful results.

### Application

Although it is nearly impossible to match a test or assessment instrument to all the diverse characteristics of the client/patient/examinee, psychologists endeavor to familiarize themselves with the characteristics and cultural origin of their instruments in order to decide whether the test taker would be fairly assessed and fairly compared with the normative group. In

particular, psychologists strive to use test versions that better match the most prominent characteristics and culture of the examinee. One method to advance acceptable levels of matching is to include a psychological examination with a thorough psychosocial interview that provides the client information necessary to match to the best possible instruments.

Psychologists endeavor to assess the language dominance and capacity of nonmonolingual clients to determine the appropriateness of the intended assessment instrument. Psychologists strive to interpret and report the results of the assessment within a cultural context, and when in doubt, psychologists may consult with experts at any time of the assessment process in order to ensure cultural competence (Acevedo-Polakovich et al., 2007).

Psychologists are encouraged to familiarize themselves with the characteristics of translated versions when they need to use them, including the impact of language adaptations on the construct of the test. Examiners attempt to avoid choosing translations that have not been researched regarding whether they show results that have different meaning across groups (AERA et al., 2014). Moreover, psychologists strive to avoid using ad hoc translations. When assessing the applicability of an instrument that has not been adapted to the characteristics of the client (e.g., culture of origin, acculturation, language dominance, ethnicity, education, gender, age), psychologists endeavor to find literature or pilot studies that allow them to assess the validity of the test for the specific client/examinee including the differences in performance (APA, 2017b).

When the evaluation requires the use of nonresearched accommodations or modifications of the test administration, stimuli, or method to accommodate the client's needs or characteristics, the psychologist seeks to make the appropriate adjustment to the interpretations and report this caveat in the report of findings with a detailed description of the modified circumstances.

Psychologists strive to incorporate in the interpretation and conceptualization of test results the cultural factors that have been found to impact psychological assessments, including issues such as historical trauma effects, spirituality, traditional belief systems, collectivistic orientation, and acculturation.

# TRAINING AND SUPERVISORY QUALIFICATIONS AND ROLE

GUIDELINE 13

**Psychologists who educate and train others in testing, assessment, and evaluation strive to maintain their own competence in training and supervision and competency in assessment practice.**

### Rationale

Consistent with the APA Ethics Code, psychologists not only develop competence but make efforts to ensure they maintain competence (2.03 Maintaining Competence; APA, 2017a). Such efforts should optimally be deliberate given studies have found that the rapid increase in the amount of information available leads to a decreased ability to keep up to date (Neimeyer, Taylor, & Rozensky, 2012). Training and supervision are considered core competencies in health service psychology that require deliberate training (Falender et al., 2004). Further, the APA Standards of Accreditation for Health Service Psychology require training in assessment as a profession-wide competency. Licensing boards may view training and supervision as the practice of psychology, thereby introducing the standards of practice in assessment as expectations of competence. Those who teach assessment strive to be knowledgeable of the standards of practice, test instruments and their applicability, ongoing revisions of assessment and evaluation measures, and new methods of assessment and evaluation that are applicable to the populations for whom psychologists or their students are providing services. As articulated in the APA Guidelines for Clinical Supervision in Health Service Psychology, psychologists who supervise are encouraged to include supervision in their efforts to maintain competence (APA, 2015a).

### Application

A primary mechanism through which psychologists can maintain their competence is through participation in CE or continuing professional development (Neimeyer et al., 2014). At their most effective, such programs include both a didactic component and an interactive component (Neimeyer, Taylor, & Cox, 2012). Psychologists who educate and train students strive to remain aware of and to meet the current profession-wide competency expectations in assessment that include specific knowledge, skills, and application experiences beyond general programmatic requirements. These may include coursework in psychometrics, cognition and intelligence, administration and interpretation of performance-based and self-report measures of personality, integration of data, and reporting of results and the application of findings to recommendations and treatment. Psychologists who educate and train students strive to be aware of the developmental competency expectations of students at the practicum, internship, and readiness for practice levels (APA, 2012). Psychologists who teach assessment and evaluation but do not provide the experiential practicum or clinical experience are encouraged to ensure that the external supervisors meet the professional practice and knowledge-based competencies that they are expected to supervise.

