6-2020

Exploration of Factors that Influence the use of Psychological Testing in the Treatment and Evaluation of Children: Development of a Clinical Decision-Making Model

Joy Bui
Florida School of Professional Psychology at National-Louis University

Exploration of Factors that Influence the use of Psychological Testing in the Treatment and Evaluation of Children: Development of a Clinical Decision-Making Model

Joy T. Bui

Florida School of Professional Psychology at National Louis University

Patricia S. Dixon, Psy.D.
Chair

Eric L. Rosen, Ph.D.
Member

A Clinical Research Project submitted to the Faculty of the Florida School of Professional Psychology at National Louis University in partial fulfillment of the requirements for the degree of Doctor of Psychology in Clinical Psychology.

Tampa, Florida
April, 2020

The Doctorate Program in Clinical Psychology
Florida School of Professional Psychology
at National Louis University

CERTIFICATE OF APPROVAL

Clinical Research Project

This is to certify that the Clinical Research Project of

Joy T. Bui

has been approved by the
CRP Committee on April 17, 2020
as satisfactory for the CRP requirement
for the Doctorate of Psychology degree
with a major in Clinical Psychology

Examining Committee:

Patricia S. Dixon, Psy.D

Committee Chair: Patricia S. Dixon, Psy.D.

Eric L. Rosen, Ph.D., C.Ht., FPPR

Member: Eric L. Rosen, Ph.D.

Abstract

Mental health disorders are on the rise among children. Within the field of clinical psychology, psychological testing is considered a professional competency and a formative diagnostic tool. Simultaneously, testing can help clarify differential diagnosis, guide treatment for both medical and mental health conditions, and provide accurate monitoring of treatment over time. Although there is substantial evidence supporting the use of testing in practice, there is limited scientific data linking these same benefits to the assessment process. One explanation for this is that studying psychological assessments is a challenging task, given the complexities of the process. Subsequently, many have described conducting psychological assessment as a continual, dynamic, decision-making process. However, what attributes to a psychologist's clinical judgment when choosing to conduct testing remains unclear. The purpose of this clinical research project was to explore the contributing factors behind clinical child psychologists' decision-making in relation to determining the necessity of conducting assessments with children. An analysis of the existing literature was conducted for psychological assessment methods with children, clinical decision-making in psychology, and healthcare related fields. The findings of this research project suggested several patterns across the literature. Specifically, the emerging factors suggested that clinical psychologists' decision-making closely resembles a Naturalistic Decision Making model, in turn, a clinical decision-making model for clinical child psychologists was proposed. Finally, after analyzing factors weighing into effective clinical decision-making, a proposed model of action steps outlining psychological assessment methods with children was created to inform clinical practice.

DEDICATION

For my beloved parents, Rachel and Bobby. For my little sister with the kindest heart, Jayde. For my best friend and soulmate, Marlon Daniel.

Rejoice Always. Pray without ceasing. In all circumstances give thanks.

1 Thessalonians 5:16-18

ACKNOWLEDGEMENTS

To my parents, my sister, and the love of my life, this would not have been possible without you. There are no words that will ever truly capture or express how blessed I am for each of you. Your unyielding support, love, faith, and encouragement have been the source of my inspiration and strength throughout years. Thank you again for your countless sacrifices during my journey and especially this project. Thank you for your unconditional love, patience, and prayers. Thank you for inspiring me to pursue and complete this doctorate!

I would also like to give special recognition and gratitude to committee members, Dr. Eric Rosen, and Dr. Patricia S. Dixon. Your continuous support, guidance, and feedback with this project helped make my dream of becoming a psychologist a reality!

To my friends, colleagues, mentors, and supervisors - you are deeply treasured. Thank you all so very much for playing a valuable role in every step of my journey. I appreciate your support, wisdom, time, dedication, encouragement, and compassion. You have fostered my growth personally, professionally, and spiritually.

As always, I give thanks to God, for all of your blessings.

Philippians 4:13

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER I: FOUNDATIONS OF TESTING AND DECISION-MAKING

The use of psychological testing and assessment with children has steadily increased over the past century (Cuthbert & Insel, 2013; Domino & Domino, 2006; Eignor, 2013; Kamphaus, Petoskey, & Rowe, 2000; Meyer et al., 2001; Miller & Lovler, 2018; Shilkofski & Hunt, 2015). At present, the American Psychological Association (APA, 2013d) "acknowledges the validity of psychological tests are similar to medical tests" (para. 2). APA (2013d) further compares a health care provider ordering blood tests to better understand a patient's physical symptoms to a psychologist's evaluation providing information to help develop a treatment plan. Meyer et al. (2001) emphasized the value of psychological testing as the process provides information about an individual's abilities (e.g., cognitive skills, behaviors, educational) as well as their functionality and therapeutic requirements. Further, Meyer et al. (2001), as well as many others (e.g., APA, 2013a; Bornstein, 2017; Russo, 2015; Sattler & Hoge, 2006), acknowledge psychological assessments' significance in the field in that their results have the capacity to alter the course of an individual's life.

Despite the well-researched validity of instruments, the use, application, and status of psychological testing within clinical psychology has generated controversy. Specifically, there are some that question the utility, appropriateness, overuse, and efficacy of psychological assessment (Anastasi, 1985; Gottfredson & Saklofske, 2009; Sanders & Katz, 2013) while others value the practice of psychological testing, suggesting a need to increase utilization for the past several decades (Harvey, 2006; Heffer, Barry, & Garland, 2009; Hurt & Tomoyasu, 1995; Pelco, Ward, Coleman, & Young, 2009; Sattler, 1988; Sattler & Hoge, 2006; Swerdlik & Cohen, 2013).

Despite contrasting views, following a referral, ultimately a clinician must decide whether a child will undergo psychological testing or not. When discussing the act of decision-

making across disciplines, formal decision-making models are heavily researched and viewed as vital approaches to problem-solving (Gigerenzer & Gaissmaier, 2011). There are several decision-making frameworks that exist across fields, including businesses, health care systems, and legal institutions (Gigerenzer & Gaissmaier, 2011; Hammond, Keeney, & Raiffa, 2006; Klein, 2015; Perversi et al., 2018). Within healthcare, some commonly used terms are *clinical decision-making*, *shared decision-making*, and *clinical judgment* (Croskerry & Norman, 2008; Garb, 1997; Kicklighter, Geisler, Barnum, Heinerichs, & Martin, 2018; Muntean, 2011; National Academies of Sciences, Engineering, and Medicine; Institute of Medicine; Board on Health Care Services; & Committee on Diagnostic Error in Health Care, 2015; Splinter, 2010). Despite the nomenclature, this phenomenon can be formidable to define and is characteristically described more conceptually given that clinical judgement takes into account abstract thought processes that are incorporated by the provider in practice (Hammond et al., 2006; Muntean, 2011). Moreover, the APA's Ethical Principles of Psychologists and Code of Conduct (APA, 2017a) states, "Bases for Scientific and Professional Judgments: Psychologists' work is based upon established scientific and professional knowledge of the discipline" (Section 2.04). Therefore, clinical judgment can be understood as the act of applying reasoning with training, scientific theory, practical experiences, patient reports, and a level of uncertainty (APA, 2017a; Hammond et al., 2006; Kazdin, 2008; Knauss, 2001; National Academies of Sciences, Engineering, & Medicine et al., 2015; Smith, Higgs, & Ellis, 2008; Wills & Holmes-Rovner, 2006).

As expected, with advanced training and practical experience, providers or experts tend to enhance their decision-making skills and begin to easily synthesize each patient's complex variables into effective decisions that inform treatment plans (Charles, Gafni, & Whelan, 1999; Gigerenzer & Gaissmaier, 2011; Podgorelec, Kokol, & Rozman, 2002). However, even at the

expert level, research also suggests professionals are infrequently aware of their heuristics used to make clinical judgments as with more experience gained, their decision-making process changes and becomes almost autonomous (Charles et al., 1999; Croskerry, 2009; Croskerry & Norman, 2008; Gigerenzer & Gaissmaier, 2011). Given the complexity and uniqueness of each case a clinician encounters, decision-making models have been used as a guideline to aid in limiting individuals from biases and ethical pitfalls by offering a standard of considerations across patients (Splinter, 2010). In the age of evidence-based practices, including Evidenced Based Assessments (EBA), there is a need for more best-practice guidelines. The APA (2006) Presidential Task Force on Evidence-Based Practice stated, "One approach to implementing evidence-based practice in health care systems has been through the development of guidelines for best practice" (p. 271).

Within the field of clinical psychology, psychological testing is considered a professional competency, a formative diagnostic tool, and a compliment to psychological services. However, it is challenging to precisely quantify the annual number of psychological test batteries administered by psychologists (Hunsley & Mash, 2007; Piotrowski, 1999; Ready & Veague, 2014). To aid in understanding the vast applications of testing, data condensed from the Department of Education by the National Center for Education Statistics reported an increase from 6.4 million to 7.0 million students receiving services under the Individuals with Disabilities Education Act (IDEA) in the 2017-18 academic year (U.S. Department of Education, 2018). Estimates from other literature range from 100 to 200 million child evaluations conducted annually, which includes cognitive, socioemotional, and several other measures (Kamphaus et al., 2000; Pikulski, 1990; Sattler & Hoge, 2006; U.S. Department of Education, 2018).

Despite the incremental growing number of psychological assessments with children and the millions of children impacted, decision-making models for testing and assessment with children are not well researched (Garb, 2005; Kolen & Hendrickson, 2013). To better understand the value, role, impact, and clinical utility of examining emerging factors within decision-making prior to testing, it is helpful to review the foundation of psychological testing and the advancements that have occurred over the past century.

History of Psychological Evaluation and Testing

Psychological testing has become interwoven into the fabric of the discipline. Since the 1940s, psychological testing has been recognized as a "diagnostic procedure in clinical practice" (Hunt, 1946, p. 311) and is now utilized globally (Anastasi, 1985; Anastasi & Urbina, 1997; Gottfredson & Saklofske, 2009; Gregory, 2015; Hurt & Tomoyasu, 1995; Van Ree, 2014). The historical roots of psychological testing provide some explanation on the modern-day forms of testing, including its role with both children and the mental health field (Domino & Domino, 2006; Miller & Lovler, 2018; Sattler, 2001; Sattler & Hoge, 2006).

Many regard Alfred Binet as inventing the first modern intelligence test (Binet & Simon, 1916; Minton, 1988; Mülberger, 2017; Sattler, 1982). However, historians contend over almost 4000 years ago in China, basic forms of cognitive testing were developed and commonly utilized to assess various skills, such as problem solving, thinking styles, and creativeness (Higgins & Sun, 2001; Oakland, Poortinga, Schlegel, & Hambleton, 2001, p. 5). Similar to today, some of this testing was used to make vast determinations, such as to regularly assess the fitness for duty of the Chinese emperor (Gottfredson & Saklofske, 2009; Higgins & Sun, 2001). Further, Higgins and Sun (2001) contend that what is modernly termed standardizing testing procedures and proficiency testing can be linked to early assessments in China. According to McReynolds

and Ludwig (1987), a rudimentary form of a rating scale has been found linked to Galen and the creator of the first rating scales that contained psychological components is thought to be Christian Themistius in the 1700s (Brett, 1921, p. 26). In regards to children, Heffer et al. (2009) suggested that forms of child assessment have existed since the first-time parents or caregivers have observed, tracked, and compared milestones in children's development. Although it is incredibly difficult to calculate the origins of such an informal assessment, one can argue these early queries into gauging human abilities are origins for the burgeoning science of child assessments (Heffer et al., 2009). Nevertheless, prior research considers the pervasive application of psychological testing to have begun in the late nineteenth century, although it was not until the twentieth century that nations began to replicate psychological assessment practices (Gregory, 2015; Hurt & Tomoyasu, 1995; Mülberger, 2017; Sattler, 2001).

Before Binet, there were several well-known pioneers that helped build a foundation for psychological testing. German psychologists, including Gustav Fechner, Ernst Heinrich Weber, and Wilhelm Wundt, laid the scientific groundwork for psychology (Brett, 1921; Geisinger, 2000; Kelley, Sexton, & Surbeck, 1990; Minton, 1988; Mülberger, 2017). Then, in the mid-1800s, European psychologist Francis Galton and his colleagues constructed the first battery of tests (Brett, 1921). At the time, Galton and many others defined cognitive abilities through evolution and emphasized individual differences (McReynolds & Ludwig, 1987). Wundt was able to obtain consistent results across trials, although there were differences found between examinees (Bersoff, DeMatteo, & Foster, 2012). Some viewed the differences between people as errors; in contrast, others such as American psychologist James Cattell, viewed these variations as significant (Geisinger, 2000). In 1890, Cattell outlined the utility and applications of psychological instruments in the well-known paper "Mental Tests and Measurements" that

facilitated the accepting of psychological testing in the United States. Shortly after, Binet, Lewis Terman, and Edward Thorndike addressed standardized testing procedures providing the framework for what is used presently (Minton, 1988; Sattler, 1988). A combination of economic, social, and cultural conditions during this time Europe and the United States provided a platform for cognitive and intelligence testing to further thrive (Anastasi, 1985; Kelley et al., 1990; Minton, 1988).

Binet persisted and sought to reliably identify intellectual disabilities, formally known as "mental retardation" within children (Binet & Simon, 1916). The Binet-Simon (1905) scale was developed and used in North America and Europe (Binet & Simon, 1916; Wasserman, 2018). Binet's scales were found to be reliable thus paving the way for increased use of cognitive testing and stimulating advancements within scale development, statistical analyses, and interpretations of results (Binet & Simon, 2916; Gottfredson & Saklofske, 2009; Kamphaus et al., 2000). After improvements, this measure became a model for future scales, and is now known as the Stanford-Binet (Sattler & Hoge, 2006). Although some cognitive testing was being used within schools at this time, the need for identifying capable troop recruitment procedures during the first World War drove the development of psychological testing globally, especially within the United States (McIntire & Miller, 2007, p. 60; see also Haynes, Richard, & Kubany, 1995; Yoakum & Yerkes, 1920). During this time, personality measures were created (Gibby & Zickar, 2008; Mischel, 1968). Early education and aptitude testing gained momentum as well, with Yerkes designing measures for specific military skill sets (Harrell, 2017; Yoakum & Yerkes, 1920). Until World War II (WWII), aptitude testing did not have as much traction compared to intelligence measures (Goodenough, 1949; Sattler, 2001). With a similar trend, children's tests in general, seemed to lag behind, with the Wechsler Intelligence Scale for

Children (WISC) emerging in 1949 and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) in 1967 (Weiss et al., 2016).

During the 1900s, apart from intelligence and aptitude tests, various personality and psychopathology measures emerged (Anastasi & Urbina, 1997; Mischel, 1968). Projective measures, such as the Rorschach inkblots was invented in 1920s, the Thematic Apperception Test (TAT), was introduced in the 1930s, and the Minnesota Multiphasic Personality Inventory (MMPI) in the 1940s (Costa & McCrae, 1992; Gibby & Zickar, 2008; Goodenough, 1949; Mischel, 1968). As development of test measures grew, so did the advancement of psychometrics, new scoring systems, increased generalizability, and detailed manuals (Kelley et al., 1990; Leong, Park, & Leach, 2013; Swerdlik & Cohen, 2013).

Ethical Considerations

Globally, ethics is a broad term that attempts to integrate and adapt given the context (Leong et al., 2013). Within professional psychology, ethical standards are produced by varying task forces to merge endorsed values, ethical principles, and guidelines for protecting the public and profession (APA, 2017a; Leach & Oakland, 2007; Thorndike & Thorndike-Christ, 2010). In regards to psychological testing and assessments within the United States, there are presently three governing sources of ethics standards: (1) the Ethical Principles of Psychologists and Code of Conduct (APA, 2017a), (2) the Standards for Education and Psychological Testing (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US) (2014), and (3) the Guidelines for Computer-Based Tests and Interpretations (APA, 1986, 2017a).

These available guidelines work together to help to address ethical dilemmas when conducting assessments. The APA's Ethical Principles of Psychologists and Code of Conduct states:

> This Ethics Code applies only to psychologists' activities that are part of their scientific, educational, or professional roles as psychologists. Areas covered include but are not limited to the clinical, counseling, and school practice of psychology; research; teaching; supervision of trainees; public service; policy development; social intervention; development of assessment instruments; conducting assessments; educational counseling; organizational consulting; forensic activities; program design and evaluation; and administration. This Ethics Code applies to these activities across a variety of contexts, such as in person, postal, telephone, Internet, and other electronic transmissions. These activities shall be distinguished from the purely private conduct of psychologists, which is not within the purview of the Ethics Code. (APA, 2017a, pp. 495-496)

Further, to specifically address use of assessments, the Ethical Principles of Psychologists and Code of Conduct outlines 11 guidelines in Standard Nine to adhere to for Assessment, which includes:

> 9.01 Bases for assessments. 9.02 Use of assessments. 9.03 Informed consent in assessments. 9.04 Release of test data. 9.05 Test construction. 9.06 Interpreting assessment results. 9.07 Assessment by unqualified persons. 9.08 Obsolete tests and outdated test results. 9.09 Test scoring and interpretation services. 9.10 Explaining assessment results. 9.11 Maintaining test security. (APA, 2017a, pp. 508-510)

Along with the United States' main sources of ethical standards, a consensus of seven overlapping ethical issues emerge when conducting an evaluation with specifically children (Leong et al., 2013; Oakland, 2004; Redman, 2006). When evaluating children, ethical concerns include parental consent and involvement, children's rights, confidentiality, identification of patient, separation of the parent and child during assessment, nondiscriminatory assessment practices, and the use of multiple sources of information and appropriate measures (Elwyn et al., 2012; Heffer et al., 2009; Leong et al., 2013; Sattler & Hoge, 2006; Swerdlik & Cohen, 2013).

Regarding to Evidenced Based Practices (EBPP), the APA (2006) Council of Representatives explained it is the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences. APA (2006) explicitly indicates the purpose of EBPP "is to promote effective psychological practice and enhance public health by applying empirically supported principles of psychological assessment, case formulation, therapeutic relationship, and intervention" (p. 237). Evidence Based Assessment (EBA) combines EBPP along with research and theory to help aid in choices about what and how to measure as well as how to interpret and implement results (Hunsley & Mash, 2008).

Professional competencies. Although forms of assessment have been applied in various ways, over many years, the type, intensity, and consistency of training practices within graduate programs has undergone only periodic exploration since 1960 (Sundberg, 1961; Watkins, 1991). Recent data gathered from APA accredited programs by Ready and Veague (2014) illustrated an increase in assessment training, with similar forms of assessment instruction across Clinical-Science, Scientist-Practitioner, and Practitioner-Scholar training models. Consistent with other surveys conducted, the predominant instruments (varying editions) that were taught included Wechsler Adult Intelligence Scale, Wechsler Intelligence Scale for Children, Minnesota Multiphasic Personality Inventory, the Beck Depression Inventory, and the Woodcock Johnson Tests of Achievement (Ready & Veague, 2014; Sattler, 1988; Tuma & Pratt, 1982; Watkins, 1991).

APA includes "assessment" within the core competencies and ethics of training (Leong et al., 2013, p. 267). Further, to ensure that "practitioners are competent" (Clay, 2010, p. 48), the APA has outlined in detail core competencies or benchmarks within several topics, such as

scientific knowledge and methods, cultural diversity, and assessment. Krishnamurthy et al. (2004), a work group, presented eight core benchmarks within the practice of assessment, which included: (1) understanding psychometric theory; (2) foundations of psychological assessment (e.g., skill, empirical, theoretical); (3) assessment techniques; (4) assessing outcomes; (5) critical evaluation skills; (6) collaborative professional relationships; (7) assessment and intervention relationship; and (8) technical assessment skills.

Overall, when conducting assessments with children, the evaluator should be cognizant of the evolving ethical codes and competencies. By staying informed, being self-reflective, and consulting, professionals practicing assessment can maintain proper procedures and be able to effectively apply these guidelines to the unique needs of children and their families.

Limitations of Assessments with Children

The history, value, and ethics guidelines help support the advantages of psychological assessments with children. To further produce accurate results and abide by the ethical guidelines, it is critical to review the limitations of conducting assessments (Heffer et al., 2009). Several studies agree that when evaluating children, there is a lack of appropriate measures available (Eisman et al., 2000; Gregory, 2015; Loyens, Kirschner, & Paas, 2012; Sattler, 1988; Swerdlik & Cohen, 2013). The sparse amount of valid, reliable, and standardized instruments available for assessing young children are often adapted from procedures normed with older children (Panter & Bracken, 2013). Another limitation is the variations in interpretations and application of findings (Oakland, 2004). To clarify, uniformity between two different psychologists obtaining similar data may introduce disparate interpretations, in addition to pursuing different courses of treatment (Pope & Vaquez, 2007). Another concern is the validity within measures (Leach & Oakland, 2007). Throughout history, psychological tests have been

criticized for attempting to measure variables that are hypothetical constructs, not easily observable, or inferences (Krishnamurthy et al., 2004; Leach & Oakland, 2007; Thorndike & Thorndike-Christ, 2010). As theories that guide tests adapt, become more refined, or new theories emerge, older instruments become obsolete; therefore, to maintain valid and reliable tests, measures must be updated often (APA, 2006). It is considered unethical to use and misrepresent the results from outdated or inappropriately normed measures (Krishnamurthy et al., 2004; Loyens et al., 2012).

There are obvious limitations in relation to cultural considerations and generalizability of psychological instruments to the growing population. Specifically, many measures utilized in the United States were standardized in English and have limited generalizability to individuals who speak other languages (Krishnamurthy et al., 2004). In regards to diversity, socioeconomic status, and cultural backgrounds, unfortunately, many psychological tests have been normed using White, middle-class individuals (Adair, Bélanger, Dion, & Sabourin, 1998; Dana, 1998). Thus, evaluators should strive to ensure that tests, procedures, and psychometrics are conducted in ways that are interpreted with caution, properly represented, and protect the rights and promote the well-being of the children being assessed (Adair et al., 1998; APA, 2017a; Krishnamurthy et al., 2004; Russo, 2015; Sattler & Hoge, 2006).

Definition of Terms

Understanding the terminology utilized in the literature pertaining to psychological assessment with children is a necessary skill set for clinicians to possess to accurately apply the concepts addressed throughout this project. For the purposes of this clinical research project, the following list of terms were adapted from the literature reviewed.

Assessment/Evaluation. The comprehensive process of deriving meaning from data, achieving a broad but detailed description and understanding of individuals, behaviors, environments, and the reciprocal interactions among each of these elements. The dynamic integration of information collected from numerous of sources, some of which likely include formal psychological test data. Although tests and assessments are not the same thing, a substantial portion of assessment activities depend on the use and interpretation of results obtained from individual tests (American Educational Research Association et al., 1999; Kolen & Hendrickson, 2013; Loyens et al., 2012).

Child/Youth. Any person under the age of 18 as legally defined by The United Nations (1989). Of note, in this project, the term *child* will appear in most references to children and adolescents. However, when analyzing prior research and distinguishing age groups becomes significant, *child* will be used for all children between 0 and 12 years of age and *adolescent* for those between ages 13 and 18.

Clinical interviews. Clinical information gathering, structured or unstructured, may include parents, caregivers, children, teachers, and the like.

Informal assessments. Any evaluation measure that lacks structure or scientific validity, such as brief observations to obtain language samples from the child.

Instrument/measures. Psychometric inventories, surveys, and questionnaires that are widely accepted tools in the field of behavioral health (Kolen & Hendrickson, 2013).