Psychologists who train and supervise students, employees, and others consider engaging in CE that specifically focuses on advancements in the teaching, supervision, and practice of testing, assessment, and evaluation. Consultation and supervision of supervision are often effective mechanisms to maintain one's competence. Psychologists

are mindful of the recommendation in the Guidelines for Clinical Supervision in Health Service Psychology (APA, 2015a) that,

> Education and training in supervision should include the following: models and theories of supervision; modalities; relationship formation, maintenance, rupture and repair; diversity and multiculturalism; feedback, evaluation; management of supervisee's emotional reactivity and interpersonal behavior; reflective practice; application of ethical and legal standards; decision making regarding gatekeeping; and considerations of developmental level of the trainee. (p. 36)

GUIDELINE 14

**Psychologists who supervise employees or individuals who lack training in testing, assessment, and evaluation strive to ensure that supervision ultimately provides examinees/clients with testing, assessment, and evaluation that meets the ethical and professional standard of care and scope of practice.**

### Rationale

Several sections of the APA Ethics Code speak to the needs to ensure that psychologists only delegate work to others who they know have the requisite competencies to perform such work (2.05 Delegation of Work to Others; APA, 2017a). Further, psychologists specifically do not allow testing, assessment, or evaluation to be done by unqualified persons, except when conducted for training purposes and with appropriate supervision (9.07 Assessment by Unqualified Persons; APA, 2017a).

### Application

Psychologists who supervise individuals who will engage in testing, assessment, or evaluation are first encouraged to gain understanding of the relevant legal and regulatory documents that dictate allowable scope of practice for the supervisee's profession. Next, a baseline understanding of the supervisee's level of competence can be obtained. Self-report from the supervisee is typically considered a necessary, but not sufficient, method to assess competence. The supervisor strives to observe the individual conducting the testing, assessment, or evaluation with someone other than a patient/client to make decisions about level of competence and need for training. Although training might include didactic presentations or readings, ideally the individual who is supervised could watch the supervisor engage in the activity they will be conducting and then the supervisor could watch the individual. Further, psychologists strive to ensure that assessment of continued competence is conducted at regular intervals to ensure fidelity to the requirements for proper administration.

## TECHNOLOGY

GUIDELINE 15

**Psychologists who use technology when testing, assessing, or evaluating psychological status strive to remain aware of technological advances; of the influence of technology on assessment; and of standard practice, laws, and regulations in telepsychology.**

### Rationale

In the past 50 years, advances in technology have greatly impacted the field of psychological assessment (Butcher, 2006). Originally, use of computers in psychological assessment practice was primarily limited to automated scoring of paper-and-pencil tests, reporting of scores on these tests, and occasionally presentation of simple interpretative hypotheses based on these scores. However, in more recent times, this use has expanded to include internet-based administration platforms facilitating access to multiple test instruments, some of which are electronic presentations of traditional (legacy) psychometric instruments originally designed for paper-and-pencil administration. Other assessment instruments are specifically designed for electronic presentation, taking advantage of the unique presentation, response, reporting, and data-gathering capabilities of this medium (Butcher, 2006; Butcher et al., 2009; Way & Robin, 2016). Finally, and most recently, advances in computer technology, big data analysis, and gaming design have facilitated the emergence of completely new paradigms of assessment using interactive video for traditional interviews and/or real-time role-play simulations, virtual reality exercises, big data analysis, and other specialized diagnostic techniques (Wahlstrom, 2017). As with all other tests and assessments, the essential criteria for evaluating technology enhanced measures are reliability, validity, and fairness.

### Application

Inasmuch as computer technology, test instrument usage, and new instrument design are constantly evolving, the responsibilities and challenges to the psychologist practitioner using these modalities are likewise substantial. Insofar as many or most (legacy) tests are now electronically mediated in one way or another in regard to scoring, administration, and/or interpretation (Wahlstrom, 2017), when using legacy tests adapted for electronic presentation, psychologists are encouraged to review with care available information and validation evidence documenting the process through which these instruments have been adapted, including issues of equivalence in regard to internal consistency, predictive

validity, and/or interpretation across modalities (Butcher et al., 2009).

It should also be noted that use of assessment instruments incorporating elements of interactive real-time video is also likely to fall within the purview of telepsychology and, in this regard, psychologists are strongly encouraged to be familiar with the *Guidelines for the Practice of Telepsychology* (APA, 2013), as many of the issues described in those guidelines are likely to be directly relevant to use of these instruments.