Observations. Behavioral observation is one of several measurement approaches available to clinicians. It refers to the act of carefully and attentively listening to, watching, and studying the child and their environment (Bakeman & Quera, 2012, p. 207).

Direct. Obtained by the clinician and clinician performs the act of watching a child within a setting (e.g., office, school).

Indirect. Information that is based on the reports of others who completed the direct observing.

Test battery. Incorporating multiple methods to obtain information, cross-check hypotheses, and to choose tests and measures to complete an assessment (Meyer et al., 2001). These methods include self-reports, performance tasks, observations, and information derived from behavioral or functional assessment strategies (Haynes et al., 1995; Sattler, 1988).

Tests/Testing. A subset of the assessment process, representing one method among many (e.g., observations, interviews, developmental histories) of obtaining important information about individuals. This includes the administration of written, visual, or verbal psychological instruments and measures to assess the cognitive, socio-emotional, and behavioral functioning of children. Of note, the testing process often begins with a clinical interview to obtain relevant background information (e.g., family background, social relationships, school history, medical/developmental/mental health history, and overall presenting concerns). Tests often provide information that is essential to a comprehensive assessment (American Educational Research Association et al., 1999; Kolen & Hendrickson, 2013; Loyens et al., 2012).

Statement of Problem

In an era of evidence-based practices, increased health care demands, and clinical decision-making, it is important to explore factors contributing to the decision-making behind conducting assessments with children. Psychological assessment research is robust, including validity (e.g., Achenbach, 2005; Messick, 1995), interpretation techniques (e.g., Benson & Donohue, 2018; Weiss et al., 2016), and report writing (e.g., Harvey, 2006; Thorndike &

Thorndike-Christ, 2010). However, much of the research subsumes that psychologists have implemented their clinical judgment and deemed testing necessary. Setting, nature of referral, and clinical context collectively influence the process by which psychological testing and evaluations of children are employed, be it clinical, educational, employment, or forensic in nature (APA, 2006). Regardless of the setting, the literature pertaining to assessment including APA's Code of Ethics (2017a), have variations alluding to the use of "clinical judgment" in determining the practice of testing with children (Cronbach, 1949; Eignor, 2013; Krishnamurthy et al., 2004; Russo, 2015; Sattler & Hoge, 2006; Watkins, 1991). Prior research (e.g., APA, 2006; Gregory, 2004; Heffer et al., 2009; Loyens et al., 2012) offers broad recommendations that do not benefit the assessor, as they still must synthesize several domains (e.g., current ethics, evolving procedures, theories, and literature) prior to administration. Yet, prior studies have not compiled universal guidelines to establish practical, evidence-based conceptual and practical frameworks for decision-making to deem assessments necessary following a referral question.

Significance of the Study

In the healthcare field, using suggested guidelines as an aid in decision-making along with clinical judgment is common practice across disciplinaries (Croskerry, 2015; Garb, 2005; National Academies of Sciences, Engineering, & Medicine et al., 2015). For example, research has provided evidence for several broad evidence-based theories within medical decision-making, such as Fishbein the theory of reasoned action and Prochaska the transtheoretical model (Reyna, 2008). Over time, extensive literature has developed for decision-making models regarding ethical decisions within psychology (Riggin & Lack, 2018), as well as approaches to choosing assessments for a specific presenting problem within the medical field (McGinnis, Hack, Nixon-Cave, & Michlovitz, 2009). To limit clinical errors, ethical pitfalls, and establish

evidence-based decision-making, clinical guidelines are essential. Yet, there is a paucity of research addressing universal, clinical guidelines, or identifying salient factors governing a clinician's decision-making in creating assessment batteries.

Purpose of the Study

There are many psychological tests and methods supported by literature to evaluate a child's psychological functioning. However, the factors that attribute to an evaluator's clinical judgment when choosing to conduct testing remains unclear. Thus, the purpose of this clinical research project was to explore the contributing factors behind clinical child psychologists' decision-making in relation to determining the necessity of conducting assessments with children. The factors and themes that emerge prior to a psychologist conducting a psychological evaluation with a child were explored through a review of salient clinical literature. Areas of future research are suggested in effort to expand upon the present foundation of knowledge. In efforts to aid in the development of evidence-based, clinical guidelines for assessments with children a decision-making model are proposed.

Literature review questions. This research project sought to identify factors and guidelines applicable to clinical decision making in relation to the selection and use of psychological tests, such as instruments and measures, for child psychological assessments. Furthermore, this project provides an examination of the existing research and literature relevant when conducting psychological evaluations and testing with children to answer three research questions:

Literature Review Question #1: What is the role of psychological testing in the practice of child clinical psychology?

Literature Review Question #2: What is the clinical utility of psychological assessment with children?

Literature Review Question #3: What approaches are recommended as best practices in clinical decision-making and child psychological assessment?

Lastly, the author proposes factors for clinical decision-making to augment incorporation of psychological testing in clinical practice.

Methods

Search strategy. An evaluation of the peer-reviewed literature generated from PsycARTICLES, PsycINFO, PsycEXTRA, PsycTESTS, PsycBOOKS, and PubMed. The following limitations were placed on the search: peer-reviewed sources; English-language publications; and full text of the article must be available. The terms for psychological (e.g., clinical psychology, psychology), assessment (e.g. testing, evaluation), decision-making (e.g., clinical judgement), and children (e.g., youth, child, adolescent), along with Boolean connectors (i.e., and, or) were applied. Of note, terms such as "adult" or "teen" were utilized sparingly to illustrate commonalities or comparisons within clinical practice. Using these criteria, the search produced a vast number of articles. Given the extensive research literature available on psychological assessments, a comprehensive review of all relevant materials was not conducted. Instead, to identify articles that met the inclusion criteria, the reference lists were hand searched of each article relevant to psychological assessment decision-making including review. Several of the articles only briefly mentioned the search terms and were deemed not relevant to the topic of the current clinical research project. After systematically examining the peer-reviewed results, the articles with strong relevance to clinical decision-making and psychological assessment with children sustained a more detailed and thorough review for this project.

For the purpose of collecting recent statistics from respected polling services (e.g., United States Department of Education), as well as reviewing the APA (2017a) principles, ethics, and code of conduct, substantial effort was placed into ensuring all non-peer-reviewed information, such as government agencies and reports, were relevant, purposeful, and accurate.

CHAPTER II: PSYCHOLOGICAL TESTING AND CHILD CLINICAL PSYCHOLOGY

Psychological testing is viewed as a professional competency and a formative diagnostic tool (Meyer et al., 2001), while within the child population, there has been an incremental increase of psychological testing (U.S. Department of Education, 2018). To better understand the patterns of use, application, and training in relation to children and child psychologists, this chapter explores the utilization and clinical utility of psychological testing with children.

Understanding the Target Population

Youth and mental health. The onset of many mental illnesses occurs in childhood or adolescence (Castellanos-Ryan et al., 2016; Kessler et al., 2007). Other studies (i.e., Leadbeater, Thompson, & Gruppuso, 2012) show that symptoms often continue or intensify throughout adulthood, especially without services. Numerous studies have shown the benefits of early interventions and the importance of accurately diagnosing (Ivnik et al., 2000; Stone et al., 1999). Specifically, the value of accurate diagnoses of children is well supported in the literature, as it leads to effective, early, targeted interventions (Ivnik et al., 2000). In turn, these interventions are correlated with lifelong restorative effects including decreased symptomology in adulthood (Majnemer, 1998; Panter & Bracken, 2013). Thus, when a child is experiencing atypical behavioral, social, or academic problems, they are often referred for some form of psychological testing or screenings (Kieling et al., 2011). Across current studies, psychological testing with children is frequently broken down into categories (e.g., internalizing or externalizing disorders) or diagnostic rule ins and rule outs of neurodevelopmental disorders, specific learning disabilities, sensory processing disorders, anxiety disorders, mood disorders, sleep or eating disorders, and trauma related disorders (Altmaier & Tallman, 2013; American Psychiatric Association, 2013; Bracken & Nagle, 2017; Zero to Three, 2016).

Prevalence of mental health in youth. Globally, childhood mental health disorders rates are rising. More specifically, within the United States, the National Comorbidity Survey by Merikangas et al. (2010) found that approximately 32% of children and adolescents experience symptoms of mental illness, while other studies estimate that upwards of 40% of youth have impairment related to emotional or behavioral difficulties (Kieling et al., 2011; The Mental Health Foundation, 2016). Further datum reported by the National Alliance on Mental Illness [NAMI] (2019) indicated that 7.7 million (16.5%) of children ages 6-17 were diagnosed with a mental health disorder. While current national trends denote steady increases in mental health disorders among youth (Merikangas et al., 2010), more comprehensive studies of all children under the age of 18 predict upwards of 17.1 million children having a diagnosable mental disorder (Kieling et al., 2011). It is estimated about one in every five youth in the United States meets criteria for a mental health disorder, with severe impairment across their lifetime (Merikangas et al., 2010) that impacts their functioning significantly at home, school, with peers, and within their communities. The most common diagnosis among youth fall within anxiety disorders (31.9%), followed by behavioral disorders (19.1%), mood disorders (14.3%), and lastly substance use disorders (11.4%) (Merikangas et al., 2010).

Referrals. Several models for referral sources to mental health services for youth currently exist and are well supported (Burns et al., 1995; Bush & Iannotti, 1990; Carise & Gurel, 2000; Stiffman, Pescosolido, & Cabassa, 2004; Young & Rabiner, 2015). Common referral sources or the "gateway providers" (Stiffman et al., 2004, p. 190) to mental health services can be divided into two categories: informal (e.g., parents, friends) or formal (e.g., specialists, pediatricians, teachers, child welfare, juvenile justice) gateway providers (Mechanic, Angel, & Davies, 1992; Stiffman et al., 2004). Stiffman et al. (2004) created the Gateway

Model, further supporting access to mental health services for youth who rely heavily on these individuals in that they have initial contact with the youth in need before mental health specialists.

Utilization of mental health services. Even with the aid of national agendas working to provide access to mental health services, national data suggests only one-half of adolescents receive mental health services (Merikangas et al., 2010). Other findings indicate between 70% and 80% of children and adolescents who meet criteria for a mental health diagnosis are undiagnosed, untreated, and do not receive any form of services (Castellanos-Ryan et al., 2016; Eisman et al., 2000; The Mental Health Foundation, 2016). With regard to the small amount who do gain access to services, the data demonstration disparities. Specifically, children with internalizing disorders (e.g., anxiety), while making up much of the impacted population, are steadily less likely to receive services with a utilization rate of less than 20% (Merikangas et al., 2010), in comparison to children with an externalizing disorder (e.g., ADHD) who receive care over three times more often (73.4%) (Merikangas et al., 2010). Youth with mood disorders fall within the middle at 37.7%, and the lowest rates of utilization are with eating (12.8%) and substance use (15.4%) disorders (Merikangas et al., 2010). A common explanation for these findings is the growing acceptance and understanding of externalizing disorders, such as ADHD as well as the overt, disruptive, easily identifiable symptoms usually observed by common referral sources (e.g., pediatricians, teachers) to mental health (Lemberger, Morris, Clemens, & Smith, 2013; Williams, Palmes, Klinepeter, Pulley, & Foy, 2005). Arguably, internalizing disorders are covert within a school setting, and require awareness from caregivers and youth (Merikangas et al., 2010), suggesting less responsiveness by referral sources.

Barriers to mental health services. Mental health services with children and adolescents are historically underutilized. Despite the increasing number of youth who meet criteria for a mental health diagnosis (Domino & Domino, 2006; Merikangas et al., 2010), research frequently demonstrates barriers (Kieling et al., 2011; Owens et al., 2002) to mental health care, including a lack of problem identification from gateway providers (Stiffman et al., 2004; Young & Rabiner, 2015). The prior research literature also supports strong relationships between low rates of mental health use and youth, ethnic minority, and insurance coverage (Brondolo, 2015; Chow, Jaffee, & Snowden, 2003; Holm-Hansen, 2006; Reardon & Portilla, 2016; Young & Rabiner, 2015).

One study found families who participated in a prevention-intervention project reported barriers due to three main areas (Owens et al., 2002). Owens and colleagues found barriers to mental health for children including (a) perceptions of services (25.9%) such as stigma or distrust of mental health providers; (b) view of mental health symptoms (23.3%) or caregivers' or medical providers' inability to identify the issues; and (c) structural barriers (20.7%), such as lack of access related to financial, transportation, or no provider available. Notably, perceived difficulties with parenting were found to weigh into all three overarching barriers (Owens et al., 2002).

Regarding gateway providers, reportedly between 46 and 96% of formal providers (e.g., pediatricians) fail to identify mental health symptoms and diagnosis (Stiffman et al., 2004). The high occurrence of misidentification or recognition may lead to the absence of referral to the appropriate provider(s), misdiagnosis, or increase of symptomatology (Holm-Hansen, 2006; Kieling et al., 2011; Liang, Matheson, & Douglas, 2016; Owens et al., 2002).

Diversity Considerations

Inequalities. Within the United States, existing research has established children of color utilize mental health services at a much lower rate than their White counterparts (Eisman et al., 2000; Kieling et al., 2011; Owens et al., 2002; Pastor & Reuben, 2005) despite comparable prevalence rates of mental health symptoms/diagnoses (Brondolo, 2015; Knauss, 2001; Merikangas et al., 2010). A survey conducted by Young and Rabiner (2015) added to these findings, revealing parents of color endorsed more barriers (e.g., stigma and financial) when accessing mental health than medical services for their children.

Referrals. In addition to the lack of identification and subsequent referral by gateway providers, White youth comprise the majority of the referrals from professional, health care individuals (Stiffman et al., 2004). In comparison, youth within ethnic groups are more likely referred by educational sources following behavioral incidents or legal agencies (e.g., child welfare) (Daryanani, Hindley, Evans, Fahy, & Turk, 2001; Holm-Hansen, 2006; Reardon & Portilla, 2016; Stiffman et al., 2004). Within the youth of color, more significant disparities have been correlated with Black children, as they often obtain a majority of mental health referrals after being placed in more restrictive facilities such as foster care, detention centers, or residential treatment (Holm-Hansen, 2006; Stiffman et al., 2004; Young & Rabiner, 2015).

Assessment. Inequalities are present in both barriers and quality of services provided for children of color (Brondolo, 2015; Holm-Hansen, 2006). The small number of children within racial/ethnic groups who can access services often receive care that is deemed "inappropriate, fragmented, or inadequate" (Holm-Hansen, 2006, p. 2; see also Chow et al., 2003; Reardon & Portilla, 2016). Across the literature, assessment practices conducted with youth of color denote bias measures which lack applicability with diverse populations (Liang et al., 2016; Sladek,

Umaña-Taylor, McDermott, Rivas-Drake, & Martinez-Fuentes, 2020) thus "leading to inaccurate diagnoses and inappropriate treatment plans" (Holm-Hansen, 2006, p. 8). Within practice, across disciplines, assessment bias has a lasting impact. In support, another study conducted by Pastor and Reuben (2005) collected data from the National Health Interview Survey (NHIS) 1997-2001, which is a cross-sectional household survey conducted by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC). The data obtained from the NHIS was then entered into a statistical program (SUDAAN) to generate nationally representative figures. The sample included 21,294 children between the ages of 6 and 11 years. Overall, Pastor and Reuben's (2005) findings suggested that the frequency of ADHD and the use of prescribed medication was at lower rates for Latinx and Black children, compared to their Caucasian peers. Additionally, they found adjusting for confounding socioeconomic variables (e.g., income and health insurance) did not reduce the existing disparities (Pastor & Reuben, 2005).

Testing in Practice

Reasons for testing. Psychological tests can be administered to children for several reasons. Hunsley and Mash (2008) denote seven primary areas in their frequently cited work:

> (a) diagnosis (i.e., determining the nature and/or cause[s] of the presenting problems, which may or may not involve the use of a formal diagnostic or categorization system), (b) screening (i.e., identifying those who have or who are at risk for a particular problem and who might be helped by further assessment or intervention), (c) prognosis and other predictions (i.e., generating predictions about the course of the problems if left untreated, recommendations for possible courses of action to be considered, and their likely impact on the course of the problems), (d) case conceptualization/ formulation (i.e., developing a comprehensive and clinically relevant understanding of the patient, generating hypotheses regarding critical aspects of the patient's psychosocial functioning and context that are likely to influence the patient's adjustment), (e) treatment design/planning (i.e., selecting/developing and implementing interventions designed to address the patient's problems by focusing on elements identified in the diagnostic evaluation and the case conceptualization) (f) treatment monitoring (i.e., tracking changes in symptoms,

> functioning, psychological characteristics, intermediate treatment goals, and/or variables determined to cause or maintain the problems), and (g) treatment evaluation (i.e., determining the effectiveness, social validity, consumer satisfaction, and/or cost-effectiveness of the intervention). (p. 6)

Within these reasons for assessments, often specialty referrals, such as forensic or career referrals, can fall within one of these seven categories. For example, when a psychologist receives a forensic referral to rule out an intellectual disability, then the forensic referral for diagnosis or might be considered screening depending on the exact question. Of note, Hunsley and Mash's (2008) work provides an example of clinical decision to perform testing being subsumed within literature and clinical judgment already being implemented. Comparatively, the authors' frequently referenced work offers an integral component to the foundational guidelines for testing. Specifically, these seven categories for administering testing are essential for clinicians to be aware of given that before the evaluator administers any test, in regards to conducting EBA, it is assumed or expected to first identify the intention for testing (Hunsley & Mash, 2008; Prinstein, Youngstrom, Mash, & Barkley, 2019; Sattler & Hoge, 2006).

Prevalence of child psychologists. According to APA's Work Force Studies (Lin & Stamm, 2016), within the United States there are approximately 106,000 licensed psychologists while the U.S. Bureau of Labor Statistics (2018) reports there are roughly 124,750 psychologists employed. Between 23% to 34.2% of psychologists provide services to children and adolescents (APA, 2016). As discussed in the previous chapter, the scope of child psychological testing is vast, making it difficult to calculate the precise number of evaluations conducted annually.

Current Utilization of Testing

Educational systems. The National Center for Education Statistics (NCES) reported seven million students receiving services under the Individuals with Disabilities Education Act

(IDEA) in the 2017-18 academic year. A further breakdown of the seven million students highlights the classifications of the students who received special education services under the IDEA. In the 2017-18 academic year, the majority (34%) had specific learning disabilities; 19% of students had speech or language impairments; 14% had other health impairments (e.g., rheumatic fever, asthma, diabetes); and 10% of the students were diagnosed with autism, developmental delays, intellectual disabilities, and emotional disturbances. Less than 2% of the students served under IDEA are reported to have multiple disabilities or physical impairments (e.g., hearing, visual, orthopedic, traumatic brain injuries) (NCES, 2019).

Limitations. These students receiving services under IDEA account for 14% of total public-school enrollment, which has steadily increased (NCES, 2019). This percentage means that over seven million children in the United States participated in some form of a psychoeducational evaluation to qualify for these services. Although these figures reinforce the upsurge of psychological testing, they also have several limitations. First, these statistics do not account for the children who underwent testing and did not meet the threshold to obtain services. They also do not specify how many children received an initial or additional evaluation from other sources (e.g., private practitioner, hospital). Additionally, these numbers are not representative of the entire child population (e.g., private schools, homeschooling).

The U.S. Census Bureau (Davis & Bauman, 2013) reported there are 59 million total students enrolled (ages three and over) within an academic year; including five million in nursery school, four million in kindergarten, 33 million in first through eighth grades, and 17 million in ninth through twelfth grades. More recent studies reported private school enrollment for kindergarten through twelfth grade to be around 4.4 million in 2018 (United States Census Bureau, 2019). This means that if the number of students enrolled in private schools required

services at a similar rate (14%) to those receiving services under IDEA, the number of students that would require some form of psychological testing would be over 615,000 students. This figure is over six times as many active, licensed psychologists currently, and does not account for the seven million children already receiving services under IDEA. Still, other studies estimate anywhere between 100 to 200 million child evaluations are conducted annually that includes psychoeducation, cognitive, forensic, emergency, and other types of evaluations (Bersoff et al., 2011; Heffer et al., 2009; Kamphaus et al., 2000; Pikulski, 1990; Sattler & Hoge, 2006; U.S. Department of Education, 2018). These figures support the increased use and need for testing with youth.

Testing within clinical practice. Numerous early studies (e.g., Cronbach, 1949; Tuma & Pratt, 1982; Watkins, 1991) as well as current work (Wright et al., 2017) agree; psychologists, despite their practicing setting, provide assessment services. In addition, Meyer et al. (2001) highlighted that assessment is considered the second most crucial clinical activity for psychologists following psychotherapy. Further, a study found a significant focus on assessment fluency as a requirement when investigating job description adverts for clinical psychologists (Kinder, 1994). Kinder's work further demonstrates that education and competencies in the psychological assessment are crucial to psychologists.

Other studies have attempted to examine the practice habits of clinical psychologists. Out of their average number or clinical hours per week, psychologists serving children spend between 33% to 75 % conducting assessments (Palmiter, 2004; Tuma & Pratt, 1982; Watkins, 1991; Wright et al., 2017). Tuma and Pratt (1982) surveyed to gain a better understanding of the practice and training experiences of clinical child psychologists. Specifically, their survey identified in a sample of 358 respondents from APA's Section 1 of Division 12; respondents

spent an average of 33% of their practice in assessment with children and 51% in treatment. Tuma and Pratt noted psychologists working with children spend more time in assessment, compared to the 24% reported by psychologists conducting assessments with adults. Recent data collected through APA Survey of Psychology Health Service Providers (APA, 2016) yielded similar results. Out of the 5,325 participants who completed the survey, approximately 75% were APA members and 25% were non-members (APA, 2016). Of the total, "fifty percent of respondents reported never providing services to children" (APA, 2016, p. 22). Data from APA's survey (2016) as well as others (Miller & Lovler, 2018; Russo, 2015) have repeatedly found when working with adults, clinicians endorsed spending less time on assessment services compared to those working with children. These studies offer valuable insight, as they suggest across settings, psychologists serving children spend more time practicing testing than those working with adults.

Gap between training and practice. Despite the increase in demands for testing and core competencies including assessment, several findings allude to deficits in training for child psychologists (Cook & Coyne, 2005; Psihogios, Gutiérrez-Colina, Iskander, Wasserman & Ramsey, 2019). Over three decades ago, Tuma and Pratt's (1982) survey compiled deficiencies into five main categories: lack developmental and clinical child courses; child assessment and therapy training; direct services with children across settings; supervision; and "too much emphasis on techniques suitable only for adults" (p. 33). Unfortunately, recent surveys still report training falling short in areas of assessment (Psihogios et al., 2019). Of note, regarding the areas of training that were deemed most advantageous for practicing child psychologists, included child coursework, specifically child assessment coursework (Cashel, 2002; Kaplan & Saccuzzo, 2008; Psihogios et al., 2019; Tuma & Pratt, 1982). Furthermore, Wright et al. (2017)

conducted a study with 1,973 licensed psychologists who were members of APA Practice Organization (APAPO). Wright et al. (2017) found 73% of the psychologists endorsed additional training (e.g., continuing education) in assessment skills as a significant barrier to conducting assessment in practice.

Previous research has illuminated areas in need of improvement within graduate training of psychological assessments and as a result, some psychologists have proposed ways to procure competency in psychological assessment. One psychologist (i.e., Acklin, 2002) recommended implementing advanced postdoctoral specialization in assessment, while others (e.g., Kaslow et al., 2004) agreed, but added it may be necessary to implement credentialing for specific areas of assessment expertise to solve the problem. Although these specific approaches remain briefly addressed in the literature, many others argue the need to develop and standardize parameters for psychological tests used to ensure the assessors' competency. In further agreement, there has been a demand for assessment guidelines and scientifically sound standards regarding psychological test usage and assessment globally (Bartram, 2001).

The lack of continuity between graduate training and clinical practice has led to a closer look into possibly correlated factors, such as psychologists' models of training or managed care. After conducting a review of training and clinical practices of psychological assessment over a span of 30 years, Watkins (1991) found that assessment tools remained incredibly stable even across practice settings (e.g., medical center, inpatient, child, adult). In turn, he suggested three reasons for the stability. First, Watkins hypothesized that these instruments might be the best tools available at the time. Second, Watkins (1991) noted that “old habits can be hard to break or modify. Because these identified tools have such a long history of practice, usage, and (in some cases) research behind them, they may well be the methods about which psychologists first think

when it is time to perform a psychodiagnostic assessment" (pp. 431-432). Finally, Watkins proposed a "reinforcing cycle" (p. 432) by training programs. The researcher explains as the same instruments are taught in training programs, they are then utilized most in practice, thus shaping usage patterns in research that reinforce training programs to continue focusing on similar assessment tools in classes (Watkins, 1991, p. 432). Upon further observation, different training models, such as the scientist-practitioner or practitioner-scholar model (Mihura, Roy, & Graceffo, 2017; Ready & Veague, 2014) and managed care (Cashel, 2002) have not significantly impacted assessment training in graduate programs. Instead, it has been successfully established that across training models, there has been consistency in types of classwork offered, such as instruments taught and psychometric training (Cook & Coyne, 2005; Evans & Finn, 2017; Lovitt, 1974; Watkins, 1991). Still, the apparent split between training received and requirements for practice suggests the demand for modifications and further improvements of assessment training (e.g., graduate education, post-graduate, and continuing education).

Summary

Research has shown a steady rise in childhood mental health rates internationally (Kieling et al., 2011; Merikangas et al., 2010; National Alliance on Mental Health, 2019). Unfortunately, between 46 and 96% of common referral sources (e.g., pediatricians) to mental health fail to identify mental health symptoms and diagnosis (Stiffman et al., 2004). Other findings indicate between 70% and 80% of children and adolescents who meet criteria for a mental health diagnosis are undiagnosed, untreated, and do not receive any form of services (Castellanos-Ryan et al., 2016; Eisman et al., 2000; The Mental Health Foundation, 2016). There are several explanations for the underutilization of mental health services in children. Some barriers with strong support include, stigma, insurance coverage, and being a child of color

(Brondolo, 2015; Chow et al., 2003; Holm-Hansen, 2006; Reardon & Portilla, 2016; Young & Rabiner, 2015).

Psychological assessment has been established as an integral service component of practicing psychologists across diverse clinical settings (Krishnamurthy et al., 2004; Merikangas et al., 2010; Piotrowski, 1999). Many psychologists utilize psychological assessment to inform diagnosis and treatment recommendations (Lemberger et al., 2013; Mash & Hunsley, 2005; Mülberger, 2017; Youngstrom et al., 2017). The importance of psychological assessment has been emphasized within training, core competencies, and ethical standards (APA, 2017a; Bersoff et al., 2012; Plake & Wise, 2014; Sattler, 1988). Studies have consistently found clinicians serving children endorsed spending more time in practice conducting assessments compared to those also working with adult populations (APA, 2016; Piotrowski, 1999; Tuma & Pratt, 1982; Wright et al., 2017).

Despite growing mental health rates in childhood and required competencies in psychological testing, several findings allude to deficits in training for child psychologists (Cook & Coyne, 2005; Psihogios et al., 2019). Although psychologists view psychological assessments as having high clinical utility, studies have found that barriers to conducting assessments include limited child focused training in graduate school (Tuma & Pratt, 1982; Watkins, 1991).

CHAPTER III: CHILDREN, PSYCHOLOGICAL TESTING, CLINICAL DIAGNOSTICS, AND TREATMENT

As discussed in the previous chapter, across settings, reasons for referrals vary yet child psychologists tend to spend more time conducting assessments with youth compared to psychologists also serving adults (APA, 2016; Piotrowski, 1999; Tuma & Pratt, 1982; Wright et al., 2017). Adequate assessment of children and adolescents are critical for clinical diagnosis, treatment plan formulation, treatment evaluation, and research. This chapter explores the relationships between clinical diagnoses, testing, and treatment outcomes.

Disorders and Testing Use

Previously noted, the common mental disorders among youth include ADHD, behavior problems, anxiety, and depression (Merikangas et al., 2010). Other mental health conditions include specific learning disabilities and developmental disabilities, autism, substance use, and self-harm (American Psychiatric Association, 2013; Hersen & Reitman, 2008). Within these prevalent diagnoses, surprisingly only two specific learning disorders and intellectual disability mention using "psychometric evidence" (p. 69) and "psychometrically sound tests of intelligence" (p. 39) in DSM-V's diagnostic criteria (American Psychiatric Association, 2013). Throughout assessment and testing literature, there are no universally accepted guidelines in regards to determining if an instrument has adequate scientific evidence to necessitate its use (Cicchetti, 1994; Hunsley & Mash, 2008; Ivnik et al., 2000).

In terms of organizational stances, APA (2019) explicitly reports, "APA's Testing Office does not maintain, sell, or endorse any tests" (para. 4). Despite the association's unwillingness to endorse any assessments, APA has published *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014). These standards provide

broad criteria to be followed in developing and using psychological instruments. One significant limitation is that the APA's standards do not indicate the minimum psychometric values necessary to indicate that an instrument meets the benchmark of being reliable or accurate. Mash and Hunsley (2005) suggested a reason for the lack of exact values is due to psychometric characteristics not being properties of specific instruments; instead, are "properties of an instrument when used for a specific purpose with a specific sample" (p. 38). Largely, this offers little benefit and increased difficulties in practice in that, without those set precedents, clinicians and researchers must rely on "clinical training and judgment required to interpret test results" (American Psychiatric Association, 2013, p. 37), including maintaining their ethical obligations of ensuring valid assessment tools (Hunsley & Mash, 2008). In contrast, psychologists have the option to utilize alternative resources. For example, Buros Center for Testing offers critical reviews of tests and operates as an independent, non-profit organization. Further, Buros is approved to sponsor APA continuing education, but requires a fee to access any of their materials. Moreover, Hunsley and Mash (2008) have made continued attempts to bridge the gap by identifying Evidenced Based Assessment Practices including investigating instruments with questionable psychometric validity. At present, APA does not enforce the use of alternative organizations or regulate test usage; in turn, resources such as Buros, are often viewed as supportive rather than necessity.

When examining common disorders such as child anxiety disorders, the literature strongly suggests a need for more reliable and valid assessment measures (Kessler et al., 2007; Schniering, Hudson, & Rapee, 2000). Moreover, current critiques of existing anxiety measures for children include the inability to distinguish between diagnostic groups (e.g., social or general) (Schniering et al., 2000). Similarly, DSM-V criteria for autism spectrum disorder does not

require psychodiagnostic testing, although the Autism Diagnostic Observation Schedule (ADOS) is considered a *gold standard* measure within several peer-reviewed articles (Falkmer, Anderson, Falkmer, & Horlin, 2013; Kamp-Becker et al., 2018; Luyster et al., 2009). The ADOS is known as the "gold standard" because "good diagnostic accuracy and interpersonal objectivity have been demonstrated for the ADOS in research setting" (Kamp-Becker et al., 2018, p. 1193). Within healthcare, the term gold standard refers to a benchmark or best available assessment with a substantial evidence base (Versi, 1992). However, as mentioned, within psychology, studies investigating specific assessment tools do not have a consistent, enforced, or even recommended value to justify and generalize the title good standard without a minimum benchmark provided as a standard (Mash & Hunsley, 2005). Another consideration is the influence of managed care. For example, when requesting an autism evaluation, some insurance provider's state test requirements must include the Autism Diagnostic Observation Scale (ADOS-2) to access treatment (Tricare, 2018).

Diagnostics

Support for testing. Formal assessments or testing is a differentiating skill within psychological practice compared to other health care providers. For example, a primary care physician obtains specialized training to perform procedures (e.g., biopsies, injections) and some are even trained to deliver babies; however, not all choose to do this in practice (Curet, Darosa, & Mennin, 1999). Similarly, a psychologist conducting testing is like one of these procedures, perhaps even demanding greater acuity. Yet, it is interesting that in practice, some child psychologists may choose not to practice psychological testing or become credentialed in specialized therapies (e.g., play).

Nevertheless, the utility of assessment has been securitized, especially following immense changes in managed health care (Eisman et al., 2000). Around the time of these changes, other studies began to cite a significant decline amid the overall use of testing in professional practice (Cashel, 2002; Piotrowski, 1999). In stark contrast, publications related to the use and application of testing in youth populations have almost doubled every decade since the 1960s (Hersen & Reitman, 2008, p. 8).

In response to managed care's threats to the overall use of psychological assessment services, APA's Board of Professional Affairs established a Psychological Assessment Work Group (PAWG) to evaluate and compile evidence of the efficacy of assessment in clinical practice (Meyer et al., 2001). The PAWG consisted of Meyer and colleagues (2001); together, the group examined 241 meta-analyses on test validity and 800 samples examining multimethod assessment. The PAWG's findings further substantiated various, known clinical utilities for testing. First, the ability to provide details of clinical symptomatology and differential diagnosis. There was utility in psychological assessment as a form of treatment in itself. Additionally, testing was shown to have the capability of describing everyday behavior and predicting its functional capacities. In turn, testing provided differential treatment needs for both medical and mental health conditions, as well as accurate monitoring of treatment over time (Meyer et al., 2001).

Meyer et al. (2001) also made several excellent points to support the value of testing, such as standardization can minimize unconscious bias, thus can reduce legal and ethical issues. Mirroring findings of studies prior, the PAWG was able to demonstrate psychological test validity, is comparable to medical test validity, distinct assessment methods can provide unique sources of information, and the clinicians who did not conduct testing or relied entirely on

interviews were prone to making increased diagnostics errors and missing pertinent information (Meyer et al., 2001). Further, testing has been associated with enhanced abilities to detect subtle behavioral indications of psychological complications (Carlson & Geisinger, 2009; Majnemer, 1998; Meyer et al., 2001).

Notably, clinical interviews (i.e., structured or informal) are a common first tool used in assessment practices (Gresham & Elliott, 1984; Schniering et al., 2000; Williams et al., 2005). While several (Palmiter, 2004; Sattler & Hoge, 2006) are highly supportive of using clinical interviews in practice, others (cf., Mash & Hunsley, 2005; Meyer et al., 2001) suggest it is better practice to add formal psychological assessment to circumvent errors associated with using clinical interviews exclusively.

With the support and some evidence for clinical utility and testing comes limitations. Specifically, the clinical utility of assessment has been defined in various ways, for example, Nelson-Gray (2003) defined "the phrase *the treatment utility of assessment* to refer to the degree to which assessment is shown to contribute to beneficial treatment outcome" (p. 963). Others have extensively studied tests used in assessments (e.g., Anastasi, 1985; Cicchetti, 1994; Meyer et al., 2001; Plake & Wise, 2014); however, these studies almost entirely focus on psychometric criteria (e.g., validity) instead of applied value for the purposes and clinical populations they are intended to assist (Hunsley & Mash, 2008). Through decades of research, individual measures have been shown to aid in diagnostic determination (Binet & Simon, 1916; Minton, 1988); however, less support exists regarding the extent the integration of tests (i.e., assessment methods) adds to clinical utility (Fernández-Ballesteros et al., 2001; Nelson-Gray, 2003). Gaining diagnostic clarity can be derived from assessments and can aid in selecting evidence-based treatment options (Hunsley & Mash, 2014). Specifically, studies (Prinstein et al., 2019;

Roysircar, 2005; Russo, 2015) have exposed that often diagnosis is only one factor in deciding which treatment is best for a client; others (Hunsley & Mash, 2014; Nelson-Gray, 2003) have shown there is limited evidence concerning the extent that the integrated information gathered during assessments progresses treatment outcomes.

Psychodiagnostics

American Psychiatric Association (2013) has emphasized more behavioral and operational criteria for mental health diagnoses. However, the value of psychodiagnostic testing is dependent on the premise that there is a classification system for diagnoses that is valid and has interrater-reliability. Developing a diagnostic classification system has proven to be an increasingly challenging task. Over 61 years, the APA has published six versions of the DSM. For DSM-V alone, APA formed the DSM-5 Task Force to begin revising the manual as well as 13 workgroups focusing on various disorder areas (Regier, Narrow, Kuhl, & Kupfer, 2011). Each consecutive manual has provided more explicit and comprehensive diagnostic criteria. However, many have pointed out that the procedure of differential diagnosis lacks reliability across recent versions of the DSM (Cuthbert & Insel, 2013; Hyman, 2011; Regier et al., 2011).

Importance of diagnosing. Having a diagnostic system with reliability is valuable for many reasons. For example, Jensen-Doss and Weisz (2008) found diagnostic agreement between clinicians and researchers serving youth have been correlated with higher therapy engagement, reductions in reported internalizing problems during treatment, and predicting successful treatment. While other researchers and clinicians support these findings, adding most importantly diagnostic agreement and accuracy allows children access to the proper interventions (Whaley, 2001). Understanding the etiology of a child's symptoms is important, given the diagnosis the child receives can immensely alter the course of evidenced-based treatment

(Ebesutani, Bernstein, Chorpita, & Weisz, 2012). Studies have shown treatment (e.g., therapeutic or pharmaceutical) can be ineffective if the underlying problem is not properly attended to (Wolraich et al., 2019). Because of several factors such as race and gender, children have been found to be commonly misdiagnosed with ADHD (Bruchmüller, Margraf, & Schneider, 2012). In practice, if a child presents with inattention and disruptive behavior and is misdiagnosed with ADHD, they can be prescribed stimulant medication as a result. This, in turn, unnecessarily introduces the incorrect evidence-based treatment (e.g., pharmacologic agents) when the diagnoses are incomplete, incomplete, or inaccurate. Furthermore, the medication will likely be ineffective or can exacerbate symptoms of the true diagnosis, such as anxiety or autism. Equally pertinent, the evidence of early interventions is heavily supported given the strong relationships between benefits in academic achievement, positive behavior, reduction in delinquency and criminality, long-term gains, and educational progression and attainment, among many other domains (Bailey, Hebbeler, Scarborough, Spiker, & Mallik, 2004; Kessler et al., 2007; Panter & Bracken, 2013).

Regarding young children, Bailey and colleagues (2004) collected a nationally representative sample of over 3300 parents of children with or at risk of disability. In line with other research (e.g., Stiffman et al., 2004), medical professionals such as pediatricians were gateway providers to mental health referrals (Bailey et al., 2004). Of the children later diagnosed with a delay, about 86% of families discussed their concerns with a medical professional. According to Bailey et al.'s findings, on average, parents reported concerns at 7.4 months; however, a diagnosis was reported 1.4 months (8.8 months of age) later. Unfortunately, it took almost seven months (average age of 14 months) after the parents' initial concerns for the children to receive a referral or services, while within another cluster of children with less severe

delays, they began to receive services even later, between 23 and 30 months. Between 76% and 82% of the families of children who received early interventions reported their child receives were sufficient in terms of quantity, and approximately 93% indicated that the services were either good or excellent. These findings are significant as they demonstrate early childhood intervention is helpful to reduce the factors that place children at risk of poor outcomes, support the importance of early access to mental health services, and value of accurate problem identification (i.e., diagnosis) (Bailey et al., 2004).

Further diagnostic considerations. Ethically, in the United States, psychologists use the DSM-5 to make a diagnosis. As stated, the DSM-5 does not require psychometric evidence to diagnose most of the developmental disorders prevalent in childhood and adolescence. However, this is not congruent with findings suggesting psychologists serving children spend significantly more time conducting assessments (APA, 2016; Piotrowski, 1999; Tuma & Pratt, 1982; Wright et al., 2017). This then poses the question: why are child psychologists testing more? As of now, there are no scientifically sound tests to provide a mental health diagnosis in the same ways some tests confirm the medical diagnosis (e.g., diabetes) (APA, 2013d). Instead, psychologists are able to incorporate psychological testing to expose empirical findings with regard to symptom clusters, aligned with clinical diagnostic disorder requirements, and then bolster their diagnoses and henceforth treatment plans with greater diagnostic clarity (Kamp-Becker et al., 2018; Leong et al., 2013; Ready & Veague, 2014; Wright et al., 2017). Many note the limitations secondary to the lack of inclusion of psychometric evidence (Cashel, 2002; Eisman et al., 2000; Piotrowski, 1999). Specifically, without the "requirement" or criteria of psychometric evidence to make a diagnosis, managed care (i.e., third-party reimbursors) often

require psychologists to submit for “preauthorization” to indicate “medical necessity” before conducting psychological testing (Eisman et al., 2000, p. 133).

Hymen (2011), a member of the DSM-5 Taskforce and chair of the Mental and Behavioral Disorders chapter for the International Advisory Group in the International Classification of Diseases (ICD-11), has expressed his personal views about the DSM-5. Hymen conveyed a need for increased reliability among clinicians or researchers when using the diagnostic system. Further, Hymen (2011) "wished for" the next diagnostic system to be "constrained" by objective tests (p. 6). Hymen's notion demonstrates the absence of testing within diagnostic criteria. Few would dispute that psychological assessment is a powerful tool for psychologists; however, the effectiveness of testing strongly relies on the skill, knowledge, and competency of the assessor administering and interpreting the test. This limitation with testing has been considered in great detail for decades. For example, in 1992, Masling established that clinical psychologists’ capability to use psychological assessment methods a unique qualification and valuable skill. Masling (1992) also observed that what tests are intended to reveal and how well psychologists are able to apply the tests remain unspecified.

Regarding the use of testing diagnostically for children, there is a need to differentiate between diagnosis due to the different evidenced based treatments. For example, externalizing and internalizing disorders have different modalities for treatment or between subtypes of conduct disorder versus other disruptive behavior disorders (Bornstein, 2017; Hunsley & Mash, 2018; Posada, 2004). A significant concern during psychodiagnostic testing is the similarities across criteria for children, such as the overlap and possible similar presentations between ADHD, autism, and anxiety. For example, if a child has disruptive behavior in the classroom and a Functional Behavioral Assessment (FBA) is generated, then what typically follows

psychoeducationally is a Behavior Intervention Plan (BIP), or Positive Behavior Intervention Plan (PBIP). However, if the intervention is based on externalizing behaviors, and the source of the errant behavior actually is anxiety, depression, or trauma, precipitants that are internalized can go undetected without psychological measures or a keen conceptual eye.

Moreover, when working with children, psychologists receive most of their diagnostic data from adults close to the child. As a result, several studies caution the clinician to be mindful of decision-making errors, as the assessor can be easily influenced by the secondary information gathered by others, such as caregivers or teachers (Baker-Ericzén, Jenkins, & Brookman-Frazee, 2010; Falkmer et al., 2013; Wiener & Costaris, 2012).

Still, diagnoses within practice are viewed as relatively objective (Carlson & Geisinger, 2009) and facilitates communication across providers (Harvey, 2006). Other benefits associated with giving a diagnosis is providing qualifying information that unlocks access to resources, treatment, and other essentials for the child. For illustration, many school systems require classifications and diagnosis to qualify for an individual educational plan (Eignor, 2013; Naglieri, 2013).

Drawbacks of diagnosing. One implication is adding to the stigmatization of mental health (Young & Rabiner, 2015). Often with children, stigma has been linked to both families and children feeling devalued and experiencing discrimination, from themselves and the general public (Mukolo, Heflinger, & Wallston, 2010). Children with the same diagnosis, such as ADHD, can have very different presenting problems. For example, within a school setting, someone with a limited understanding of a child's unique profile, can make vastly inaccurate assumptions about the child and thus treat them differently, according to negative stereotypes (Brondolo, 2015; Riggin & Lack, 2018). Children can also experience negative peer interactions

secondary to observable behaviors such as hyperactivity or tics. Dissimilarly, several authors argue that some children experience relief or comfort following diagnostic label, as they are better able to explain their behaviors to peers and have normalizing experiences (e.g., support groups), consequently helping to mitigate stigma (Mukolo et al., 2010; Young & Rabiner, 2015).

Limitations. According to the Clinical Practice Guidelines for ADHD by the American Academy of Pediatrics, "The use of neuropsychological testing has not been found to improve diagnostic accuracy in most cases, although it may have benefit in clarifying the child or adolescent's learning strengths and weaknesses" (Wolraich et al., 2019, p. 6). However, the authors do not provide citations to studies substantiating this, further illustrating the lack of empirical data available to offer support or discount the necessity of assessments in practice. Furthermore, this poses another consideration of assigning a diagnosis: the lack of diagnostic accuracy across clinicians. Several studies have demonstrated misdiagnosis of mental health disorders are common within children (Bruchmüller et al., 2012; Liang et al., 2016; Merten, Cwik, Margraf, & Schneider, 2017; Pastor & Reuben, 2005). Within other populations, since DSM-III, it has been found repeatedly that agreement between clinical and criterion-based diagnoses (e.g., DSM-5, ICD-10) remain low (Blashfield & Herkov, 1996; Morey & Benson, 2016).

In assessment literature for children, when completing standardized measures on the child's symptoms, studies have continuously exposed inconsistencies among informants (e.g., child, caregivers, teachers); in turn, this can lead to inconsistent diagnoses between clinicians (Beidas et al., 2016; De Los Reyes & Kazdin, 2005; Sallis et al., 2019). Depending on the assessment approach a psychologist uses, it can increase the likelihood of misdiagnosis, inflated rates of prevalence, and low diagnostic reliability (Beidas et al., 2016; Comer & Kendall, 2004;

De Los Reyes & Kazdin, 2005; De Los Reyes & Langer, 2018; Morey & Benson, 2016). Some have recommended a recommended way to address these discrepancies is to employ the "or rule" (De Los Reyes & Kazdin, 2005, p. 504) particularly in the absence of EBA guidelines (Beidas et al., 2016). The "or rule" encourages the evaluator to consider all symptoms reported across measures by each informant to meet diagnostic criteria, thus treatment can be targeted towards current symptoms (De Los Reyes & Kazdin, 2005, p. 504). In contrast, the "and rule" decreases sensitivity and requires all informants to report the same symptoms to meet diagnostic criteria (Comer & Kendall, 2004, p. 884). However, there are limitations to using the "or rule." One consideration is over pathologizing from merging the reports (De Los Reyes & Kazdin, 2005). For this reason, several authors have strongly recommended developing informant-specific diagnostic classification systems for children (De Los Reyes & Kazdin, 2005; Offord et al., 1996).

These studies have exemplified despite changes in the DSM, both clinician factors (e.g., demographics) as well as client factors (e.g., demographics, symptomatology) can lead to misdiagnosis and low diagnostic reliability for children. This literature adds to the strong influence of a psychologist's decision-making as well as the need for more systematic approaches in assessment methods.

Assessment Criticisms

Alongside the limitations discussed above, other literature suggests additional critiques of assessment include psychological assessment is too costly to justify their use, some instruments lack tests validity, tests cannot be used effectively in multicultural contexts, some test content is outdated though still being used, and instruments are face value, thus are easily manipulated by

respondents who wish to engineer a particular outcome (Anastasi, 1985; Carlson & Geisinger, 2009; Sanders & Katz, 2013).