Insofar as some concern has been expressed regarding the use of adapted legacy tests (i.e., those developed and standardized prior to electronic enhancements in administration, scoring, and interpretation), it is important to ascertain the degree to which assumptions are based on the results of electronically presented tests. Psychologists strive to evaluate carefully the publisher's technical manual data in regard to equivalence or norms, effects of electronic presentation, and/or scoring, as well as validation of interpretative predictions made on the basis of electronically administered and scored version of the test in question (Butcher et al., 2009). Individuals differ substantially in their technology experience and proficiency. These differences have important implications for the administration and interpretation of technology-based measures. For example, older adults have less technology experience and proficiency compared with younger adults (Czaja et al., 2019). The reliability, validity, and acceptability of technology-based measures must be examined in diverse populations, including in older adults.

Psychologists endeavor to recognize that it is also important that instruments primarily designed to take advantage of opportunities and advances in computer technology have been carefully researched regarding fundamental scientific utility for measurement and interpretation of results. Similarly, when evaluating new instruments and assessment procedures designed exclusively for electronic presentation, psychologists strive to review carefully the test manual and/or detailed documentation of the underlying technical information describing the rationale, construction, response characteristics, internal consistency, and validation evidence supporting use of the instrument and fairness (AERA et al., 2014; Butcher et al., 2009; Wahlstrom, 2017). Psychologists are also encouraged to consider potential pitfalls of computer-based assessment and computer-generated results (e.g., potential disruption of the diagnostic alliance; see Rosen et al., 2016).

GUIDELINE 16

**Psychologists who conduct services using technology for online or in-person testing, assessment, and evaluation make every effort to ensure their own competency.**

**Rationale**

The past 20 years have seen explosive growth and development in various areas of information technology (IT). Developments in areas of new hardware, software, networking, and data storage have become ubiquitous in everyday life and, as such, have directly impacted the practice of psychology in general and psychological assessment in particular (Montag et al., 2016). Concurrent with these developments, the prevalence and sophistication of security threats and breaches of data integrity have also greatly accelerated. Similarly, as the number of psychologists and frequency of technological utilization within psychological testing and assessment escalates, the need to achieve and maintain technological competence is critical to maintaining professional competence in general as well as standards of ethical practice (Lustgarten, 2015; Rigg, 2018). Considering the preceding, psychologists are strongly encouraged to develop and/or enhance their competence in technology by seeking out available technical and training resources and/or CE opportunities in computer science and technology.

**Application**

The APA and APA Services, Inc. provide information about building and maintaining a technology infrastructure for psychology practices and a list of available resources for psychologists with regard to software reviews, privacy and security, automation, and basic technology. Psychologists who provide services using technology are strongly encouraged to review/access this resource periodically to maintain requisite knowledge and competence in critical technology-related areas.

In addition, certain organizations are dedicated to the development and presentation of training programs in computer technology focusing on education and psychology. For example, the Association for the Advancement of Computing in Education maintains a searchable database of more than 100,000 articles and abstracts in computer technology.

Last, psychologists whose assessment practices are deeply rooted in technology and/or are in positions of primary responsibility for developing and maintaining technology-based assessment programs and products within their organizations may wish to consider advanced training leading to formal certification in various information-technology-related functions.

GUIDELINE 17

**Psychologists who use technology-based assessment instruments are encouraged to take reasonable steps to ensure the security, transmission, storage, and disposal of data. Psychologists also strive to ensure that security measures are in place to protect data and information related to their clients/patients/examinees from unintended access, misuse, or disclosure.**

**Rationale**

Psychologists using testing and assessment instruments and processes that are technologically involved for scoring, administration, or interpretation and/or who gather and store assessment responses and other client information are strongly advised to periodically review their current practices in regard to data protection, storage, and protection of client confidentiality.

The 21st century has brought with it significant increases in technology and advances in accessibility (Lustgarten, 2015). Psychological tests, and other assessment instruments and processes that utilize this

technology, do so within the context of a rapidly growing and changing threat environment as well. As such, in addition to maintaining general competence in technological areas impacting practice, it is particularly important for psychologists in general and assessment practitioners in particular to strive to be aware of threats to data integrity and client confidentiality.

In response to this growing threat environment, in recent years a number of federal and state rules and regulations have been enacted that directly impact the storage of data and protection of confidentiality. Compliance with federal and state privacy laws and regulations can also be expected to intersect with areas of ethical practice as well as with APA (2007) record-keeping guidelines. As such, psychologists are encouraged to maintain knowledge and competence in these areas as well.