Across the assessment literature, such as forensic, educational, and behavior, several investigators suggest the need for standardized assessment procedures for the assessor to help aid in overuse or misuse of tests, as well as suggesting further research on when is the appropriate time to use testing (Knauss, 2001; Pope & Vaquez, 2007; Ready & Veague, 2014; Sanders & Katz, 2013). When discussing the future of testing over almost 75 years ago, Hunt (1946) suggested, at times, the use of clinical judgment within measures can reduce the validity and reliability of the measure itself. An illustration of this may look like a psychologist making modifications in the assessment administration that alters the constructs assessed or the comparability of scores due to a child's unique needs (Phillips, 1993). On the other hand, assessment *accommodations* are changes in testing format (e.g., response, setting, or timing) that are often confused with assessment *modifications*. In other words, accommodations level the playing field of testing, while modifications change the field you are playing on; yet, still both require clinical decision-making. Specifically, assessment accommodations, such as adapting the setting for a child with a disability or changes in the method of response, do not modify in any significant way what the test measures or the comparability of scores (Phillips, 1993). For example, a psychologist may encourage an anxious child to answer nonverbally by pointing to their answers.

These changes happen frequently in assessments with children; however, the evaluator must determine whether the change is an accommodation or a modification (Cashel, 2002). Thus, the psychologists must employ their clinical judgment before administering instruments, as their assessment approach (e.g., methods and conceptualizations) when choosing then

interpreting tests, can leading to the misuse by misrepresentation of a measure (Bersoff et al., 2011; Cashel, 2002; Hunt, 1946; Pope & Vaquez, 2007; Sanders & Katz, 2013).

As mentioned earlier in this chapter, within the field of clinical psychology, an overwhelming amount of literature has been devoted to the development of tests, use of specific instruments, interpretation methods, and statistical procedures (Achenbach, 2005; Getz, 2011; Hunsley & Mash, 2008; Minton, 1988; Prinstein et al., 2019). In other words, much of the current research, standards, and guidelines have narrowly focused on psychological tests (e.g., subscales, validity, norms) instead of the usefulness of assessment procedures (Fernández-Ballesteros et al., 2001; Hunsley & Mash, 2018). Although easier to study, psychological tests (e.g., instruments, measures) make up only one piece of the assessment process (Fernández-Ballesteros et al., 2001). Thus, the plethora of empirical data available for tests (e.g., measures, subscales) is an essential element to a larger body of assessment research and should not be considered a replacement for studies, guidelines, and standards regarding psychological assessments (Hunsley & Mash, 2014).

Within multiple works, Hunsley and Mash (2008, 2014, 2018) propose the focus and abundance of literature across these areas have created a gap in literature exploring the specific methods and activities of psychological assessment. The authors further hypothesize clinical psychologists have become "complacent" and often expect their peers are capable of competently performing psychological assessments; therefore, "the usefulness of contemporary assessment instruments and practices has rarely been questioned" (Hunsley & Mash, 2014, p. 77). Hunsley and Mash expand on previous studies (Fernández-Ballesteros et al., 2001; Fletcher, Francis, Morris, & Lyon, 2005; Russo, 2015), arguing there is a little research available to support clinical utility of conducting assessments or contend the efficacy for psychological

evaluations to progress treatment outcomes. Moreover, several instruments that are reported to be used regularly in assessment practices have lack empirical evidence or have inconsistent findings (Fletcher et al., 2005; Norcross, 1991). Consequently, the lack of studies exploring the utility of assessment creates greater barriers within the field. For example, in terms of qualifying the psychologist's expended effort, economic expenses, and clinical need for assessments, many believe the field is unable to produce supporting empirical data to address these areas (Hunsley & Mash, 2014; Meyer et al., 2001). In practice, many articles have found between 40 and 68% of child psychologists reported reimbursement issues as the prevalent barrier to conducting psychological assessment (Cashel, 2002; Eisman et al., 2000; Wright et al., 2017), further suggesting that the paucity of current empirical support may have an impact on low reimbursement rates for assessments (e.g., Eisman et al., 2000; Meyer et al., 2001; Wright et al., 2017). Another consideration is that constrictions of managed care have been correlated with misdiagnosis and mismatched treatment (Miller & Luft, 1997). Given that upwards of 72% of psychologists have reported adjusting their test utilization as a direct result of managed care (Piotrowski, Belter, & Keller, 1998), delineates the influence of managed care within the clinical decision-making process. Specifically, psychologists have admitted to tailoring their battery to fewer tests or not administering tests due to managed care or third-party payors (Dhami & Mumpower, 2018; Eisman et al., 2000; Meyer et al., 2001; Wright et al., 2017). Within another study, psychologists endorsed such significant limitations by managed care policies on their ability to conduct tests; in turn, they openly acknowledged their diagnostic decision-making was inhibited (Wright et al., 2017). Therefore, research offering supporting data for the clinical utility of assessment remains to be conducted. The lack of studies leaves psychologists unable to provide scientific data in response to "requests from program administrators, third-party payers,

or clients to justify the professional time and costs associated with psychological assessment" (Hunsley & Mash, 2014, p. 77).

Studying psychological assessments is a formidable task given the complexities of the process, such as integrating test data from various measures, background information, and collateral data into a neatly wrapped package (i.e., report) (Fernández-Ballesteros et al., 2001; Mash & Hunsley, 2005). Studies citing the difficulties in creating guidelines and exploring factors within the assessment process are well documented (Fernández-Ballesteros et al., 2001; Hunsley & Mash, 2014; Jensen-Doss & Hawley, 2010; Krishnamurthy et al., 2004; Wolraich et al., 2019). To explain this literature gap, numerous frequently cited works describe psychological assessment methods as a constant decision-making process (APA, 2013c; Hunsley & Mash, 2014; Kicklighter et al., 2018; Krishnamurthy et al., 2004; Norcross, 1991; Spengler, Dixon, Strohmer, & Shivy, 1995).

Currently, ethics (APA, 2017a), guidelines (Krishnamurthy et al., 2004), and training (Garb, 2005; Lovitt, 1974; Spengler et al., 1995) all encourage the use of clinical judgement in regards to assessment practices. Yet, Kazdin (2008) pointed out as a field, psychologists have just accepted statements involving meeting patients' needs and applying clinical judgment despite several concerns, including a lack of solid research available to aid psychologists in this process. Further, a large number of existing studies in the broader literature (e.g., Bruchmüller, Margraf, & Schneider, 2012; Leadbeater et al., 2012; Pastor & Reuben, 2005; Richardson et al., 2015) have echoed the need for research investigating the use of clinical judgement and decision-making to guide treatment. Kazdin (2005) explicitly noted the necessity of a methodical application of clinical judgement as well as the use of systematic measurement to evaluate the impact of psychological services. Another suggestion regarding the lack of empirical support

can be explained by the divide between research and clinical practice (Kazdin, 2008). As Kazdin (2008) highlights, scientific data is missing in two essential areas: "reliability in decision-making (consistency within therapists over time and among different therapists at a given point in time) and validity (that the decision makes a difference in the outcome when compared with a less flexible algorithm or an alternative case formulation model)" (p. 149).

Kazdin's findings (2008) supplement other studies that have revealed decision-making errors within clinical practice and assessment procedures (Bruchmüller et al., 2012). In terms of misdiagnosis, Bruchmüller et al. (2012) suggested that the diagnostic decisions of psychologists were influenced by heuristic errors (e.g., representativeness, discussed further in chapter four). A systematic review of the literature conducted by Merten et al. (2017) explored overdiagnosis of mental disorders in children and adolescents. Within the study, Merten et al. (2017) found several factors that influenced misdiagnosis including the use of heuristics biases across assessment information (e.g., test data, informant information) in addition to limitations set by health care settings such assigning a diagnosis to justify treatment.

Managed Care Implications

According to Krishnamurthy et al. (2004), within clinical psychology, many view assessments as a distinctive and essential feature of the field, specifically within the context of multidisciplinary child healthcare settings and services. As discussed, assessment is not without fault and has experienced criticism, especially from a managed care perspective (Eisman et al., 2000) to the extent that health care administrators differentially limit reimbursement for psychological tests in comparison to medical tests (Meyer et al., 2001). Many argue that such actions are not justifiable based on empirical evidence, including test validity (Cashel, 2002; Meyer et al., 2001; Oakland et al., 2001; Piotrowski, 1999).

Child practitioners employed within a variety of settings, such as independent practice, hospitals, outpatients, and school-based settings throughout the United States, consistently reported significant limitations in psychological testing due to managed-care policies (Bersoff et al., 2011; Cashel, 2002). A proposed suggestion to counteract managed care constraints is to bolster empirical data that demonstrates psychological tests are a central factor across treatments, in turn validating assessments is not optional or a supplementary service (Camara, Nathan, & Puente, 2000; Meyer et al., 2001; Wright et al., 2017). In support, APA and several authors recommend that clinical psychologists will need to accomplish two main objectives: studies will need to clearly demonstrate that conducting assessments for treatment-planning purposes both significantly enhances diagnostic accuracy as well as reduces the cost or length of treatment (Cashel, 2002; Eisman et al., 2000; Hunsley & Mash, 2008; Meyer et al., 2001; Richardson et al., 2015).

In terms of training methods, Ready and Veague (2014) found an increased focus on assessment within graduate programs with similar forms of assessment instruction across training models. However, other researchers (e.g., Belter & Piotrowski, 2001) suggested that assessment competency is better understood as "contextually specific" (Krishnamurthy et al., 2004, p. 730) explaining in regards assessment practice, as tests are taught and used within a specific cultural and social context. In alignment with Watkins' (1991) explanation for stability in testing practices (i.e., the reinforcement cycle), Belter and Piotrowski (2001) and other studies (e.g., Cashel, 2002) have found managed care and reimbursement restrictions have not significantly altered assessment training practices in graduate programs that, yet again, affirms the divide between clinical practice and graduate training as well as the need for continuity. Specifically, greater understanding of assessment methods within clinical application to help

unify research, graduate, practical, and post-doctoral training for clinical child psychologists. Without a framework for current best practice guidelines in place or diagnostic criteria requiring psychometric evidence, in practice, psychologists' clinical rationale and decision-making remain largely affected by economics.

Managed care influences on psychological assessment. It has been documented that several managed care teams who are responsible for authorizations and creating protocols for psychological assessments, are not psychologists (Eisman et al., 2000). When psychologists are a part of the teams, they have been deemed lacking competency in psychological assessment protocols (Eisman et al., 2000, p. 132). Further, some studies indicated that the individuals responsible for authorizing psychological evaluations often had inadequate knowledge of assessment instruments or clinical decision-making (Cashel, 2002; Eisman et al., 2000), thereby leading to denials in reimbursement for the services. The increase of denials and limitations set by managed care organizations deprives the provider of an essential source of support both clinically and ethically, such as objective data to defend any malpractice claims related to adverse treatment or diagnostic decisions (Pope & Vaquez, 2007). When deliberating legal ramifications, managed care organizations have responded that they have no ethical or legal liability related to assessment or treatment denials because making clinical decisions is out of their scope; instead, they are responsible only for managing the patient's benefits (Eisman et al., 2000).

Many child psychologists are limited in practice by managed care organizations, as mentioned (Cashel, 2002). Some studies suggest the restrictions of managed care are heavily influencing a psychologist's determinations about which tests are acceptable or optimal during assessments, rather than their clinical judgment (Eisman et al., 2000; Piotrowski, 1999; Pope &

Vaquez, 2007). For the purposes of illustration, one managed care company, Beacon Health Options has free, easily obtainable, online accessible policies and procedures for their clinical criteria in various areas of healthcare. In policy and procedure "3.1101 Psychological/ Neuropsychological Testing (Child/Adolescent)" last reviewed in November 2015 regarding approval for ADHD testing, the policy states:

> The expectation is that the diagnosis of ADHD can in most instances be made on the basis of the most current version of the DSM-5 criteria alone and such diagnosis does not necessarily require psychological testing. Extended testing for ADHD is not authorized prior to the provision of a thorough evaluation, which has included developmental history of symptoms and administration of rating scales. (Collection of rating scale data from additional objective sources is highly preferred.) One unit of 90791 and one or more units of 90834 is usually authorized for the rating scale review and feedback. In the vast majority of cases, the diagnostic question is answered with this work-up. If this initial work-up shows indications for ADHD, referral to a physician would be appropriate. (Beacon Health Options, 2015, p. 2)

Interestingly, this managed care company specifically references the lack of required psychometric evidence as DSM criteria for ADHD. In turn, the policy seems to use this as rationale of psychological testing to be a supplementary process when assigning a diagnosis. In terms of limitations, the policy alludes to conducting a "thorough evaluation" yet does not reference evidenced-based assessment practices for ADHD. Furthermore, the managed care company does not have easily procurable information regarding how these policies were created or what empirical support guided the development.

Consequently, another factor within the decision-making process before conducting assessments emerges. Eisman et al. (2000) argue decisions regarding which tests are appropriate is a clinical issue and should not be made according to a company protocol. This author suggests that using clinical judgment to make decisions within the limitations of managed care has become an area of competence for many practicing psychologists. Clinicians must maintain a

sound rationale to adapt the protocol that fits the need of the client on the basis of accepted clinical indicators within the realm of managed care. Thus, questions that then naturally arise include: could more reliable results be produced if diagnostic criteria were inclusive of psychometric evidence and evidence based guidelines were in place to support decision-making preceding test administration?

Summary

To conclude, testing can provide details of symptomatology and help clarify differential diagnoses (Meyer et al., 2001; Sattler, 1988). The assessment process, including empathic feedback, is viewed and can be a form of treatment and therapeutic intervention (Leadbeater et al., 2012; Meyer et al., 2001; Youngstrom et al., 2017). Simultaneously, testing can help guide differential treatment needs for both medical and mental health conditions, as well as accurate monitoring of treatment over time (Meyer et al., 2001). In other words, psychological testing assists psychologists in identifying and understanding the nature of a child's difficulties, which consequently aids their ability to develop the best treatment plan for that child.

Although there is substantial evidence supporting the use of testing in practice, there is a limited amount of research linking these same benefits to the assessment process. One explanation for this is that studying psychological assessments is a challenging task, given the complexities of the process (Fernández-Ballesteros et al., 2001; Mash & Hunsley, 2005). Subsequently, many articles have begun to describe psychological assessment methods as a persistent, dynamic, decision-making process (APA, 2013c; Hunsley & Mash, 2014; Kicklighter et al., 2018; Krishnamurthy et al., 2004; Norcross, 1991; Spengler et al., 1995).

Another concern regarding current testing research is the absence of guidelines (Fernández-Ballesteros et al., 2001; Hunsley & Mash, 2014; Jensen-Doss & Hawley, 2010;

Krishnamurthy et al., 2004; Wolraich et al., 2019). While APA has published *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014), presently, the standards only offer general criteria to be followed in developing and using psychological instruments. One noteworthy limitation is that the APA's standards do not indicate the minimum psychometric values necessary to indicate that an instrument meets the benchmark of being reliable or accurate. Moreover, most of the current diagnostic criteria (e.g., DSM, ICD) does not include psychometric evidence. The proposed implications of these limitations include the lack of empirical data to support assessment methods and affect due to managed care organizations, in addition to diminished and inconsistent decision-making regarding diagnostics, assessment methods, and consequently treatment-planning among psychologists.

CHAPTER IV: CURRENT APPROACHES

Clinical Decision-Making in Clinical Psychology

As discussed in chapter one, clinical decision-making can be understood as the act of applying reasoning with training, scientific theory, practical experiences, patient reports, and a level of uncertainty (APA, 2017a; Bellman & Zadeh, 1970; Hammond et al., 2006; Knauss, 2001; National Academies of Sciences, Engineering, & Medicine et al., 2015; Smith et al., 2008; Wills & Holmes-Rovner, 2006). The evidence-based practice in psychology (EBPP) movement has established the basis of criteria in several areas (APA, 2006; Frick, 2007), although it has not explicitly defined clinical judgment in practice. Some fields of healthcare encourage using evidence-based practices (EBP) as the gold standard for guiding professional decisions (Magnavita, 2016; Podgorelec et al., 2002).

Over time, some preliminary literature has developed on decision-making within mental health practices. For example, Magnavita (2016) outlined five overarching pillars for effective decision-making in mental health: (a) access to high-quality empirical evidence, (b) developing clinical expertise, (c) using sound theoretical constructs, (d) including ethical considerations, and (e) foundation in decision theory (p. 9). Magnavita later emphasizes the value of understanding the clinical decision-making process, stating:

> Decision-making influences every aspect of clinical practice and is increasingly important for behavioral and mental health clinicians, as well as all health care providers, because of the inherent uncertainty in many aspects of clinical science. Many decisions are made that do not derive from an empirical evidence base, and that necessitate a comfort with uncertainty. (p. 16)

Many others agree with this stance, although also recommend that it is of higher value for the clinician to be aware of their clinical decision-making process in order to identify potential biases throughout their choices (National Academies of Sciences, Engineering, & Medicine et

al., 2015; Suhr, 2015; Watkins, 2009). Magnavita (2016) later notes that clinical mental health is far behind on developing decision-making guidelines when compared to the progress of current medical decision-making models.

Limitations. Other authors (e.g., Smith et al., 2008) have found similar findings as Magnavita; however, several argue that literature investigating psychologists' clinical decision-making is sparse (Fernández-Ballesteros et al., 2001; Hunsley & Mash, 2008; Kazdin, 2008; Mash & Hunsley, 2005; Prinstein et al., 2019). Specifically, in previous literature, components of clinical decisions are frequently only mentioned secondary to the focus of the study. For instance, across articles (e.g., APA, 2006; Bornstein, 2017; Hunsley & Mash, 2007), EBPP is frequently defined as "clinical practice that is informed by evidence about interventions, clinical expertise, and patient needs, values, and preferences and their integration in decision-making about individual care" (Kazdin, 2008, p. 147). However, within many of these studies, clinical judgment and decision-making are not explicitly defined. In regards to this gap in the literature, Kazdin noted there are currently no formal or replicable procedures for psychologists to know how to apply clinical decision-making, nor is a framework available to decide in what amount and sequence should treatment be delivered to families (Kazdin, 2008, p. 149). In turn, many studies implore for literature exploring "clinical decision-making, judgment, and expertise as a guide to individual treatment" (Kazdin, 2008, p. 149; see also Mash & Hunsley, 2005; Morey & Benson, 2016). These studies demonstrate, within the field of psychology, there is a need for research to specifically explore the reliability and validity of psychologists' clinical judgment and decision-making (Hunsley & Mash, 2014; Kazdin, 2005, 2008; Watkins, 2009). Conversely, within the medical field, clinical judgment and decision-making has been increasingly investigated (Croskerry, 2009; Croskerry & Norman, 2008; Engel, 1977; Makary & Daniel,

2016; Makoul & Clayman, 2006; Reyna, 2008). The medical fields' prior research on clinical decision-making provides a framework to apply in areas of the field of psychology. Given these are adjacent applications of clinical decision-making, for the purpose of this project, the relationship of medical decision-making as it pertains to psychology and psychological assessment will be further explored within this chapter.

Value of Clinical Decision-making

When a family expresses concerns and seeks recommendations from a gateway or mental health provider, they often assume the provider's clinical judgment of their condition is accurate (Owens et al., 2002; Stiffman et al., 2004). However, as discussed in chapter two, findings reveal between 70% and 80% of children who meet criteria for a mental health diagnosis are undiagnosed, untreated, and do not receive any form of services (Castellanos-Ryan et al., 2016; Eisman et al., 2000; The Mental Health Foundation, 2016). Additionally, upwards of 96% of formal gateway providers (e.g., pediatricians) fail to identify mental health symptoms and diagnoses (Stiffman et al., 2004). Despite the large discrepancy in symptom identification of the referral sources, much of the broad assessment literature, such as report writing (Wiener & Costaris, 2012), ethics (Smith, 2003), and interpretation guidelines (Ivnik et al., 2000), put a heavy focus on answering the referral question solely. For case illustration, envision a child presenting in a psychologist's office after finally receiving a referral for a conduct disorder assessment. Some of the many considerations for the child psychologists to consider include the large discrepancy in symptom identification of the referral sources mixed with over 32,000 combinations of symptoms presentations (Perepletchikova & Kazdin, 2005) that can meet criteria for this disorder. When contemplating just these two factors, it poses the ethical question: should solely answering the referral question be the main objective of the assessment

or should it be answering the symptomology? In this case example, the referral question is does the child have conduct disorder: if the psychologist conducting this evaluation is only answering and building their test battery around the referral question and does no additional testing, does this violate their code of ethics such as, "professional and scientific responsibilities" (APA, 2017a, Principle B)? Another comparison is being referred to a primary care physician with concerns of high blood pressure, yet the physician would only conduct tests to inform on high blood pressure, report the patient does not have high blood pressure, and then not explore the cause of the patient's symptoms.

As discussed, some frequent referrals for children include ADHD, autism, and anxiety; however, these all can have similar presentations and comorbidities (Merikangas et al., 2010). Consequently, classifying a mental health diagnosis is particularly difficult with children in that commonly occurring symptoms can have multiple causes and can mimic several diagnoses, thus increasing the likelihood of diagnostic mistakes (Liang et al., 2016; Merten et al., 2017). Heuristic or cognitive errors have also shown to be a factor in the misdiagnosis of several mental health disorders in children (Merten et al., 2017). Notably, studies have demonstrated that provider's clinical decisions are strongly influenced by heuristic errors, instead of following scientific data such as diagnostic criteria (Bruchmüller et al., 2012). This problem is not unique to clinical psychology. A study by Johns Hopkins found that more than 250,000 deaths per year are due to medical errors in the U.S., making medical errors the third leading cause of death in the U.S. (Makary & Daniel, 2016), while older works, such as by the Institute of Medicine (2000), have illustrated that medical errors occur due to poor decision-making or human errors. High occurrences of poor patient outcomes, with some resulting in deaths due to medical errors, have ignited a surge in medical studies to better understand and create systematic approaches for

medical decision-making (Institute of Medicine, 2000; Makary & Daniel, 2016; Reyna, 2008; Rylander & Guerrasio, 2015).

Understanding Decision-Making

The origins of decision-making are far-reaching and have been historically pursued. Numerous pioneers constructed a foundation, such as Aristotle's systematic empiricism, Occam's ideas of logic, Francis Bacon's inductive reasoning, and the renowned modern philosopher, Descartes publishing Discourse on Method in 1637 (Buchanan & O'Connell, 2006), while many within the healthcare field make connections from Simon's (1959) findings as he connected the effect and satisfaction of motivating forces. Simon's (1959) theory regarding decision-making is comparable to the "good enough notion," meaning an individual works until a result is good enough, rather than the ideal or perfect outcome. This theory is valuable as it proposed decision-making to be profoundly affected by an individual's perception, or in other words, their internal and external forces such as demographics, morals, and setting.

In terms of clinical judgment and decision-making, several researchers have pursued an advanced understanding of the phenomena. In turn, diverse expressions and paradigms fundamentally to describe similar experiences have emerged. As might be expected, across disciplines, there is no standard definition of clinical judgment or decision-making. However, within existing philosophies and constructs, such as rational and intuitive approaches, the power of experience and contextual elements as well as informal and formal factors appear (Charles et al., 1999; Makoul & Clayman, 2006; Splinter, 2010; Whaley, 2001). In Hamm's (1988) earlier works studying clinical judgment, it was found that medical professionals with more experience made clinical decisions without adhering to a systematic method, despite being taught formal methods in training to decision-making (Dhami & Mumpower, 2018). Hamm later proposed the

cognitive continuum theory suggesting that various forms of cognition (e.g., intuitive, analytical, common sense) create a spectrum with intuitive and analytical processing at opposite ends (Dhami & Mumpower, 2018). Further, the theory proposes variations in reasoning (e.g., cognitive control, awareness of cognitive ability) have a role along the continuum, and cognitive performance is controlled by matching the task properties and mode of cognition (Dhami & Mumpower, 2018). In other words, in terms of clinical decision-making, Hamm proposed that intuition exists during uncertainty and can bridge the gap in clinical judgment. To put it another way, when making a decision, an individual's intuition consists of clinical experiences, training, and knowledge thus promoting the ability to predict outcomes effectively (Dhami & Mumpower, 2018). Used collectively, both analytic and intuition facilitates clinical decision-making across a continuum of a professional's experiences, training, contextual factors, and situations (Croskerry, 2009; Dhami & Mumpower, 2018; Reyna, 2008). These ideologies tend to be the closest fit to the current definition of EBP within the field psychology.

Decision-Making Frameworks and Healthcare

Within healthcare, clinical decision-making studies often align with one decision-making paradigm then utilize the model's terms throughout. However, each paradigm tends to conceptualize constructs and definitions of the decision-making process differently (Gigerenzer & Gaissmaier, 2011; Klein, 2008, 2015; Lilienfeld & Lynn, 2015). For example, *intuition* is described in opposing ways across some models, such as in (a) Naturalistic decision-making (NDM) that suggests intuition is grounded on several patterns acquired through experiences and results in different forms of tacit knowledge, (b) Fast and Frugal Heuristics (FFH) define intuition in relation to general-purpose heuristics, and (c) Heuristics and Biases (HB) model that views intuitions as a significant source of bias and error (Klein, 2015).

Heuristics. Within many of the heuristics and biases models, studies have repeatedly revealed humans, including healthcare professionals, rely on heuristics to solve problems and make judgments quickly and efficiently despite education, training, scientific data, and known strategies (Croskerry, 2009; Hammond et al., 2006; Lilienfeld & Lynn, 2015; Rylander & Guerrasio, 2015). Although heuristics are often helpful and efficient, one common explanation for heuristic errors is that purely logic-based choices take a substantial amount of time given the process is comprised of evaluating endless factors, alternatives, risks, and benefits (Croskerry, 2009; Garb, 1998; Gigerenzer & Gaissmaier, 2011; Rylander & Guerrasio, 2015).

Classical decision making. Several healthcare studies (e.g., nursing, medicine) apply the classical decision-making (CDM) model that implies a rational or scientific inquiry methodology. Specifically, the decision-maker considers all possible options and consequences, then initiates a formal selection process until the ideal alternative is selected (Shaban, 2005). However, the limitations of this model are that it assumes the decision-maker is aware of all possibilities, has complete certainty, is always in an ideal, controlled setting in which the problem is clearly defined, and all possible actionable alternatives are known (Klein, 2015; Klein, Calderwood, & Clinton-Cirocco, 2010; Shaban, 2005). Given the CDM does not account for chaotic, real-world, clinical decision-making, other studies have found decision-makers with more practical experience do not use this approach (Klein et al., 2010). Conversely, studies found challenging decisions involve multifaceted problems and vague objectives that cannot be identified prior; instead, objectives become refined as they are pursued (Klein, 2008; Klein et al., 2010; Pliske & Klein, 2003). These findings challenged the work of earlier researchers and pushed the movement of Naturalistic Decision Making (NDM) (Klein, 2015; Klein et al., 2010; Shaban, 2005).

Naturalistic decision making. Several researchers across disciplines (e.g., businesses, medicine, military) attempted to study how individuals made complex judgments under demanding conditions such as high stakes, time, uncertainty, instability, and ambiguously defined problems (Klein, 2008; Shaban, 2005; Simon, 1959). Klein's (2015) work within the NDM paradigm gained attention as it emphasized a more dynamic, behavioral approach in decision-making that acknowledges human experiences (e.g., cognitive limitations in bounded rationality) and contextual factors. Namely, unlike the classical approach, the naturalistic model recognizes that many settings are uncontrolled, changing, or may have time constraints (Klein, 2008, 2015). Further, the role of time "is a significant factor, requiring assessment, interpretation, and assimilation of multiple data from multiple sources, often in high stakes settings" (Shaban, 2005, p. 2; see also Klein et al., 2010). In clinical practice, when using the NDM model, it is expected that the problem will not be clearly defined, the clinician will rely on their experiences and will have limited knowledge of all alternatives and risks, although it presumes that the clinician will respond adeptly based on the available data and their prior experiences (Klein, 2008, 2015).

Interestingly, expert decision-makers select options based on their *tacit knowledge* or knowledge gained from subjective experience that is more difficult to quantify (Klein, 2015). Experts, such as skilled, experienced individuals within an area (e.g., doctors, nurses, soldiers), make decisions without being aware of criteria or procedures (Klein & Wright, 2016). Instead, their decision-making process becomes highly adaptive and proficient when selecting an option (Kicklighter et al., 2018; Klein et al., 2010). In other words, experts hone into the problem or task and quickly discount irrelevant considerations of a wide scope of unlikely alternative diagnoses or explanations (Klein & Hoffman, 1992). However, before NDM studies, it was

thought that expertise was dependent on *explicit knowledge* such as learning procedures, guidelines, and rules (Klein et al., 2010; Klein & Hoffman, 1992; Pliske & Klein, 2003). Meaning, at the start of a task or problem, experts tend to use their own mental models to efficiently sift through an abundance of material quickly and determine what is considered data and what is extraneous (Klein, 2015). Experts then relentlessly engage in this process while adjusting their model as new material is obtained (Pliske & Klein, 2003).

Two important findings that have surfaced from NDM research are ways to reduce uncertainty and improve performance (Klein, 2008). Early literature in clinical-decision making suggested identifying assumptions and the need to gather as much information possible (Klein & Wright, 2016). However, authors within NDM have shown listing assumptions are not helpful, as often the decision-maker is unaware of their biases and could never self-generate a list (Klein, 2015; Klein & Wright, 2016). Additionally, studies have demonstrated that increased uncertainty and diminished decision-making abilities occur both when too much data is attained (e.g., inadequate framing of information) as well as during the lack of data (Klein & Wright, 2016). In a more recent work, Klein and Wright (2016) also recommended macro cognitive models as tools that can be used in cognitively challenging activities to enhance performance.

Similarly, psychological assessment is understood to be "a decision-making process that includes various tasks, operations, and actions (conducted in a given sequence), in order to answer the client's question, and that requires basic psychological knowledge and professional abilities" (Fernández-Ballesteros et al., 2001, p. 188). Given the parallel of considerations between conducting psychological assessment and naturalistic decision-making literature, this paradigm can act as a framework to advance understanding of child clinical psychologists' decision-making and judgment. Furthermore, Klein and Wright (2016) also suggest the

ShadowBox approach, which is another macro cognitive tool. This scenario-based method of training is used to facilitate novices' ability to tap into experts' tacit knowledge without experts being present. The training provides several demanding, thought-provoking, practical situations with decision-making points. First a panel of subject-matter experts provide their rankings and rationale for their choices, which are then synthesized. A trainee goes through the scenario, ranking the best options about which pieces of data to gather, areas to monitor, outcomes to prioritize, and course action to take. After submitting their rankings, the trainee will have access to the expert's responses and gain a look through their mental models. Klein's work suggests this reflective process further promotes the development of intuition and insight, thus fostering the progression of novice to expert thinking (Klein, Hintze, & Saab, 2013; Klein & Wright, 2016). Because of the deficiencies in child psychologist's educational training for child assessment and decision-making, Klein's ShadowBox approach could be one possible way to aid graduate students in developing their clinical judgment. Another consideration is that there is a strong support for explicit knowledge being the basis for problem solving, such as in evidence-based practices (Watkins, 2009). On the contrary, NDM research has suggested in times of uncertainty, experts rely more on tacit knowledge and then reference explicit knowledge (Klein, Calderwood, & Clinton-Cirocco, 2010).

Clinical Decision-Making in Mental Health

Ethical standards, guidelines, and competencies, as well as several articles imply psychologists are expected to use clinical judgment and decision-making (APA, 2013c, 2017a; Garb, 1998, 2013; Hogarth, 2010; Redelmeier, Ferris, Tu, Hux, & Schull, 2001; Spengler et al., 1995). Across the literature, the clinical decision-making process is described as a dynamic process involving consideration of evolving factors to achieve an optimal outcome (Croskerry,

2009; Kazdin, 2008; Magnavita, 2016; Smith et al., 2008; Youngstrom, Choukas-Bradley, Calhoun, & Jensen-Doss, 2015). Across healthcare disciplines, most early studies as well as current works, have identified main factors correlated with clinical judgment and decision-making. Decision-making factors that apply to the psychological assessment process can be organized into both macro and micro-level influences (see Figure 1).

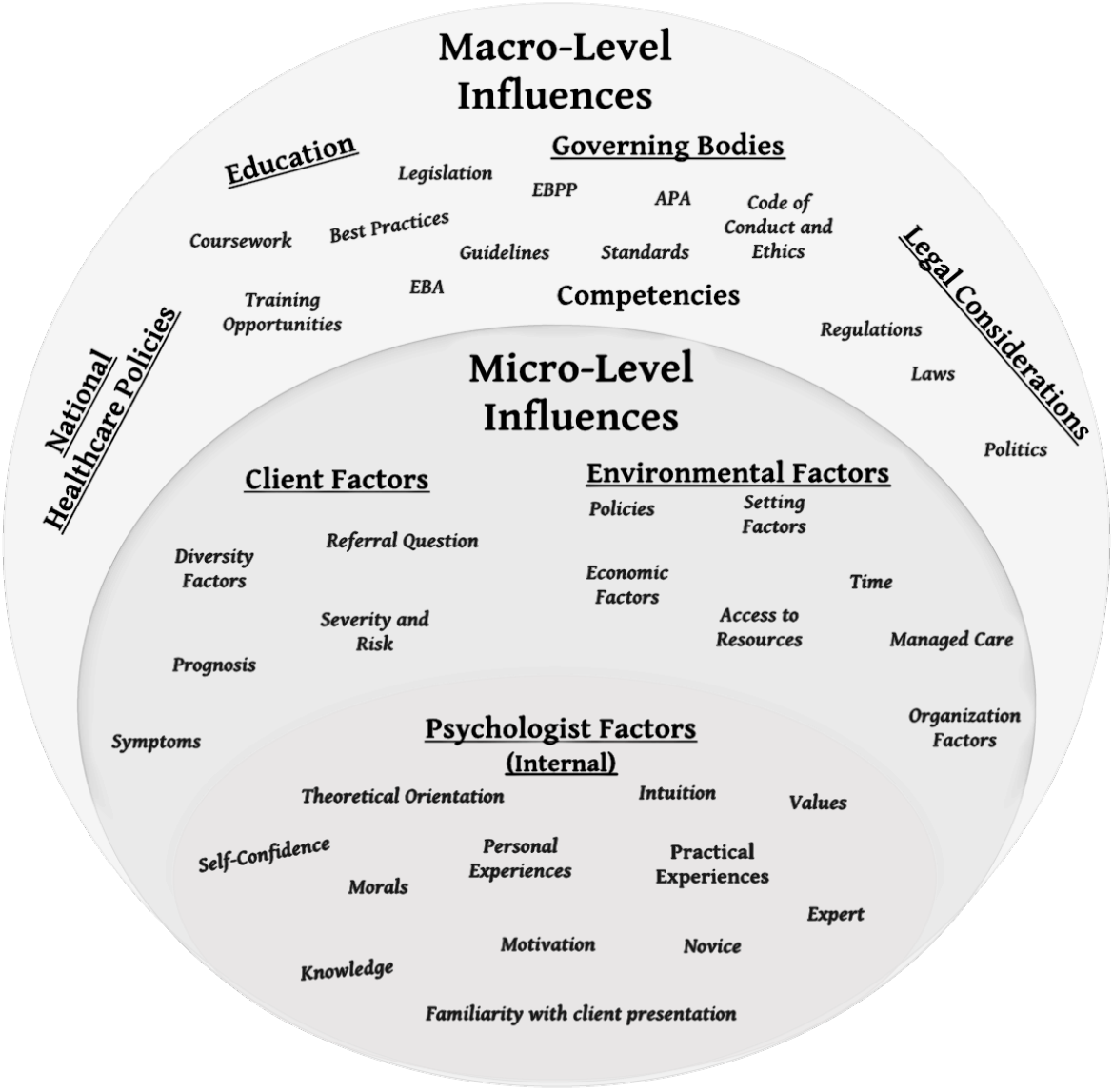

Figure 1. Illustration of Macro- and Micro-level Influences on Clinical Decision-Making (Bui, 2020)

On the macro level, emerging factors include (a) national healthcare policies (Legare, Ratte, Gravel, & Graham, 2008); (b) education (e.g., curriculum, training opportunities) (Poland, Thurlow, Ysseldyke, & Mirkin, 1982); (c) professional governing bodies (e.g., ethics, guidelines, standards) (APA, 2017a; Jensen-Doss & Hawley, 2011; Kazdin, 2005); and (d) legal considerations (e.g., law) (APA, 2006, 2017a). Micro influences include (a) environmental (Garb, 2013); (b) internal and external clinician factors (Magnavita, 2016); and (c) client factors (e.g., demographics) (Bornstein, 2017; Garb, 1997).

Within the micro-level, clinician factors are one of the heavily researched areas (Garb, 2005; Magnavita, 2016; Piotrowski, 1999). Clinician factors can further be divided into subgroups including attitudes, values, morals (Hogarth, 2010; Pope & Vaquez, 2007); self-confidence, knowledge, skills, (Hagbaghery, Salsali, & Ahmadi, 2004; Hammond et al., 2006; Riggin & Lack, 2018); and practical experiences (e.g., practice site) (Kazdin, 2008, 2015; Lovitt, 1974). Another emerging factor includes intuition (Hogarth, 2010; Klein, 2015). Intuition is frequently mentioned in articles; however, this phenomenon is difficult to study independently thus creating a greater challenge in integrating into research on decision-making (Hogarth, 2010). Additional works by Garb (1997, 1998, 2013) have established that patient factors strongly influence clinical judgment. Ultimately, race, socioeconomic status, and gender were found to significantly affect clinical decision-making (Garb, 1997). In comparison to gender and socioeconomic status, studies exploring clinical judgment suggests that race bias is more pervasive (Brondolo, 2015; Dana, 1998; Garb, 1997, 2013). Further studies reveal that these biases have led to poor clinical decision-making, such as failure to collect information or overlooking diagnostic exclusion criteria (Bruchmüller et al., 2012; Garb, 1997, 1998; Merten et al., 2017). Garb (1997, 2013) recommended to help combat these biases, clinicians must be

mindful when their biases are likely to occur, follow guidelines and adhere to diagnostic criteria, and use statistical prediction rules (e.g., base rates).

When examining the systemic power of national healthcare policies, the influences of managed care organizations contribute to the clinical decision-making on the macro-level. Moreover, managed care can also serve as a micro-level factor, such as when psychologists are impacted by organizations limiting quality of care. Studies have found many psychologists alter their diagnostic techniques and therapies due managed care influences (Wright et al., 2017). For illustration, consider that after constructing a unique, clinically sound test battery for a child, the evaluating psychologists receives a pre-authorization form from an insurance company. Given the new circumstances, the psychologist may be compelled to make adjustments based on the authorization criteria (e.g., economics or pragmatics) rather adhering to their clinical judgment or best practices.

Taking a broader view, healthcare costs have gradually increased worldwide and the influences are widespread across disciplines (Jacobs & Fischer, 2012). In the United States, the National Health Expenditure Accounts (NHEA) reported healthcare costs account for 17.7% of the nation's Gross Domestic Product (U.S. Centers for Medicare and Medicaid Services, 2019). In 2018, American spending on healthcare increased 4.6% in 2018, which amounted to $3.6 trillion or $11,172 per person (U.S. Centers for Medicare and Medicaid Services, 2019). In response, articles have discussed comprehensive cost-control is needed for healthcare clinics (Jacobs & Fischer, 2012). However, these determinations are usually performed by clinic professionals or organizational administrators (Jacobs & Fischer, 2012) rather than the healthcare providers. For better patient outcomes, Jacobs and Fischer (2012) argued that providers need to

be a part of the equation because they are better equipped to make decisions regarding the best practices for care. The authors further explained that providers,

> Should accept the limitation of resources and take responsibility to improve their clinical cost-reimbursement ratio. To achieve this, they will need basic education in clinic management to control and adjust costs and reimbursement, without impacting professional quality of care. Rational use of diagnostics and therapy should be implemented and frequently verified. (Jacobs & Fischer, 2012, p. 632)

Further, they provide steps and guidelines to aid providers in becoming a part of the decision-making associated (see Jacobs & Fischer, 2012).

Clinical Decision-Making in Assessment

Prior research has reiterated the importance and application of decision-making in assessments; namely, APA (2013b) acknowledged that the assessment process involves a conceptual, methodical process of collecting reliable, pertinent data about an individual to arrive at a well-informed decision (see also APA, 2006). In the relevant literature of clinical decision-making in assessments and testing, similar expectations emerge. For instance, established criteria for sound decision-making within assessments include expertise or knowing: what information to gather; psychometrics; and reliable, valid, measures (APA, 2013b; Hunsley & Mash, 2008; Krishnamurthy et al., 2004; van Meter et al., 2014; Watkins, 2009).

These elements are obligatory and require much more than just memorizing criteria or cut off points. Regardless of how many aids or elements exist, they are not intended to replace a psychologist's knowledge, experience, intuition, or judgment (Pope & Vaquez, 2007). This means the clinician is expected to have proficiency as well as the clinical skills to accurately integrate the fundamental theories, science, and conceptualization behind each element (APA, 2019). In a like manner, clinicians are expected to utilize available resources to complement their skills and enhance the quality of care provided (APA, 2013c). Psychologists should also be

continually reflective and able to identify their decision-making errors or biases during every step of this process (APA, 2017a; Croskerry, 2015; Garb, 1997; Liang et al., 2016; Teresi & Jones, 2013).

Researchers have identified numerous heuristics, and several have implications in clinical decision-making (Hammond et al., 2006; Lilienfeld & Lynn, 2015; Teresi & Jones, 2013). The previous studies have been able to reveal the ways heuristics can both aid and detract from clinical decision-making. As discussed, many facets of assessment exist including the decision-making process and some concerns have been identified in decision-making with conducting assessments. Notably, without guidelines in the decision-making process, this becomes problematic in training continuity and more so in practice (Garb, 2013; Lovitt, 1974; Wright et al., 2017).

Common errors in decision-making. Child clinical psychologists likely utilize several hundred heuristics daily. An example of one is fidgeting during a clinical interview as a marker for nervousness or adding further supporting evidence to an anxiety diagnosis with adolescents. In that one sentence, there were several cognitive errors or biases, depending on the clinician's heuristics and past cases. Cognitive biases are inevitable; thus, understanding the heuristics informing the decision-making process during assessments is essential so that corrective action can be implemented (Watkins, 2009). Within psychology and other healthcare disciplines, cognitive biases have been correlated with diagnostic inaccuracies and treatment errors leading to mismanagement or inadequate utilization of resources, as well as substantial ethical dilemmas (Croskerry, 2015; Elwyn et al., 2012; Garb, 2005; Liang et al., 2016). In relation to psychological assessment practices, common biases were extracted and developed from the literature related to (a) clinical decision-making, (b) diagnostic outcomes, and (c) mental health.

Affective error. The tendency to convince oneself that the subjectively favored outcome is true instead of less appealing alternatives. For example, countertransference is a subset of an affective error (Gigerenzer & Gaissmaier, 2011; Lilienfeld & Lynn, 2015).

Anchoring bias. The tendency to maintain one's initial impression despite evidence pointing to the contrary (Rylander & Guerrasio, 2015). During assessments, anchoring errors can delay arriving at the correct diagnosis or even more detrimental, leading to incorrect diagnoses (Garb, 2005). Anchoring errors can also skew results. For example, if the clinician becomes overly focused on one finding or hypothesis, they may discount the client's full story by not exploring additional information that does not correlate with their view of the client's symptoms (Hammond et al., 2006; Lilienfeld & Lynn, 2015).

Attribution errors. The tendency to making decisions based on negative stereotypes that can lead clinicians to ignore or minimize the possibility of accurate conceptualizations. An example of this during the initial assessment process is attributing an acting out behavior as aggression in Black children and not exploring a possible hypothesis of anxiety (Gigerenzer & Gaissmaier, 2011; Kolen & Hendrickson, 2013; Magnavita, 2016).

Availability error. The bias of establishing an analysis of readily available material (e.g., source, information, or retrieval bias). This error is strongly influenced by recent cases a provider has seen (Lilienfeld & Lynn, 2015; Rylander & Guerrasio, 2015). In psychological assessments, this error can repeatedly occur throughout the case formulation and decision-making process. For example, a subset of this error, called *primacy bias,* can occur at the initial referral question then again after speaking to the first informant (e.g., parent), or a *recency bias* after receiving data back from a school (Croskerry, 2015; Lilienfeld & Lynn, 2015; Magnavita, 2016).

Confirmation bias. The tendency to allow more considerable influence to data that support a preliminary diagnosis or hypothesis, while selectively seeking and interpreting evidence that is confirmatory and failing to seek or discounting contradictory evidence (Lilienfeld & Lynn, 2015). This bias is one of the most well-known and researched cognitive errors (Gigerenzer & Gaissmaier, 2011; Magnavita, 2016). In terms of assessment, confirmation bias plays a significant role. Given the clinician's initial working hypotheses likely have been influenced by other errors (e.g., representative, availability), this following phase of the decision-making process should be methodical and not focused solely on skewed, confirmatory evidence. Confirmation bias poses a significant threat to the assessment process in that it can unknowingly skew data gathering and lead to inaccurate conclusions. Consider a redacted case illustration. Jason, a child with Down syndrome, evidences avoidant behavior and elopes when in large groups at school. The psychologist utilizes a behaviorist approach, employing Applied Behavioral therapy, although misses the presence of an anxiety disorder. In doing so, Jason's avoidant behavior intensifies and he manifests depressive symptoms then aggression. Psychological testing via behavior rating measures may have directed treatment toward other etiologies and treatment outcomes. In practice, a clinician may unintentionally seek only confirmatory data by asking limited questions related to their initial hypothesis or might be more likely to administer tests that confirm their working diagnosis. For example, the psychologist may view a child’s symptoms and create a working diagnosis of anxiety. The psychologist may be focused on reports by parents, schools, and children of anxiety symptoms and, in turn, administer instruments and interviews to confirm an anxiety disorder while dismissing or giving less attention to contradicting data. Then, during the assessment, the clinician follows an evidence-based assessment methodology, carefully choosing psychometrically sound instruments

appropriate for the referral question. However, the lack of attention to cognitive errors during decision can lead to not administering other relevant measures related to differential diagnoses; in turn, the client receives an inaccurate diagnosis (Hammond et al., 2006; National Academies of Sciences, Engineering, & Medicine et al., 2015).

Premature closure. Before obtaining an accurate diagnosis, this error is the tendency to discontinue seeking other possibilities during the diagnostic decision-making process. This bias tends to occur frequently following confirmation bias and especially under time constraints (Bornstein & Emler, 2000). Given the current confines of managed care organizations, some studies suggested the demand for diagnosis after the initial visit and reliance on screeners as confirmatory evidence might have increased this error within the mental health field (Eisman et al., 2000; Piotrowski, 1999; Wright et al., 2017).