Manipulation through human interaction and social engineering is a common source of initiation of data breaches (Tetri & Vuorinen, 2013). As such, psychologists strive to ensure ongoing training and regular review of all practice employees and/or others with access to confidential information as an essential component of any practice security plan.

Finally, a number of institutional guidelines, best practices and standards for protecting security, data integrity, and confidentiality have been developed by APA and various technological organizations. Psychologists using technologically involved tests and other assessment processes are strongly advised to integrate these practices and standards within their routines and practices.

**Application**

In addition to the APA Ethics Code (APA, 2017a) and the *Record Keeping Guidelines* (APA, 2007), and the laws and regulations on privacy and confidentiality, Lustgarten (2015) suggested a number of best practices for protecting both data and client confidentiality, including the development of a threat model, encryption, use of HIPAA-compliant cloud providers, use of two-factor authentication, use of "air gapped" (stand-alone, non-networked, or internet-connected) computers, and review of informed consent procedures.

Ferreira and Teles (2019) studied and reviewed the way email phishing and other social engineering exploits are used to persuade employees and others to bypass established security protocols. A number of commercially available education and training programs are also for use by employees and others with access to confidential information. These programs teach participants to recognize phishing emails and other common social engineering exploits, and they periodically include unannounced audits and exercises using white flag (unannounced program generated) exploits.

In addition, certain national and international technology organizations have established listings of best practices and general standards. Psychologists are encouraged to be aware of the existence of:

- The International Organization for Standardization (ISO)
  - » ISO is an independent nongovernmental international organization with a membership of 164 national standards bodies.
  - » ISO/IEC 27001 is the best-known standard in the ISO family of standards providing requirements for an information security management system.
  - » ISO 15408 ISO/IEC 15408-1:2009 establishes the general concepts and principles of IT security evaluation and specifies the general model of evaluation given by various parts of ISO/IEC 15408, which in its entirety is meant to be used as the basis for evaluation of security properties of IT products.
- Standard of Good Practice
  - » The Standard of Good Practice for Information Security, published by the Information Security Forum, is a business-focused, practical, and comprehensive guide to identifying and managing information security risks in organizations and their supply chains.
- National Institute of Standards of Technology, under the U.S. Department of Commerce (NIST).
  - » NIST implements practical cybersecurity and privacy through outreach and effective application of standards and best practices necessary for the United States to adopt cybersecurity capabilities.

# REFERENCES


Acevedo-Polakovich, I. D., Reynaga-Abiko, G., Garriott, P. O., Derefinko, K. J., Wimsatt, M. K., Gudonis, L. C., & Brown, T. L. (2007). Beyond instrument selection: Cultural considerations in the psychological assessment of U.S. Latinas/os. *Professional Psychology: Research and Practice, 38*(4), 375–384. http://dx.doi.org/10.1037/0735-7028.38.4.375

Achenbach, T. M. (2017). Future directions for clinical research, services, and training: Evidence-based assessment across informants, cultures, and dimensional hierarchies. *Journal of Clinical Child and Adolescent Psychology, 46*, 159–169.

Alfonso, V. C. (2004, July). How to be an educated consumer of assessment instruments: Brief

tips for practitioners. *The Score, 26*, 10–11.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

American Psychological Association. (2001). Guidelines for test user qualifications. *American Psychologist, 56*(12), 1099–1113. https://doi.org/10.1037//0003-066X.56.12.1099

American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. *American Psychologist, 58*(5), 377–402. https://doi.org/10.1037/0003-066X.58.5.377

American Psychological Association. (2007). Record keeping guidelines. *American Psychologist, 62*(9), 993–1004. https://doi.org/10.1037/0003-066X.62.9.993

American Psychological Association. (2012, July). Benchmarks evaluation system [Webpage]. https://www.apa.org/ed/graduate/benchmarks-evaluation-system.aspx

American Psychological Association. (2013). *Guidelines for the practice of telepsychology*. https://www.apa.org/practice/guidelines/telepsychology

American Psychological Association. (2015a). Guidelines for clinical supervision in health service psychology. *American Psychologist, 70*(1), 33–46. http://dx.doi.org/10.1037/a0038112

American Psychological Association. (2015b). Professional practice guidelines: Guidance for developers and users. *American Psychologist, 70*(9), 823–831. http://dx.doi.org/10.1037/a0039644