Representation error. This heuristic can often be helpful in making quick decisions based on prior experiences; unfortunately, this heuristic can become a problem when it causes the clinician to ignore factors that also play a role in shaping choices (Lilienfeld & Lynn, 2015; National Academies of Sciences, Engineering, & Medicine et al., 2015). In assessments, the clinician's own subjective experiences with disorders create mental shortcuts and inform case formulations. Although this can aid the clinician by having credible case presentations more accessible to reference, it also can lead to missing important information or discounting empirical data (e.g., base rates).

Hindsight bias. The tendency for one to exaggerate the extent to which a past event could have been predicted beforehand. When evaluating a child, this is likely to occur when a child already was diagnosed (Lilienfeld & Lynn, 2015). For example, upon reviewing medical records, there are provisional or confirmed diagnoses listed. Having access to past diagnosis in

the medical field has shown providers displaying overconfidence and being more susceptible to hindsight biases (Arkes, 2013).

Regret bias. Also referred to as value-induced bias, is the tendency to allow the undesirability of a particular outcome (e.g., misdiagnosis) to alter the estimate of its likelihood of occurrence. Medical literature has argued this happens so frequently in clinical decision-making that they have developed another term *base-rate neglect* or the tendency to ignore the prevalence of a disease (Rylander & Guerrasio, 2015, p. 288). When assessing children's mental health, this can lead to overestimating base rates, overdiagnosis, or misdiagnosing in fear of a possibly severe outcome or different treatment strategy. Consider a redacted case illustration of nine-year-old Jane, who presents with labile moods and dangerous, aggressive, and risky behaviors (i.e., self-injuries). She was initially diagnosed with disruptive mood dysregulation disorder (DMDD) without the consideration of its very low prevalence rate in children and the increased likelihood of another diagnosis. Upon further exploration, it was discovered Jane was a survivor of complex trauma that best explained the onset of her symptoms and guided treatment. Given the significant overlap in presentations for children, it is especially important to consider as well as review the base-rates and differentials for diagnosis.

Diagnostic bias. The tendency to focus on specific information (e.g., personal views, stereotypes) as a diagnostic criterion instead of taking into account all of the case information and integrating empirical (National Academies of Sciences, Engineering, & Medicine et al., 2015; Watkins, 2009). When assessing a child, this might be the clinician assuming following a referral question that pathology exists to merit the referral, thus over pathologizing or allowing other biases (e.g., ethnicity stereotypes) be used to guide decision-making. This bias can be especially dangerous when differentiating the presence or severity of child pathology compared to typical development

given there are many overlaps in presenting concerns such as tantrums, crying spells, or acting out behaviors (Fletcher, 2005; Pope & Vaquez, 2007; Sallis et al., 2019).

Risk Factors and Ways to Avoid Errors

Ethical considerations. Numerous authors have recognized common ethical blind spots or pitfalls in relation to testing (Arslan, 2018; Ready & Veague, 2014; Smith, 2003). Some considerations include being aware of the referral question, not relying on third-party accounts to formulate assessments, assuring "the assessment is thorough" (Smith, 2003, para. 65), and identifying and documenting limitations (Hunsley & Mash, 2018; Krishnamurthy et al., 2004; Prinstein et al., 2019; Youngstrom et al., 2017), in addition to using statistical prediction rules (e.g., base rates), following guidelines, and adhering to diagnostic criteria (Garb, 1997, 2013). Another vital ethical standard of assessment repeatedly mentioned (APA, 2013a; Arslan, 2018; Horin, Hernandez, & Donoso, 2012; Krishnamurthy et al., 2004) is:

> When interpreting assessment results, including automated interpretations, psychologists take into account the purpose of the assessment as well as the various test factors, test-taking abilities, and other characteristics of the person being assessed, such as situational, personal, linguistic, and cultural differences, that might affect psychologists' judgments or reduce the accuracy of their interpretations. They indicate any significant limitations of their interpretations. (APA, 2017a, 9.06)

Notably, the term "judgment" appears twice in section nine of the code of conduct, with once mentioned above in standard 9.06 and the other being "See also Standard 2.04, Bases for Scientific and Professional Judgments" (APA, 2017a, 9.01), while the term "decisions" occurs twice, explicitly in:

> Obsolete Tests and Outdated Test Results (a) Psychologists do not base their assessment or intervention decisions or recommendations on data or test results that are outdated for the current purpose. (b) Psychologists do not base such decisions or recommendations on tests and measures that are obsolete and not useful for the current purpose. (APA, 2017a, 9.08)

Ethical decision-making. A technique commonly used to facilitate ethical decision-making is pursuing constant awareness of their heuristics and then challenging their biases (Charles, Gafni, & Whelan, 1999; Garb, 2005; Kicklighter et al., 2018; Riggin & Lack, 2018). When considering psychological testing with a child, this can be reviewing base rates, having a permanent differential of nonexistent pathology (e.g., avoid over pathologizing), and continuously searching for all the contradicting evidence (Arslan, 2018; Harvey, 2006; Pope & Vaquez, 2007). Other recommendations include seeking consultation (Meyer et al., 2001b), reframing or redefining the problem (Frick, 2007; Hunsley & Mash, 2008), creating a systematic review process (Wolraich et al., 2019), and acknowledging personal circumstances that may interfere with professional expectations (Pope & Vaquez, 2007; Riggin & Lack, 2018). To promote this process, assessors should pursue competency in relevant biopsychosocial and developmental information and current approaches to assessment methods needed to guide clinical decisions. In terms of cultural considerations, it is a core competency and part of the ethical guidelines for practicing psychologists to seek cultural awareness (APA, 2017a). When conducting testing with children, some cultural considerations are maintaining awareness of the implications of the test bias (see also Reynolds & Suzuki, 2012; Teresi & Jones, 2013) and challenging the threat of personal stereotypes (APA, 2013a).

Although heuristics are typically advantageous for assessors and aid a clinician throughout their decision-making process, inevitably, they also account for several errors and biases (Croskerry, 2009; Lilienfeld & Lynn, 2015; Rylander & Guerrasio, 2015). Therefore, during the decision-making process, starting from the initial hypothesis and prior to administering tests, an evaluator needs to become aware of biases that are influencing each stage of their process (Carlson & Geisinger, 2009; Suhr, 2015). In the assessment decision-making

process, it is implied the clinician begins their journey by formulating then testing hypotheses while integrating data from various sources (Kicklighter et al., 2018; Prinstein et al., 2019; Sattler, 1988). In agreement, Suhr (2015) suggests an evaluator must embark on a reflective journey from the moment the first hypothesis develops, then the evaluator begins to select questions, measures, the nature of data that will be gathered, and the method of data integration. Given the robust data obtained from tests, administering some instruments or screening measures initially can be used to improve and guide assessment decision-making (Frick, 2007; Panter & Bracken, 2013; Richardson et al., 2015; Watkins, 2009). Largely, authors agree with this approach; however, there is some concern that tests are not consistently used well (Castellanos-Ryan et al., 2016; Lilienfeld & Lynn, 2015; Reynolds & Suzuki, 2012). This is particularly true when tests are used by those who are not appropriately trained, and potentially the tests begin to detract from good decision-making (Anastasi, 1985; Cook & Coyne, 2005; Kicklighter et al., 2018; Lovitt, 1974). For this reason, it is essential to understand the current approaches to assessment methods.

General Approaches to Assessment

Structured assessment methods, for general practice, have continued to gain more traction over the past decade, which can be attributed to the evidence-based practice movement (Bornstein, 2017; Jensen-Doss & Hawley, 2011; Kazdin, 2005). However, specialties, such as forensic psychology, have had established evaluation guidelines for over two decades, namely “Guidelines for Psychological Evaluations in Child Protection Matters" (APA, 2013c). In the light of these well-supported guidelines, endorsed by APA, it is conceivable that some of the basic tenets are generalizable to a decision-making model for other types of psychological evaluations conducted with children. A closer look at these guidelines reveals many of gaps and

shortcomings in the available literature for decision-making for assessment practices. For example, the Committee on Professional Practice and Standards (COPPS, the developers of these guidelines), stated:

> These guidelines are informed by the American Psychological Association's (APA's) "Ethical Principles of Psychologists and Code of Conduct" (hereinafter referred to as the Ethics Code; APA, 2002). The term guidelines refer to statements that suggest or recommend specific professional behavior, endeavors, or conduct for psychologists. Guidelines diverge from standards in that standards are mandatory and may be accompanied by an enforcing mechanism. Guidelines are aspirational in intent. They are intended to facilitate the continued systematic development of the profession and to help facilitate a high level of practice by psychologists. Guidelines are not intended to be mandatory or exhaustive and may not apply to every professional situation. They are not definitive, and they are not intended to take precedence over the judgment of psychologists. The specific goal of the guidelines is to promote proficiency in using psychological expertise when psychologists conduct psychological evaluations in child protection matters. (APA, 2013c, p. 20)

The statement provides limitations to guidelines, while still placing significant value on the psychologist's clinical judgment. Similar to prior research examined, these guidelines suggest that clinical decision-making occurs throughout the entire assessment process. However, numerous other types of reasons for testing involve an initial decision-making process, unlike a court-mandated evaluation. Previous research has not yet presented broad-spectrum guidelines or a framework for clinical decision-making for assessments with children, although the literature has provided some comprehensive approaches and diverse functions of assessment practice. Within the application of assessments for children, a combination of the following methods surface; (a) case formulation, (b) Sattler's Four Pillars, (c) multimethod assessment, (d) Evidence-Based Assessment, (e) Three P's, and (f) Culturally Competent Assessment.

Case formulation. Numerous existing studies in the broader literature have examined theoretical approaches to assessment and case conceptualization (Hersen & Reitman, 2008). Regardless of theoretical orientation, in both treatment and assessment practices, a number of

authors have recommended using a biopsychosocial approach to gather clinically relevant information (Hersen & Reitman, 2008; Russo, 2015; van Meter et al., 2014). Since Engel's (1977) development of the biopsychosocial model, a biopsychosocial approach has been widely used as a holistic assessment tool in both mental health and medical settings (Dodge & Pettit, 2003). The biopsychosocial model assesses the intersectionality of factors across biological, psychological, and social domains that can be contributing to the concerns of an individual and suggests all influence each other in different ways (Dodge & Pettit, 2003; Engel, 1977).

In psychological assessment practice, using the biopsychosocial model—usually in the clinical interview phase—has acknowledged benefits (Richardson et al., 2015; Russo, 2015; Sattler & Hoge, 2006). One significant contributing factor is the strong initial emphasis of identifying the unique presentation of symptomology and behaviors of an individual instead of channeling the attention to solely diagnosis (Hunsley & Mash, 2018; Mihura et al., 2017; Schniering et al., 2000; Suhr, 2015). The biopsychosocial model is used to distinguish the influences of various domains to aid in decision-making, case conceptualization, and treatment (Campbell, Ruble, & Hammond, 2014; Loyens et al., 2012; Poland et al., 1982). Further case conceptualization is used throughout the assessment process and can begin with the formulation of hypothesizes (APA, 2006; Hunsley, 2009). Case formulation serves as a foundation for assessing a child in that it allows for a wide range of data to be considered and integrated (Russo, 2015). In turn, the assessor gains a comprehensive understanding of the child's unique experiences. During case formulation, some objectives include methods to gain a better understanding of presenting problems by investigating various contributing factors and informing decision-making (e.g., test selection, treatment planning) based on possible etiologies (Anastasi & Urbina, 1997; Hunsley, 2009; Wiener & Costaris, 2012). With the added context of

evidence-based practice, many current assessment approaches include case formulation within their methods.

Sattler's Four Pillars. Sattler's (2001) work established the four *Pillars of Assessment* with children. These pillars are commonly applied in various aspects of psychological evaluation practices and are often used as a theoretical framework for the assessment (Cicchetti, 1994). Sattler's (2001) four pillars are composed of (a) norm-referenced tests, (b) interviews, (c) observations, and (d) informal assessment procedures. Further, Sattler (1988, 2001) has argued to gain the best understanding of a child's unique presentation of symptoms, and a good assessment will integrate each of these four pillars.

Multimethod assessment. Sattler's four pillars supplemented the position of multimethod assessment (Hopwood & Bornstein, 2014; Riccio & Rodriguez, 2007). Specifically, Sattler's approach implies the need for evaluators to consider a wide range of methods, and consequently, information sources during the assessment process hence requiring the consideration of environmental, cultural, social, biological, cognitive, and motivational influences on the child's behaviors. Results across authors appear consistent with supporting the use of multimodal assessment approaches within both clinical and research practices (Haynes et al., 1995; Hunsley & Mash, 2018; Sattler & Hoge, 2006). However, existing research indicates many psychologists are guilty of relying on unimodal sources such as unstructured interviews (Garb, 2005; Garb, Lilienfeld, Nezworski, Wood, & O'Donohue, 2009; Hunsley & Mash, 2018; Jensen-Doss & Hawley, 2010; Schniering et al., 2000; Youngstrom et al., 2017). Many (Eisman et al., 2000; e.g., Garb, 1998; Meyer et al., 2001) have found several limitations to various forms of clinical interviews. Specifically, their studies have exposed structured or semi-structured as being tremendously time burdensome, and unstructured interviews often producing poor

accuracy and reliability. Some researchers, such as Hopwood and Bornstein (2014), have suggested this is due to "empirically validated, clinically useful models for integrating multimethod data have not been presented in a comprehensive, systematic, transtheoretical way" (p. 13). To bridge the gap between evidence-based multimodal assessment and clinical practices, Hopwood and Bornstein (2014) created a "framework for multimethod assessment and test score integration" (p. 9). The framework consists of six steps:

1. Understand the strengths and limitations of different methods.
2. Know when to collect data using multiple methods.
3. Decide which methods to use.
4. Select appropriate measures.
5. Implement a framework for integrating data from different sources.
6. Use assessment data to enhance treatment planning. (Hopwood & Bornstein, 2014, pp. 9-13)

In the first step, Hopwood and Bornstein outline the importance of knowing existing methods so the test user can differentiate between measures and cipher through extraneous variables, thus being able to apply to the individual. Next, the model encourages the use of more than one test that measures a parallel construct, especially within complex cases. The authors further contend, "Any time two tests that measure parallel constructs using different methodologies fulfill established criteria for reliability and validity, each test has the potential to add incremental validity-unique predictive value-to a test battery" (Hopwood & Bornstein, 2014, p. 10). Next, it is time to narrow down which domains require further exploration by reviewing the referral question, and relevant background data (e.g., history, records), then identifying which areas are most salient to the specific assessment because "it is important to tailor the battery to

match tests with outcome" (Hopwood & Bornstein, 2014, p. 10). In other words, it is crucial for the assessor to understand what the goals of the assessment are to inform decision-making (e.g., test selection) and competency. For example, if evaluating for autism, the evaluator should be abreast of current, empirically supported approaches to assessing for autism thereby leading to proper test selection. In step four, Hopwood and Bornstein (2014) comment preliminary test selection should not rely solely on the referral question, instead "test selection decisions should be based on validity evidence, an understanding of underlying processes engaged by different tests, cost-effectiveness, and clinical utility" (p. 10). This step also highlights the dynamic process of clinical assessment (Hopwood & Bornstein, 2014) given "initial test results guide subsequent clinical decisions when assessing" (p. 13). This means the evaluator should be ready and capable of adapting their test battery as additional questions arise (e.g., confounding test data, new information obtained). Within step five, the assessor obtains and interprets the data, creates a conceptualization with the integrated findings across test data (i.e., supporting, or contradicting data and parallel constructs), domains (e.g., cognitive, adaptive functioning), and assessment methods. Finally, in step six, all the data comes together to aid treatment planning based on the clinician's interpretations to then inform recommendations.

Evidence-based assessment. As previously outlined, assessment is a core competency within the field of psychology; moreover, Evidence-Based Assessment (EBA) is considered a vital component of EBPP (APA, 2006; Hunsley & Mash, 2007). In Hunsley and Mash's frequently cited works (2007, 2008), the authors have developed criteria for EBA including figures to aid with test selection based on empirically sound, psychometric properties. According to these researchers, EBA "emphasizes the use of research and theory to inform the selection of assessment targets, the methods and measures used in the assessment, and the

assessment process itself" (Hunsley & Mash, 2007, p. 29). Hunsley and Mash offered a concise summary of their continued efforts to develop EBA guidelines. They wrote,

> EBA involves the recognition that, even with data from psychometrically strong measures, the assessment process is inherently a decision-making task in which the clinician must iteratively formulate and test hypotheses by integrating data that are often incomplete or inconsistent. A truly evidence-based approach to assessment, therefore, would involve an evaluation of the accuracy and usefulness of this complex decision-making task in light of potential errors in data synthesis and interpretation, the costs associated with the assessment process and, ultimately, the impact the assessment had on clinical outcomes for the person(s) being assessed. (Hunsley & Mash, 2007, p. 30)

Also, Hunsley and Mash (2007) established three core principles that define EBA:

1. Research findings and specifically viable theories on both psychopathology and normal human development should be used to guide the selection of constructs to be assessed and the assessment process.

2. As much as possible, psychometrically strong measures should be used to assess the constructs targeted in the assessment.

3. Although, at present little evidence bears on the issue; therefore, the entire process of assessment (i.e., selection, use, and interpretation of an instrument and integration of multiple sources of assessment data) must be empirically evaluated. In other words, a critical distinction must be made between evidence-based assessment methods and tools, on the one hand, and evidence-based assessment processes, on the other. (p. 33)

Within EBA, guidelines for conducting an assessment with child and adolescent disorders emerged (Hunsley & Mash, 2008; Mash & Hunsley, 2005). For common disorders occurring in youth, their textbook summarizes EBA instruments as well as empirically supported methods and assessment strategies. In regards to EBA instruments, Hunsley and Mash (2008) offer concrete criteria for norms and reliability (p. 8) and validity and utility (p. 9) through the development of tables with classifications ranging from adequate, good, and excellent. EBA methods provide relevant, advantageous guidelines that are utilized throughout the research and clinical practices and have been long-established as an empirically sound framework to further support best practices (Achenbach, 2005; Kazdin, 2005; Youngstrom et al., 2017). However,

even such renowned guidelines come with several limitations within clinical practice. Many criticize current EBA methods are too time-intensive, and consequently are not used in practice (Beidas et al., 2016; Bumbarger & Campbell, 2012; Falkmer et al., 2013; Jensen-Doss & Hawley, 2010; Legare et al., 2008).

Many reiterate that aspects of EBA exhaust resources, time, and are unrealistic, thus become a burden to providers and facilities (Beidas et al., 2016; Bumbarger & Campbell, 2012; Jensen-Doss & Hawley, 2010). Unintentionally, EBA's thorough requirements create a barrier to implementation, especially in community settings (Beidas et al., 2016; Bumbarger & Campbell, 2012). Jensen-Doss and Hawley (2010) suggest to lessen the burden within certain settings, providers can select measures that are brief from start to finish (e.g., administration to interpretation), low cost, and validated for diverse application (e.g., ethnic minority, low socioeconomic status). Researchers (Beidas et al., 2016) have compiled a non-privileged, resource to aid in the implementation of EBA specifically in the context of test selection of "free, brief, and validated evidence-based assessment tools for public sector mental health settings" (p. 5). The study applies the EBA instrument criteria developed by Hunsley and Mash to easily accessible measures thereby helping break down the barrier to achieving EBA practices in public sectors. Furthermore, Beidas et al. (2016) stress the tremendous need for EBA practices, given amid well-trained providers in assessment, overall diagnostic reliability and agreement of treatment planning remains significantly low.

The Three P's. Youngstrom's (2008) work suggests using the "Three Ps" of assessment as an added approach to EBA (p. 34). The Three Ps are composed of predict, prescribe, and process. Youngstrom refers to the Three Ps as a "litmus test for assessment methods" reliant on response to the following questions: "Do they *predict* important criteria? Do they *prescribe*

specific treatments? Do they inform our understanding of the *processes* in developmental psychopathology?" (Youngstrom, 2008, p. 34). He argues if the instrument does not fall into these areas, then there is little justification to add them to a clinical or research test battery (Youngstrom, 2008). In other works, the Three Ps are applied to specific diagnostic assessment methods (e.g., Youngstrom, Freeman, & Jenkins, 2008). These supplemental resources to EBA demonstrate another area entangled in clinical decision-making before initial test administration.

Culturally informed assessments. Culturally informed psychological assessments are often difficult to achieve and can create another barrier for both clinicians and children in need (Horin et al., 2012; Liang et al., 2016). The dynamic assessment process becomes more challenging as the pool of available, valid measures with normed data is very small (Daryanani et al., 2001). Consequently, evaluators may use instruments considered culturally biased (e.g., content, validity) with limited application to diverse groups (Liang et al., 2016; Reynolds & Suzuki, 2012). Additionally, assessors may misinterpret, such as over or underestimate the influence of cultural considerations on a child's functioning; therefore, during testing, across instruments, the results can be misleading (Oakland, 2004; Whaley, 2001). Many have promoted and recommended the use of *culture fair tests* (Scarr, 1994). Getz (2011), in the book Encyclopedia of Clinical Neuropsychology, defines culture fair tests as:

> A test that is equally fair to all cultural groups. Fairness is related to a lack of bias in the interpretation or use of a test to classify or diagnose. In a culture fair test, the validity of the interpretation is similar across different cultural groups. It is unlikely that any test can entirely eliminate the influence of learning and cultural experience, given that the test content, language, directions, and validity criteria are culturally bound. However, avoiding culturally loaded items, items that are found to be unfair to certain groups of people increases the likelihood of it being a culturally fair test. Culturally loaded items, such as those that utilize pictures or general information that are differentially prevalent for certain cultures, decrease the likelihood of a culturally fair test. (Getz, 2011, pp. 755-756)

Although many instruments do not meet this standard, many researchers have strived to provide a framework for culturally sensitive assessments. For example, Roysircar (2005) published guidelines conducting multicultural assessments. The author insists on conducting assessments that are culturally informed, and psychologists must constantly seek knowledge and training of current diversity literature, including practice methods. Roysircar (2005) offers brief Practice Guidelines for Multicultural Assessment (pp. 30-32).

APA has continuously attempted to provide free resources to aid in cultural competency with psychologists. Presently, APA has offered free online access to DSM-5 Cultural Formulation Interviews (CFI), such as *Core CFI, CFI Informant Version*, and *CFI Supplementary Modules* (DeSilva, Aggarwal, & Lewis-Fernández, 2015; Roysircar, 2005). It is also expected that psychologists are knowledgeable of the current APA's (2017b) "Multicultural guidelines: An ecological approach to context, identity, and intersectionality."

Summary

Clinical decision-making is complex, requiring the integration of a plethora of abilities such as the proficiency to define clinical questions accurately, extract, then apply relevant information from literature, differentiate research methodology, select statistical procedures as well as the ability to critically evaluate studies and understand their implications for care. Clinical psychologists are expected to adhere to APA's *Standards for Educational and Psychological Testing* (2014), "Multicultural guidelines: An ecological approach to context, identity, and intersectionality" (APA, 2017b), and "Ethical Principles of Psychologists and Code of Conduct" (2017a). In that no general child assessment guidelines are available, APA (2013c) has provided "Guidelines for Psychological Evaluations in Child Protection Matters." The guidelines provide a structure to perform forensic, child evaluations, although still place a central

value on the psychologist's clinical judgment. Mirroring prior research examined, these guidelines suggest that clinical decision making occurs throughout the entire assessment process.

Regarding clinical decision-making in relation to psychological assessments, factors emerge on the macro and micro-level. Macro-level influences include (a) national healthcare policies (b) education (e.g., curriculum, training opportunities, (c) professional governing bodies (e.g., ethics, guidelines), and (d) legal considerations (e.g., law). Simultaneously, micro-level factors include (a) environmental influences, (b) external and internal clinician factors, and (c) client factors (e.g., demographics). Client factors, specifically, race, socioeconomic status, and gender, were found to significantly affect clinical decision making (Garb, 1997). Naturally, these macro and micro factors and subfactors interact with one another in a dynamic process, in turn influencing a clinical child psychologist's decision-making process regarding conducting psychological assessments. There are several decision-making models used to describe clinical judgment throughout the literature. Overall, naturalistic decision making emerges to be the best fit in terms of the demands for clinical child psychologist's decision-making in relation to assessment methods.

To gain insight on current assessment methods with children, comprehensive approaches to psychological assessments were explored. Within the application of assessments for children, a combination of the following methods surfaced: (a) case conceptualization, (b) Sattler's Four Pillars, (c) multimethod assessment, (d) Evidence-Based Assessment, (e) Three P's, and (f) Culturally Competent Assessment. Several authors have sought to describe assessment methods using different constructs and terminology to broadly describe similar considerations. Presently, no universal assessment approach exists. However, the six described assessment methods

provide general principles forming action steps, which are further synthesized in the next chapter.

CHAPTER V: SUMMARY, DISCUSSION, AND EMERGING FACTORS

Summary

Clinical psychologists, despite graduate training models (e.g., Scientist-Practitioner, Practitioner-Scholar) or diverse practice settings, often provide assessment services in practice (Krishnamurthy et al., 2004; Merikangas et al., 2010; Piotrowski, 1999). In alignment, Meyer and colleagues (2001) highlighted that assessment is considered the second most crucial clinical activity for psychologists following psychotherapy. Data collected through an APA survey (2016) mirrors other studies (e.g., Miller & Lovler, 2018; Russo, 2015) suggesting psychologists serving children often spend more time in practice conducting assessments compared to clinicians also serving other ages. Taken together, a review of the literature revealed the prevalence and significance of psychological testing in graduate training, clinical practice, and, most importantly, in supporting children's treatment.

Current national trends denote steady increases in mental health disorders among youth (Merikangas et al., 2010), and more comprehensive studies of children under the age of 18 predict upwards of 17.1 million children having a diagnosable mental disorder (Kieling et al., 2011). The onset of many mental illnesses occurs in childhood or adolescence (Castellanos-Ryan et al., 2016; Kessler et al., 2007). Without treatment, children’s mental health symptoms often persist or intensify throughout adulthood, especially without proper services (Leadbeater et al., 2012). Studies have routinely demonstrated the benefits of early interventions and the importance of accurately diagnosing to obtain best suited, evidence-based treatment (Achenbach, 2005; Ivnik et al., 2000; Stone et al., 1999).

Regarding to psychological testing and assessments within the United States, there are presently three governing sources of ethics standards (1) the Ethical Principles of Psychologists

and Code of Conduct (APA, 2017a), (2) the Standards for Education and Psychological Testing (American Educational Research Association et al., 2014), and (3) the Guidelines for Computer-Based Tests and Interpretations (APA, 1986, 2017a). When evaluating with children, added ethical considerations include parental consent and involvement, children's rights, confidentiality, identification of the patient, separation of the parent and child during an assessment, nondiscriminatory assessment practices, and the use of multiple sources of information and appropriate measures (Elwyn et al., 2012; Heffer et al., 2009; Leong et al., 2013; Sattler & Hoge, 2006; Swerdlik & Cohen, 2013). It is also expected that psychologists are knowledgeable of the current APA's (2017b) "Multicultural guidelines: An ecological approach to context, identity, and intersectionality."

APA includes "assessment" within the core competencies and ethics of training (Leong et al., 2013, p. 267). Further, APA has outlined in detail core competencies or benchmarks within several topics, such as scientific knowledge and methods, cultural diversity, and assessment. Krishnamurthy et al. (2004), a workgroup, presented eight-core benchmarks within the practice of assessment, which included: (1) understanding psychometric theory; (2) foundations of psychological assessment (e.g., skill, empirical, theoretical); (3) assessment techniques; (4) assessing outcomes; (5) critical evaluation skills; (6) collaborative professional relationships; (7) assessment and intervention relationship; and (8) technical assessment skills. While these benchmarks are detailed, Clay (2010) pointed out there is still no "mandate" (p. 49) to use them. Still, when conducting assessments with children, the evaluator should be cognizant of the evolving ethical codes and competencies. By staying informed, being self-reflective, and consulting, professionals practicing assessment can maintain proper procedures and be able to effectively apply these guidelines to the unique needs of children and their families.

The DSM-5 informs psychologists' diagnostic framework. Within prevalent diagnoses in childhood, surprisingly, only two disorders—specifically learning disorders and intellectual disability—mention using "psychometric evidence" (p. 69) and psychometrically sound tests of intelligence (p. 39) in DSM-V's diagnostic criteria (American Psychiatric Association, 2013c). Incongruent with diagnostic criteria not including psychometrics, over the past century individual measures have been shown to aid in diagnostic determination (Binet & Simon, 1916; Minton, 1988). Another differentiating factor is that clinical psychologists tend to spend more time conducting assessments with children (APA, 2016; 2016; Russo, 2015; Tuma & Pratt, 1982). One explanation for this is the aid of psychological assessments in diagnostic decision-making, especially for complex cases (Hopwood & Bornstein, 2014; Leong et al., 2013). Often, with children, the underlying cause of their problems is not clear. A familiar illustration of this in practice is a child presenting for an assessment due to difficulties at home and school. The child is presently having trouble with sleep, conflicts with peers, poor concentration, and inattention. Given this brief symptom list, the child could meet the criteria for ADHD, anxiety, or even depression. Thus, given the overlap in criteria and symptomology for children, it creates complexities that may merit testing to assist with differentiating diagnosis. Additionally, there have been multiple studies addressing the contradicting information obtained from multiple informants during the assessment process with children (Comer & Kendall, 2004; De Los Reyes & Kazdin, 2005). To minimize these discrepancies, several authors have advocated for the inclusion of psychometric evidence as diagnostic criteria (Hyman, 2011), as well as development of informant-specific diagnostic classification systems for children (De Los Reyes & Kazdin, 2005; Offord et al., 1996).

Prior psychological testing research, standards, and guidelines have continually emphasized the significance of assessment competence, clinical decision-making, and utilization of psychometrically sound instruments; additionally, APA also does not endorse any psychological tests nor do current testing standards denote minimum psychometric values necessary to indicate that an instrument meets the benchmark of being reliable or accurate. Furthermore, research supporting the utility of assessment methods and studying clinical decision making has been consistently overlooked and omitted. While there is substantial evidence supporting the use of testing in practice, there is a limited amount of research linking these same benefits to the assessment process. One explanation for this is that studying psychological assessments is a challenging task due to the intricacies of the process (Fernández-Ballesteros et al., 2001; Mash & Hunsley, 2005). Subsequently, many articles have begun to describe psychological assessment methods as a persistent, dynamic, decision-making process (APA, 2013c; Hunsley & Mash, 2014; Kicklighter et al., 2018; Krishnamurthy et al., 2004; Norcross, 1991; Spengler et al., 1995).

Nevertheless, the skewed focus on psychological tests in research, guiding standards, and legislation has had several implications. One concern echoed throughout studies includes managed care constrictions (Meyer et al., 2001; Wright et al., 2017). Within the practice of psychological assessments, upwards of 72% of psychologists reported adjusting their test utilization as a direct result of managed care (Piotrowski, Belter, & Keller, 1998). Managed care has been associated with misdiagnosis (Miller & Luft, 1997), mismatched treatment (Meyer et al., 2001; Miller & Luft, 1997; Roysircar, 2005), and inhibiting psychologist's diagnostic decision-making (Wright et al., 2017). The ramifications of managed care within psychological assessment practices further delineates the weight of managed care within the clinical decision-

making process. Another consideration for the current lack of assessment decision-making models is that psychologists tend to utilize a small segment of learned test instruments that become familiar as instrument staples in practice. In addition, the economics of increasing test costs could be another factor that may restrict psychologists' exposure and use of newer and more novel test instruments.

Clinical decision-making is crucial in everyday clinical practice for psychologists and remarkably fundamental during assessment methods. This internal process becomes a part of the psychologist's refined tacit knowledge throughout their real-world experiences. Clinical decision-making is the interaction of uncertainty with applying reasoning with training, scientific theory, practical experiences, client reports, problem-solving, assessing outcomes, eliminating extraneous data, and weighing risks and benefits (APA, 2017a; Bellman & Zadeh, 1970; Hammond et al., 2006; National Academies of Sciences, Engineering, & Medicine et al., 2015; Smith et al., 2008; Wills & Holmes-Rovner, 2006). The EBPP movement has established the basis of competencies and criteria in several areas including assessment (APA, 2006; Frick, 2007). However, literature has not explicitly defined or explored clinical judgment in psychology practice, signifying that clinical mental health professionals are far behind on developing decision-making guidelines when compared to the progress of medical decision-making models (Magnavita, 2016).

In terms of clinical judgment and decision-making, several researchers have pursued an advanced understanding of the phenomena. In turn, diverse expressions and paradigms fundamentally to describe similar experiences have emerged within the literature. As expected, across disciplines, there is no standard definition of clinical judgment or decision-making. When comparing some of the most common clinical decision-making paradigms used in other

healthcare fields, the Naturalistic Decision Making model emerges to have the most similarities to factors correlated with the procedures of psychological assessment with children. Comprehensive approaches to psychological assessments were explored to gain insight on current assessment methods with children. Within the application of assessments for children, a combination of the following methods surface: (a) case conceptualization, (b) Sattler's Four Pillars, (c) multimethod assessment, (d) Evidence-Based Assessment, (e) Three P's, and (f) Culturally Competent Assessment. Several authors have sought to describe assessment methods using different constructs and terminology to describe similar considerations broadly. Presently, no universal assessment approach exists.

Clinical Implications

Psychological assessment is widely defined as a dynamic, decision-making process (Hunsley & Mash, 2007; Kazdin, 2005; Youngstrom et al., 2017). Previous literature has shown that clinical decision-making is prone to biases (Garb, 2005). Further studies within adjacent healthcare fields have gone a step further and offered recommendations to reduce errors in clinical decision-making (Garb, 2013; Liang et al., 2016; Lilienfeld & Lynn, 2015). However, within clinical psychology, prior research, governing bodies, and guidelines have narrowly focused on test properties (e.g., psychometrics) despite this only being one factor of the complex assessment process (Mash & Hunsley, 2005).

In other healthcare fields, NDM research has progressed the understanding of clinical decision-making (Klein et al., 2010). Consequently, being able to inform interventions to improve providers' decision-making abilities by reducing uncertainty and improving performance (Klein, 2008). Before these findings, it was hypothesized identifying assumptions and gathering as much information possible (Klein & Wright, 2016) were the convenient ways to

reduce errors. However, NDM literature has shown the exact opposite: specifically, listing assumptions are primarily unhelpful because oftentimes the provider is unaware of their biases and could certainly not self-generate a list (Klein, 2015; Klein & Wright, 2016).

Additionally, studies have demonstrated that increased indecision and diminished decision-making abilities occur both when too much data is attained (e.g., inadequate framing of information) as well as during the lack of data (Klein & Wright, 2016). NDM studies have also provided insight into the differences between novice and expert decision-makers (Falzer, 2018; Klein, 2015). Novices with less practical experience tend to engage in the reverse process of an expert. Notably, novices heavily rely on explicit knowledge or analytical approaches, weigh options mentally, and engage in mental rehearsal excessively (Kicklighter et al., 2018; Klein et al., 2010; Klein & Hoffman, 1992). Alternatively, experts with significant experience, tend to respond intuitively by rapidly recognizing patterns intuitively and often not referencing explicit knowledge.

In more recent works, Klein and Wright (2016) have also recommended macro cognitive models as tools that can be used in cognitively challenging activities to enhance performance. For these reasons, there are benefits of applying the Naturalistic Decision Making model to the clinical decision-making of child psychologists. One advantage is that NDM models allow grounded, clinically informed, and methodologically complex responses to be operationalized and further studied in clinical psychology. With the need of more continuity between doctoral training and practice, NDM models and subsequent macro cognitive models provide a structure to infuse best clinical practices imparted in training with real-life practice pragmatics (e.g., economics, tacit knowledge).

Discussion

This study serves as a preliminary look into factors that attribute to a clinical child psychologists' judgment and decision-making when choosing to conduct testing with a child. Currently, ethics (APA, 2017a), guidelines (Krishnamurthy et al., 2004), and training (Garb, 2005; Lovitt, 1974; Spengler et al., 1995) all encourage the use of clinical decision-making regarding psychological assessment practices. It is implied that clinical decision-making is a critical component of everyday practice for psychologists despite several concerns. In particular, empirical research available to aid psychologists in this process has been consistently neglected (Bruchmüller et al., 2012; Hunsley & Mash, 2008; Kazdin, 2008; Leadbeater et al., 2012; Pastor & Reuben, 2005; Richardson et al., 2015). Even more concerning, within assessment literature, limited studies have investigated the reliability (consistency within psychologists) and validity (that the decision to conduct an assessment with a child makes a difference in treatment outcomes when compared to other methods) of clinical decision-making.

Findings. This clinical research project conducted a systematic review of peer-reviewed literature focusing on current theories and empirical data related to clinical decision-making, psychological assessments, and clinical child psychologists. Results attained after integrating different philosophies from psychology and adjacent healthcare disciplines are presented.

Literature Review Question #1. What is the role of psychological testing in the practice of child clinical psychology?

For psychologists serving children, psychological assessment plays an integral role in clinical practice. Many psychologists utilize psychological assessment to inform diagnosis and treatment recommendations (Lemberger et al., 2013; Mash & Hunsley, 2005; Mülberger, 2017; Youngstrom et al., 2017). The importance of psychological assessment has been emphasized

within training, core competencies, and ethical standards (American Educational Research Association et al., 2014; APA, 2017a; Bersoff et al., 2011; Sattler, 1988). Studies have consistently found clinicians serving children endorsed spending more time in practice conducting assessments compared to those also working with adult populations (APA, 2016; Piotrowski, 1999; Tuma & Pratt, 1982; Wright et al., 2017).

Literature Review Question #2. What is the clinical utility of psychological assessment with children?

The clinical advantages of testing include reducing unconscious bias by standardization thus lessening legal and ethical issues (Meyer et al., 2001). The validity of psychological tests is comparable to medical test validity (Meyer et al., 2001). Other benefits of assessment methods include providing unique sources of information. Further, it has been suggested that assessments can reduce diagnostic errors (Meyer et al., 2001). Specifically, Meyer et al. (2001) found clinicians who do not conduct testing or rely exclusively on interviews tend to make increased diagnostics errors and miss pertinent information. Further, testing has been associated with enhanced abilities to detect subtle behavioral indications of psychological complications (Carlson & Geisinger, 2009; Majnemer, 1998; Meyer et al., 2001). Concurrently, testing can help guide differential treatment needs for both medical and mental health conditions, as well as accurate monitoring of treatment over time (Meyer et al., 2001). In other words, psychological testing assists psychologists in identifying and understanding the nature of a child's difficulties that consequently aids their ability to develop the best, evidence-based treatment plan for that child. Although there is substantial evidence supporting the use of testing in practice, there is a limited amount of research linking these same benefits to the assessment process. What studies have found from investigating some aspects of the assessment process is that empathic feedback

can be a form of treatment and therapeutic intervention (Leadbeater et al., 2012; Meyer et al., 2001; Youngstrom et al., 2017). Moreover, the absence of assessment testing, secondary to managed care restrictions, has been associated with misdiagnosis (Miller & Luft, 1997) and mismatched treatment (Meyer et al., 2001; Miller & Luft, 1997; Roysircar, 2005). Furthermore, a majority of psychologists have openly acknowledged their diagnostic decision-making was inhibited when unable to administer their clinically informed test battery (Wright et al., 2017).

Transposing the clinical utility from the plethora of testing research to assessment methods lends some associations of usefulness. Unfortunately, these unexplored constructs leave psychologists unable to provide scientific data in response to requests from program administrators, third-party payers, or clients to justify the professional time and costs associated with psychological assessment (Hunsley & Mash, 2014, p. 77). In addition, without data, it perpetuates the ideology that psychological assessment practices are optional or a supplementary service (Camara et al., 2000; Meyer et al., 2001; Wright et al., 2017), rather than being established as a valuable, clinically, necessitated procedure.

Literature Review Question #3. What approaches are recommended as best practices in clinical decision-making and child psychological assessment?

The psychological assessment process entails dynamic clinical decision-making. Across assessment research examined, it is suggested that clinical decision-making occurs throughout the entire assessment process (Fernández-Ballesteros et al., 2001; Hunsley & Mash, 2007; Kolen & Hendrickson, 2013; Youngstrom et al., 2017). Clinical decision-making is complex and challenging to study given it is composed of an integration of abilities including both tacit and explicit knowledge, in addition to phenomenon such as intuition (Hogarth, 2010; Klein et al.,

2010). Yet, within psychological assessment literature, studies have failed to investigate the reliability and validity of clinical decision-making.

A review of the literature within clinical psychology reveals psychologists have an obligation to adhere to APA's Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014), Multicultural guidelines: An ecological approach to context, identity, and intersectionality (APA, 2017b), and Ethical Principles of Psychologists and Code of Conduct (APA, 2017a). At this time, general child assessment guidelines are nonexistent, although APA has developed specialty assessment guidelines (see APA, 2013c).

With the sparse data available regarding clinical decision-making within psychology, other healthcare disciplines were examined. The Naturalistic Decision Making model emerges to have the most similarities to factors correlated with the procedures of psychological assessment with children. Furthermore, universally accepted testing methods for children endorsed by governing bodies are not yet developed. In turn, comprehensive approaches to psychological assessments with children were reviewed. The methodological considerations for child psychological assessments are: (a) case conceptualization, (b) Sattler's Four Pillars, (c) multimethod assessment, (d) Evidence-Based Assessment, (e) Three P's, and (f) Culturally Competent Assessment.

Introduction of Child Assessment Clinical Decision-Making Models

Within the proposed framework, macro-level and micro-level factors identified in this literature review can independently or simultaneously influence the psychologist (decision-maker) or assessment process (e.g., decision-task, test selection). Some of the factors can be taught, fostered, operationalized, and explicitly tested while some are more difficult to explore

but not impossible. Grounded within the NDM paradigm, Figure 2 represents the hypothesized clinical decision-making model regarding psychological testing for children.

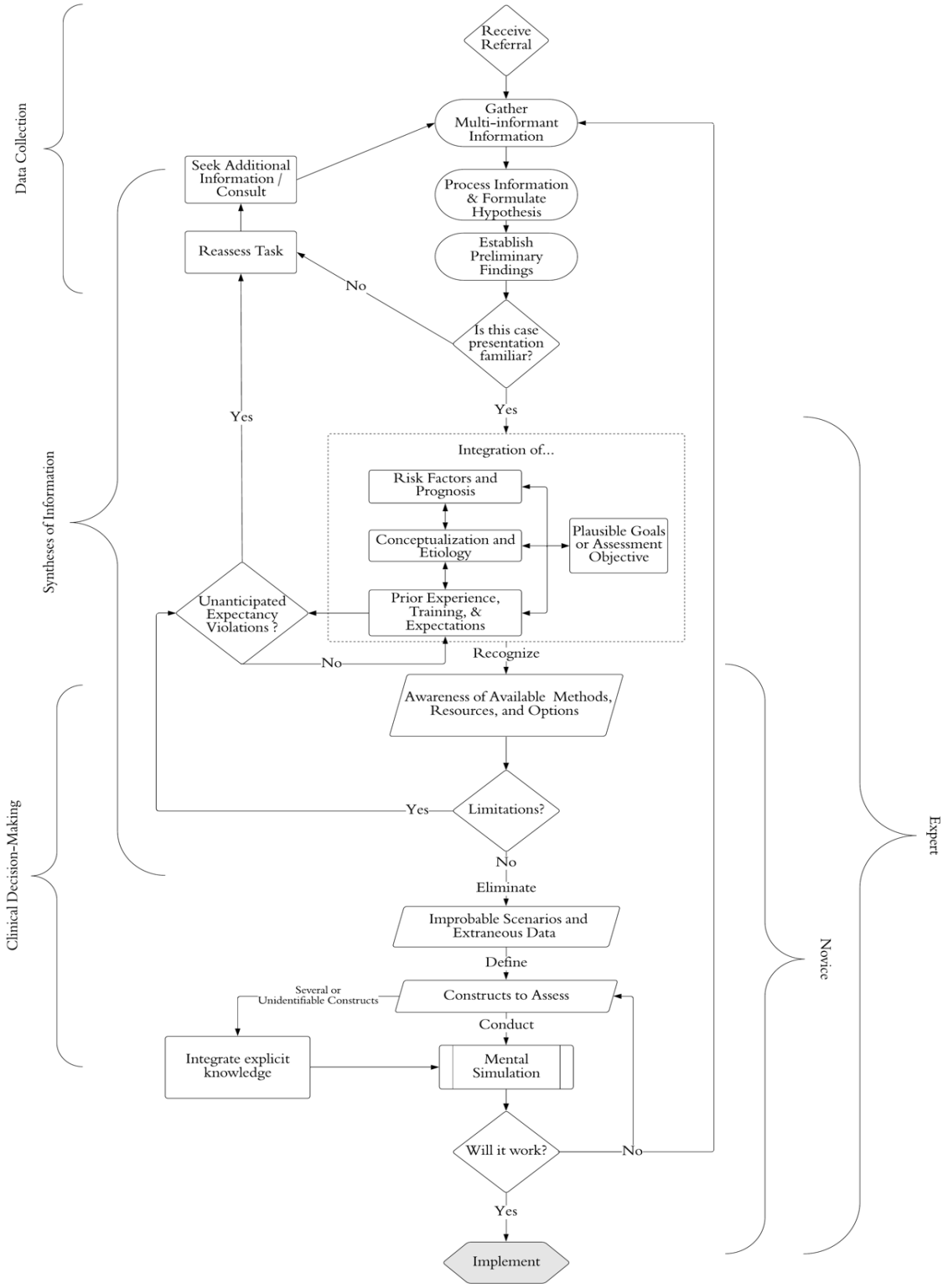

Figure 2. Proposed Model of Clinical Psychologists' Decision-Making in relation to Conducting Psychological Assessments with Children. (Adapted from G. A. Klein, 1998, *Sources of Power*)

To further illustrate the complex interactions between proposed factors, Figure 3 is provided. This model was created to better demonstrate the internal factors leading to the mental simulation of the task at hand (i.e., conducting the assessment).