American Psychological Association. (2017a). *Ethical principles of psychologists and code of conduct* (2002, amended effective June 1, 2010, and January 1, 2017). http://www.apa.org/ethics/code/index.aspx

American Psychological Association. (2017b). *Multicultural guidelines: An ecological approach to context, identity, and intersectionality*. http://www.apa.org/about/policy/multicultural-guidelines.pdf

American Psychological Association, Commission on Accreditation. (2015). *Standards of accreditation for health service psychology*. http://www.apa.org/ed/accreditation/about/policies/standards-of-accreditation.pdf

Americans With Disabilities Act of 1990, 42 U.S.C. § 12101 et seq. (1990).

Association of State and Provincial Psychology Boards. (2014). *Report of the competency assessment task force.* https://cdn.ymaws.com/www.asppb.net/resource/resmgr/eppp_2/2017_asppb_competencies_exp.pdf

Bandalos, D. L. (2018). *Measurement theory and application for the social sciences*. Guilford Press.

Beiser, M., & Gotowiec, A. (2000). Accounting for native/non-native differences in IQ scores. *Psychology in the Schools, 37*(3), 237–252. https://doi.org/10.1002/(SICI)1520-6807(200005)37:3<237::AID-PITS4>3.0.CO;2-N

Bracken, B. A., & Theodore, L. A. (in press). Clinical observation of preschool assessment behavior. In V. C. Alfonso, B. A. Braken, & R. J. Nagle (Eds.), *Psychoeducational assessment of preschool children* (5th ed.). CRC Press.

Bram, A. D., & Peebles, M. (2014). *Psychological testing that matters: Creating a road map for effective treatment*. American Psychological Association.

Brennan, R. L. (2001). *Generalizability theory*. Springer-Verlag.

Burns, G. L., & Haynes, S. N. (2006). Clinical psychology: Construct validation with multiple sources of information and multiple settings. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 401–418). American Psychological Association.

Butcher, J. N. (2006). Assessment in clinical psychology: A perspective on the past, present challenges, and future prospects. *Clinical Psychology: Science and Practice, 13*(3), 205–209. http://dx.doi.org/10.1111/j.1468-2850.2006.00025.x

Butcher, J. N., Hass, G. A., & Paulson, J. A. (2016). Clinical assessment in international settings. In F. T. L. Leong, D. Bartram, F. Cheung, K. F. Geisinger, & D. Iliescu (Eds.), *The ITC international handbook of testing and assessment* (pp. 217-230). Oxford University Press.

Butcher, J. N., Perry J. N., & Dean, B. L. (2009). Computer-based assessment. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (pp. 163-182). Oxford University Press.

Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241 (1964).

Civil Rights Act of 1991, Pub. L. No. 102-166, 105 Stat. 1071 (1991).

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Sage.

Czaja, S. J., Boot, W. R., Charness, N., & Rogers, W. A. (2019). *Designing for older adults: Principles and*

*creative human factors approaches*. CRC Press.

Dana, R. H. (2001). Clinical diagnosis of multicultural populations in the United States. In L. A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 101–131). Jossey-Bass.

Demakis, G. J. (Ed.) (2012). *Civil capacities in clinical neuropsychology: Research findings and practical applications*. Oxford University Press.

Diaz-Santos, M., & Hough, S. (2016). Cultural competence guidelines for neuropsychology trainees and professionals: Working with ethnically diverse individuals. In F. R. Ferraro (Ed.), *Minority and cross-cultural aspects of neuropsychological assessment: Enduring and emerging trends* (2nd ed., pp. 11–33). Taylor & Francis.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Erlbaum.

Equal Employment Opportunity Commission. (1978). *Code of federal regulations: Part 1607-uniform guidelines on employee selection procedures.* https://www.govinfo.gov/content/pkg/CFR-2017-title29-vol4/xml/CFR-2017-title29-vol4-part1607.xml

Falender, C. A., Erickson Cornish, J. A., Goodyear, R., Hatcher, R., Kaslow, N. J., Leventhal, G., Shafranske, E., & Sigmon, S. T. (2004). Defining competencies in psychology supervision: A consensus statement. *Journal of Clinical Psychology, 60*(7), 771–785. https://doi.org/10.1002/jclp.20013

Ferreira, A., & Teles, S. (2019). Persuasion: How phishing emails can influence users and bypass security measures. *International Journal of Human-Computer Studies, 125*, 19–31. https://doi.org/10.1016/j.ijhcs.2018.12.004

Finn, S. E. (2009). Incorporating base-rate information in daily clinical decision making. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (pp. 140–149). Oxford University Press.

First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2016). *Structured Clinical Interview for*

*DSM-5 Disorders, Clinician Version (SCID-5-CV)*. American Psychiatric Association.

Graham, J. R., Naglieri, J. A., & Weiner, I. B. (Eds.). (2013). *Handbook of psychology: Assessment psychology* (pp. 82–113). Wiley & Sons.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65–110). Praeger.

Hessen, E., Hokkanen, L., Ponsford, J., van Zandvoort, M., Watts, A., Evans, J., & Haaland, K. Y. (2018). Core competencies in clinical neuropsychology training across the world. *Clinical Neuropsychologist, 32*, 642–656.

Hook, J. N., & Watkins, C. E., Jr. (2015). Cultural humility: The cornerstone of positive contact with culturally different individuals and groups? *American Psychologist, 70*(7), 661–662. http://dx.doi.org/10.1037/a0038965

Hopwood, C. J., & Bornstein, R. F. (Eds.). (2014). *Multimethod clinical assessment*. Guilford.

International Test Commission. (2001). ITC guidelines on test use. *International Journal of Testing, 1*(2), 93–114.

International Test Commission. (2017). ITC guidelines for translating and adapting tests (2nd ed.). http://www.InTestCom.org

Janssen, R. (2011). Using a differential item functioning approach to investigate measurement invariance. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-cultural analysis: Methods and applications* (pp. 415–432). Routledge.

Kane, M. (1992). An argument-based approach to validation. *Psychological Bulletin, 112*(3), 527-535. http://dx.doi.org/10.1037/0033-2909.112.3.527

Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Education Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Kaplan, R. M., & Saccuzzo, D. P. (2017). *Psychological testing: Principles, applications, & issues.* Cengage Learning.

Kaslow, N. J., Grus, C. L., Campbell, L. F., Fouad, N. A., Hatcher, R. L., & Rodolfa, E. R. (2009). Competency assessment toolkit for professional psychology. *Training and Education in Professional Psychology, 3*(4, Suppl.), S27–S45. http://dx.doi.org/10.1037/a0015833

Keller, D. K. (2006). *The tao of statistics: A path to understanding (with no math).* Sage.

Kim, G., Chiriboga, D. A., & Jang, Y. (2009). Cultural equivalence in depressive symptoms in older white, black, and Mexican-American adults. *Journal of the American Geriatrics Society, 57*(5), 790–796. https://doi.org/10.1111/j.1532-5415.2009.02188.x

Kim, G., DeCoster, J., Huang, C.-H., & Chiriboga, D. A. (2011). Race/ethnicity and the factor structure of the Center for Epidemiologic Studies Depression Scale: A meta-analysis. *Cultural Diversity and Ethnic Minority Psychology, 17*(4), 381–396. http://dx.doi.org/10.1037/a0025434

Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). Routledge.

Lustgarten, S. D. (2015). Emerging ethical threats to client privacy in cloud communication and data storage. *Professional Psychology: Research and Practice*, *46*(3), 154–160. http://dx.doi.org/10.1037/pro0000018

Miller, I. N., Himali, J. J., Beiser, A. S., Murabito, J. M., Seshadri, S., Wolf, P. A., & Au, R. (2015). Normative

data for the cognitively intact oldest-old: The Framingham Heart Study. *Experimental Aging Research, 41*(4), 386–409.

Montag, C., Duke, E., & Markowetz, A. (2016). Toward psychoinformatics: Computer science meets psychology. *Computational and Mathematical Methods in Medicine*, 2016, Article 2983685. http://dx.doi.org/10.1155/2016/2983685

Neimeyer, G. J., Taylor, J. M. & Cox, D. R. (2012). On hope and possibility: Does continuing professional development contribute to ongoing competence? *Professional Psychology: Research and Practice, 43*(5), 476–486. https://doi.org/10.1037/a0029613

Neimeyer, G. J., Taylor, J. M., & Rozensky, R. H. (2012). The diminishing durability of knowledge in professional psychology: A Delphi Poll of specialties and proficiencies. *Professional Psychology: Research and Practice, 43*(4), 364–371. https://doi.org/10.1037/a0028698

Neimeyer, G. J., Taylor, J. M., Rozensky, R. H., & Cox, D. R. (2014). The diminishing durability of knowledge in professional psychology: A second look at specializations. *Professional Psychology: Research and Practice, 45*(2), 92–98. https://doi.org/10.1037/a0036176

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata

measure and situational judgment inventory as predictors of college student performance. Journal of Applied Psychology, 89(2), 187–207. http://dx.doi.org/10.1037/0021-9010.89.2.187

Reynolds, C. R., & Suzuki, L. A. (2013). Bias in psychological assessment: An empirical review and recommendations. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), *Handbook of psychology: Assessment psychology* (pp. 82–113). Wiley & Sons.

Rigg, T. (2018). The ethical considerations of storing client information online. *Professional Psychology: Research and Practice*, *49*(5-6), 332–335. http://dx.doi.org/10.1037/pro0000217

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*(4), 355–361. http://dx.doi.org/10.1037/0735-7028.36.4.355

Roper, B. L., Block, C. K., Osborn, K., & Ready, R. E. (2018). Education and training for clinical neuropsychologists in integrated care settings. *Archives of Clinical Neuropsychology, 33*(3), 263–268.

Rosen, D. C., Nakash, O., & Alegria, M. (2016). The impact of computer use on therapeutic alliance and continuance in care during the mental health intake. *Psychotherapy*, *53*, 117–123. http://doi.org/10.1037/pst0000022

Sandoval, J. H., Frisby, C. L., Geisinger, K. F., Ramos-Grenier, J., & Scheuneman, J. D. (Eds.). (1998). *Test interpretation and diversity: Achieving equity in assessment.* American Psychological Association.

Sattler, J. M. (2014). *Foundations of behavioral, social, and clinical assessment of children* (5th ed.). Author.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262–274. https://doi.org/10.1037/0033-2909.124.2.262

Schneider, A. L., Sharrett, A. R., Gottesman, R. F., Coresh, J., Coker, L., Wruck, L., Selnes, O. A., Deal, J., Knopman, D., & Mosley, T. H. (2015). Normative data for eight neuropsychological tests in older blacks and whites from the atherosclerosis risk in communities (ARIC) study. *Alzheimer Disease and Associated Disorders, 29*(1), 32–44.

Smith, G. (Clinical Neuropsychology Synarchy). (2019). Education and training in clinical neuropsychology: Recent developments and documents from the clinical neuropsychology synarchy. *Archives of Clinical Neuropsychology, 34*, 418–431.

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Author.

Stage, S. A., Jackson, H. G., Moscovitz, K., Erickson, M. J., Thurman, S. O., Jessee, W., & Olson, E. M. (2006). Using multimethod-multisource functional behavioral assessment for students with behavior disabilities. *School Psychology Review, 35,* 451–471.

Sue, D. W., & Sue, D. (2013). *Counseling the culturally diverse: Theory & practice* (6th ed.). Wiley & Sons.

Tetri, P., & Vuorinen, J. (2013). Dissecting social engineering. *Behaviour & Information Technology*, *32*(10), 1014–1023. https://doi.org/10.1080/0144929X.2013.763860

Thakker, J., Ward, T., & Strongman, K. T. (1999). Mental disorder and cross-cultural psychology: A constructivist perspective. *Clinical Psychology Review, 19*(7), 843–874. https://doi.org/10.1016/S0272-7358(98)00077-4

Wahlstrom, D. (2017). Technology and computerized assessments: Current state and future directions. In S. S. Bush, G. J. Demakis, & M. L. Rohling (Eds.), *APA handbook of forensic neuropsychology* (pp. 463–476). American Psychological Association.

Way, W., & Robin, F. (2016). The history of computer-based testing. In C. S. Wells & M. Faulkner-Bond (Eds)., *Educational measurement: From foundations to future* (pp. 185–208). Guilford Press.

Weiner, I. B., & Otto, R. K. (Eds.). (2014). *Handbook of forensic psychology* (4th ed). Wiley.

Wright, P. W. D., & Wright, P. D. (2015). *Special education law* (2nd ed.). House Harbor Law Press.

Yalof, J. (2019). When the assessor's limits are tested: Enactments and the assessment frame in psychological testing. *Journal of Personality Assessment.* Advance online publication. http://doi.org/10.1080/00223891.2019.1613241

AMERICAN
PSYCHOLOGICAL
ASSOCIATION