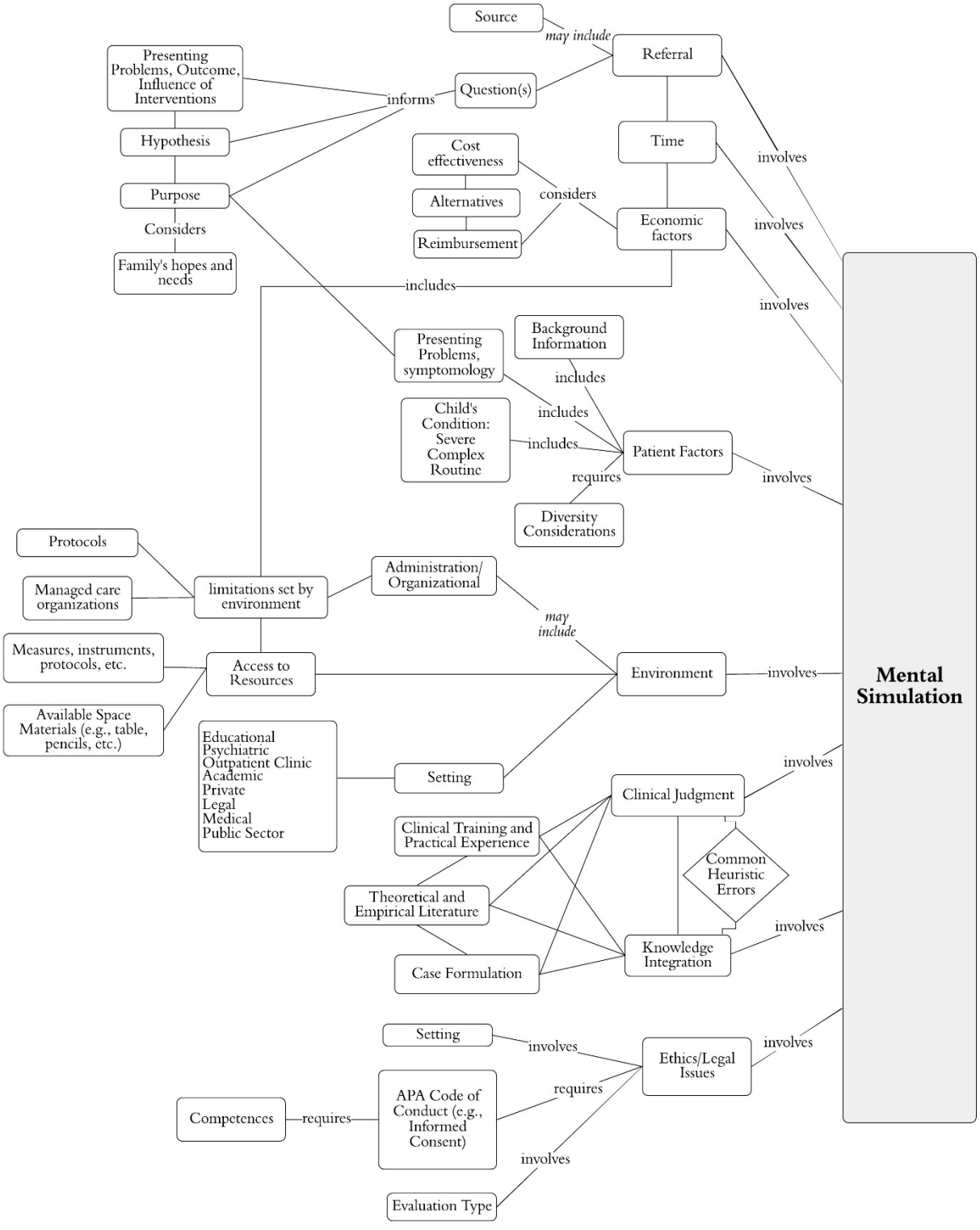

Figure 3. Interaction of the Assessor's Internal Factors in Clinical Decision-Making that leads to Mental Simulations. (Bui, 2020)

Finally, from the proposed clinical-decision making model in this project, supplemental recommendations in the form of *action steps* were created to outline effective decision-making in clinical practice (Figure 4). The action steps were guided by the emerging factors of effective clinical decision-making and current assessment methods applicable to children. The action steps are designed to inform effective decision-making for assessors at various steps of conducting assessment with children.

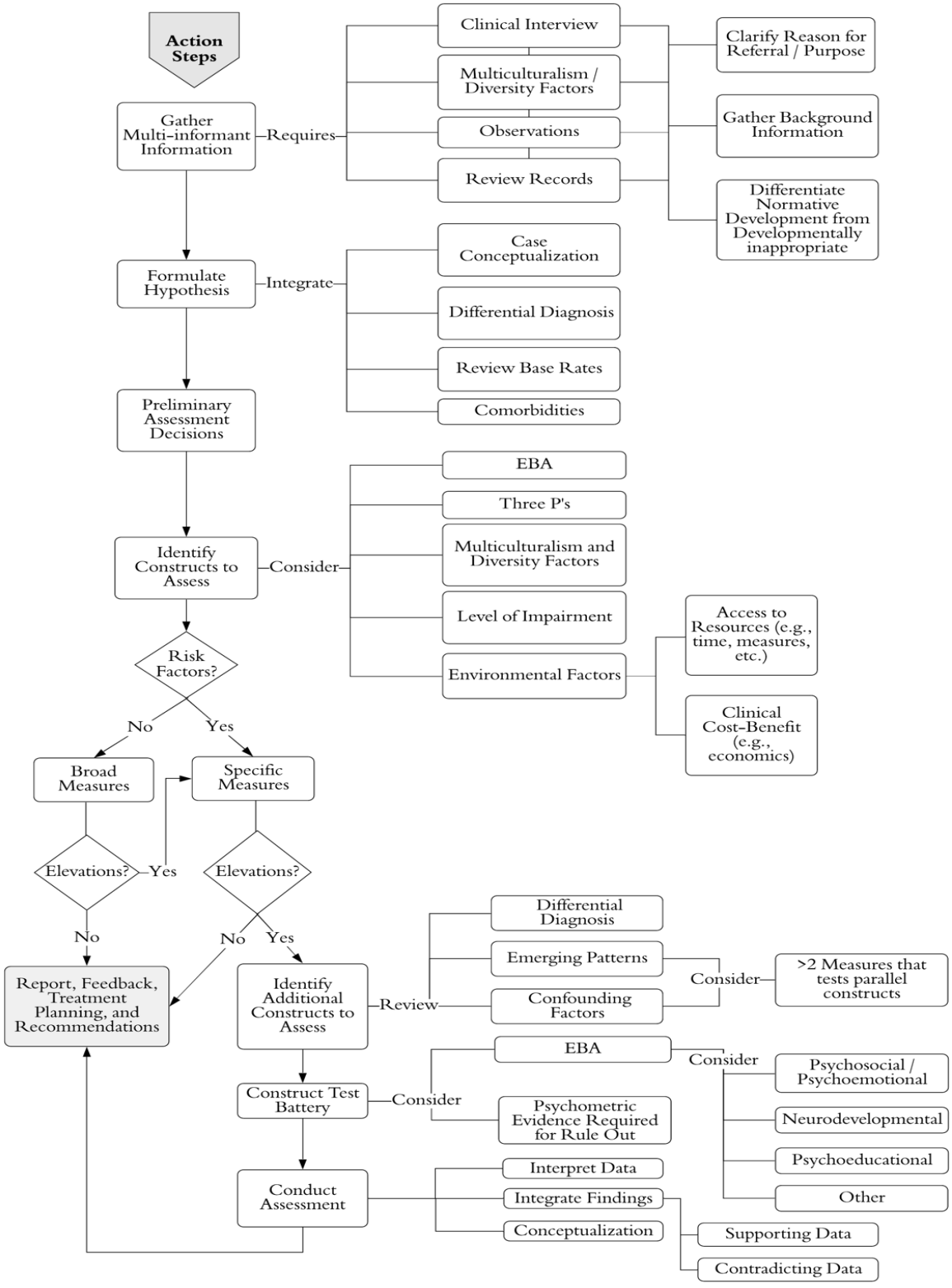

Figure 4. Effective Decision-Making Action Steps for Conducting Assessments with Children. (Bui, 2020)

Clinical Recommendations

Because of the lack of specific research on assessment utility and clinical decision-making, findings should function as recommendations to supplement practice and aid in best practices. Furthermore, to provide the highest quality of care, psychologists should seek consultation from peers in addition to staying current on evidence-based practices regarding children, assessment methods, and their profession. In addition to the proposed factors supported by the literature, the following are hypothesized factors for future consideration.

Clinician Factors

- Demographics
    - Age, gender, culture, and race
- Personal time constraints
- Workload
- Stress level
- Perceived support and safety in the setting
- Social Support
- Personal beliefs, values, morals, ethical views
- Level of Awareness, such as limitations, abilities, areas of growth
- Motives

Client Factors

- Socioeconomic status
- Attitude, behavior, and overall presentation
- Faith, spirituality, ideology
- Medical Complexities

Environmental Factors

- Culture of setting
- Geographical location of the setting
- Size of organization

Limitations of Clinical Research Project

A literature review allows for a rigorous, impartial, and inclusive evaluation of studies. Although the intent of this project was exploratory, the findings might not necessarily reflect the future quantitative data collected. The author's intent to execute a critical, comprehensive, review was challenged by the overall lack of research conducted exclusively with clinical child psychologists as it pertained to the research questions. Some of the research evaluated for clinical decision-making was collected from comparable healthcare fields thereby limiting the scope of the analysis. Given the abundance of literature on psychological testing, similarities were suggested. However, there remains a scarcity of literature concentrating on the utility of assessment methods and even less with children. Moreover, this project analysis narrowly focused on factors related to clinical child psychologists and child assessments. Further studies would benefit from investigating the possibility of a universal model for clinical decision-making for psychologists by defining universal factors through integrating any overlapping processes that emerge.

The literature and paradigms regarding clinical decision-making are extensive across disciplines. However, presently there is not a comprehensive, universal, understating, or theoretical framework widely accepted within healthcare professions. Historically, studies have exclusively ascribed to one particular theory or philosophy. This project sought to provide a comprehensive overview of decision-making models; however, similarly found one model (i.e.,

NDM) to align closer to the task demands. In turn, to develop a best-fitting model, future works are encouraged to explore multiple decision-making frameworks to apply to psychological assessment methods.

When analyzing the emerging themes in literature, no studies were found to have incorporated all of the concepts used within the proposed model. Instead, most studies purely focused on one or two of the concepts under review thus requiring cross-examination of multiple theories and benchmarks. Taking into consideration the extent of variation in nomenclature, the variables developed for this project's model first required the author to assess, interpret, conceptualize, translate, compose a description, and integrate into one of the emerging themes. As a result, author bias may have occurred during the collection data phase and the interpretation of the findings. To reduce distortions, the author made efforts to include direct quotes throughout this project and cross-reference multiple works when dissecting novel approaches.

Future Directions

Many of these limitations can provide future researchers with areas of inquiry that will expand the understanding of child psychological assessment practices and processes involved in clinical decision-making. Future research might include mixed methods approaches to explore naturalistic means of clinical decision-making. For instance, an extensive, unstructured interview has been used in other studies exploring clinical decision-making; such could provide a more in-depth examination of clinical child psychologists' tacit knowledge by detailing attitudes, feelings, and actions (see Appendix for a guide to a semi-structure interview). Additionally, sampling the broader population of mental health professionals who conduct testing with children, such as school psychologists, might ensure that the findings would have a universal application to inform systemic change. Extending research within these areas could provide new

information about evidence-based practices (e.g., EBPP and EBA) and guide improvements in diagnostic criteria, as well as ascertain the effectiveness and clinical necessity of conducting psychological assessments.

Future research is necessary to test the hypotheses suggested by the findings and model of this project. Future studies may refine this model to quantify clinical decision making and improve the validity and reliability among psychologists. This could be accomplished through exploring which components of practical experience are involved in tacit knowledge for effective decisions in clinical psychologists. Expanded data collection on individual factors, such as creativity, confidence, perceived support, stress levels, and risk-taking, could all be explored and possibly be applied to new approaches in clinical practice decisions.

With more data, studies might gain a better understanding of what models aid in effective clinical decision-making ultimately connecting children to best practices more reliably. Developing a decision-making model applicable to psychologists could help advocate for significant, macro changes (e.g., national healthcare policies). Moreover, given the clinical implications, the model can foster increased continuity across graduate training and clinical practice. One recommendation is for academic training to infuse teachings of best clinical practices with business pragmatics to enhance trainees' decision-making practices. Offering a graduate training course on economic pragmatics in various practice settings could help foster the development of tacit knowledge of trainees. Furthermore, through the NDM lens, using macro cognitive models (e.g., ShadowBox) within doctoral coursework could aid in enhancing the performance of student clinicians thereby improving client outcomes.

Conclusions

The intent of this study was to provide a preliminary look into contributing factors that impact a clinical child psychologists' decision-making when choosing to conduct testing with a child. Several factors emerged from the literature review that has been organized into macro and micro-level influences illustrated in Figure 1. Conducting psychological assessments with children is a demanding, complex, and dynamic task. Psychologists gather data about their clients and then determine what information is meaningful. In practice and in literature, this process is often referred to as clinical decision-making. However, currently, there are no guidelines in place and little research available demonstrating the utility of psychology assessment or clinical decision-making. Filling these gaps in research will require a full and comprehensive understanding of several inter-sectionalities; nonetheless, establishing the validity and reliability of clinical decision-making in psychologists is a part of the first steps.

During the literature analysis, it was determined that a comprehensive framework for clinical decision making in psychology did not exist. However, despite the lack of studies, the findings of this clinical research project suggested patterns across the literature. In addition, emerging factors suggested that clinical psychologists' decision-making process closely resembles the NDM (Figure 2). This proposed clinical-decision model serves as a stepping stone for future research towards investigating the complexities in this process (Figure 3). Finally, after analyzing factors weighing into effective clinical decision-making, a proposed model of action steps (Figure 4) outlining psychological assessment methods with children was created to inform clinical practice. The author hopes the preliminary data from this project will encourage other researchers, academics, students, and clinicians to further establish the scientific merit of

assessment use and clinical decision-making as well as enrich the education and training of future evaluators.

References

Achenbach, T. M. (2005). Advancing assessment of children and adolescents: Commentary on evidence-based assessment of child and adolescent disorders. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 541-547. https://doi.org/10.1207/s15374424jccp3403_9

Acklin, M. W. (2002). How to select personality tests for a test battery. In J. N. Butcher (Ed.), *Oxford textbooks in clinical psychology, (Vol. 2). Clinical personality assessment: Practical approaches* (pp. 13-23). New York, NY: Oxford University Press.

Adair, J. G., Bélanger, D., Dion, K. L., & Sabourin, M. (1998). *Advances in psychological science: Volume 1 social, personal, and cultural aspects*. Hove, UK: Psychology Press.

Altmaier, E. M., & Tallman, B. A. (2013). Psychological assessment in medical settings. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbooks in psychology. APA handbook of testing and assessment in psychology, Vol. 2. Testing and assessment in clinical and counseling psychology* (p. 285-302). American Psychological Association. https://doi.org/10.1037/14048-017

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US). (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

American Psychological Association [APA]. (1986). Committee on Professional Standards, American Psychological Association, Board of Scientific Affairs, and Committee on Psychological Tests and Assessment. *Guidelines for computer-based tests and interpretations*.

American Psychological Association [APA]. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271-285. https://doi.org/10.1037/0003-066X.61.4.271

American Psychological Association [APA]. (2013a). *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology*. https://doi.org/10.1037/14047-000

American Psychological Association [APA]. (2013b). *APA handbooks in psychology. APA handbook of testing and assessment in psychology, Vol. 3. Testing and assessment in school psychology and education.* (K. F. Geisinger, B. A., Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez, Eds.). https://doi.org/10.1037/14049-000

American Psychological Association [APA]. (2013c). Guidelines for psychological evaluations in child protection matters. *American Psychologist*, *68*(1), 20-31. https://doi.org/10.1037/a0029891

American Psychological Association [APA]. (2013d). Understanding psychological testing and assessment. Retrieved January 9, 2020, from https://www.apa.org/helpcenter/assessment

American Psychological Association [APA]. (2016). *2015 APA survey of psychology health service providers*. *2015 survey of psychology health service providers*. Washington, DC. Retrieved from https://www.apa.org/workforce/publications/15-health-service-providers/report.pdf%0Ahttp://www.apa.org/workforce/publications/15-health-service-providers/index.aspx

American Psychological Association [APA]. (2017a). Ethical principles of psychologists and code of conduct. *American Psychologist*. https://doi.org/10.1037/0003-066X.57.12.1060

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

American Psychological Association [APA]. (2019). *Testing and assessment*. Retrieved from https://www.apa.org/science/programs/testing/

Anastasi, A. (1985). Psychological testing: Basic concepts and common misconceptions. In A. M. Rogers & C. J. Scheirer (Eds.), *The G. Stanley Hall Lecture Series. The G. Stanley Hall lecture series, Vol. 5* (pp. 87-120). Washington, DC: American Psychological Association. https://doi.org/10.1037/10052-003

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall/Pearson Education.

Arkes, H. R. (2013). The consequences of the hindsight bias in medical decision making. *Current Directions in Psychological Science*, *22*(5), 356-360. https://doi.org/10.1177/0963721413489988

Arslan, R. (2018). A review on ethical issues and rules in psychological assessment. *Journal of Family, Counseling and Education*, *3*(1), 17-29. https://doi.org/10.32568/jfce.310629

Bailey, D. B., Hebbeler, K., Scarborough, A., Spiker, D., & Mallik, S. (2004). First experiences with early intervention: A national perspective. *Pediatrics*, *113*(4 I), 887-896. https://doi.org/10.1542/peds.113.4.887

Baker-Ericzén, M. J., Jenkins, M. M., & Brookman-Frazee, L. (2010). Clinician and parent perspectives on parent and family contextual factors that impact community mental health services for children with behavior problems. *Child and Youth Care Forum*, *39*(6), 397-419. https://doi.org/10.1007/s10566-010-9111-9

Bakeman, R., & Quera, V. (2012). Behavioral observation. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbooks in psychology. APA handbook of research methods in psychology, Vol. 1. Foundations, planning, measures, and psychometrics* (pp. 207-225). American Psychological Association. https://doi.org/10.1037/13619-013

Beacon Health Options. (2015). 3.1101 Psychological/neuropsychological testing (child/adolescent). Retrieved December 12, 2019 http://www.valueoptions.com/providers/Handbook/clinical/3-1101.pdf

Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., … Mandell, D. S. (2016). Free, brief, and validated: Standardized instruments for low- resource mental health settings Rinad. *Cognitive Behavioral Practice*, *22*(1), 5-19. https://doi.org/10.1016/j.cbpra.2014.02.002.Free

Bellman, R. E., & Zadeh, L. A. (1970). Decision-making in a fuzzy environment. *Management Science, 17*(4), B141-B273. https://doi.org/10.1287/mnsc.17.4.b141

Belter, R. W., & Piotrowski, C. (2001). Current status of doctoral-level training in psychological testing. *Journal of Clinical Psychology, 57*(6), 717-726. https://doi.org/10.1002/jclp.1044

Benson, N. F., & Donohue, A. (2018). Assessment. In S. L. Grapin & J. H. Kranzler (Eds.), *School psychology: Professional issues and practices* (pp. 97-113). New York, NY: Springer Publishing Company.

Bersoff, D. N., DeMatteo, D., & Foster, E. E. (2012). Assessment and testing. In S. J. Knapp, M. C. Gottlieb, M. M. Handelsman, & L. D. VandeCreek (Eds.), *APA handbooks in psychology. APA handbook of ethics in psychology, Vol. 2. Practice, teaching, and research* (pp. 45-74). American Psychological Association. https://doi.org/10.1037/13272-004

Binet, A., & Simon, T. (1916). *The development of intelligence in children* (The Binet-Simon Scale) (E. S. Kite, Trans.). Philadelphia, PA: Lippincott Williams & Wilkins. https://doi.org/10.1037/11069-000

Blashfield, R. K., & Herkov, M. J. (1996). Investigating clinician adherence to diagnosis by criteria: A replication of Morey and Ochoa (1989). *Journal of Personality Disorders*, *10*(3), 219-228. https://doi.org/10.1521/pedi.1996.10.3.219

Bornstein, R. F. (2017). Evidence-based psychological assessment. *Journal of Personality Assessment*, *99*(4), 435-445. https://doi.org/10.1080/00223891.2016.1236343

Bracken, B., & Nagle, R. (2017). *Psychoeducational assessment of preschool children* (4th ed.). New York: NY: Routledge

Brett, G. S. (1921). *A history of psychology: Vol. 3. Modern psychology*. (J. H. Muirhead, Ed.) (Vol. 3). Sydney, Australia: George Allen & Unwin.

Brondolo, E. (2015). Racial and Ethnic Disparities in Health. *Psychosomatic Medicine*, *77*(1), 2-5. https://doi.org/10.1097/psy.0000000000000149

Bruchmüller, K., Margraf, J., & Schneider, S. (2012). Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *Journal of Consulting and Clinical Psychology*, *80*(1), 128-138. https://doi.org/10.1037/a0026582

Buchanan, L., & O'Connell, A. (2006). A brief history of decision making. *Harvard Business Review*, *84*(1), 32-132.

Bumbarger, B. K., & Campbell, E. M. (2012). A state agency-university partnership for translational research and the dissemination of evidence-based prevention and intervention. *Administration and Policy in Mental Health and Mental Health Services Research*, *39*(4), 268-277. https://doi.org/10.1007/s10488-011-0372-x

Bureau of Labor Statistics. (2018). Occupational employment and wages, May 2018: 25-4012 Curators. Retrieved March 7, 2019, from https://www.bls.gov/oes/current/oes291131.htm

Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M. Z., & Erkanli, A. (1995). Children's mental health service use across service sectors. *Health Affairs*, *14*(3), 147-159.

Bush, P. J., & Iannotti, R. J. (1990). A children's health belief model. *Medical Care*, *28*(1), 69-86.

Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice*, *31*(2), 141-154. https://doi.org/10.1037//0735-7028.31.2.141

Campbell, J. M., Ruble, L. A., & Hammond, R. K. (2014). Comprehensive developmental assessment model. In L. A. Wilkinson (Ed.), *School psychology book series. Autism spectrum disorder in children and adolescents: Evidence-based assessment and intervention in schools* (pp. 51-73). American Psychological Association. https://doi.org/10.1037/14338-004

Carise, D., & Gurel, O. (2000). Benefits of integrating technology with treatment: The DENS project. In J. L. Sorensen, R. A. Rawson, J. Guydish, & J. E. Zweben (Eds.), *Drug abuse treatment through collaboration: Practice and research partnerships at work* (pp. 181-195). American Psychological Association. https://doi.org/10.1037/10491-011

Carlson, J. F., & Geisinger, K. F. (2009). Psychological diagnostic testing: Addressing challenges in clinical applications of testing. In R. P. Phelps (Ed.), C*orrecting fallacies about educational and psychological testing* (pp. 67-88). Washington, DC: American Psychological Association. https://doi.org/10.1037/11861-002

Cashel, M. L. (2002). Child and adolescent psychological assessment: Current clinical practices and the impact of managed care. *Professional Psychology: Research and Practice*, *33*(5), 446-453. https://doi.org/10.1037/0735-7028.33.5.446

Castellanos-Ryan, N., Briere, F. N., O'Leary-Barrett, M., Banaschewski, T., Bokde, A., Bromberg, U., … The IMAGEN Consortium. (2016). The structure of psychopathology in adolescence and its common personality and cognitive correlates. *Journal of Abnormal Psychology*, *125*(8), 1039-1052. https://doi.org/10.1037/abn0000193

Charles, C., Gafni, A., & Whelan, T. (1999). Revisiting the shared treatment decision-making model. *Social Science & Medicine (1982)*, *49*(5), 651-661.

Charles, C., Gafni, A., & Whelan, T. (1999). Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Social Science & Medicine*, *49*(5), 651-661. https://doi.org/10.1016/S0277-9536(99)00145-8

Chow, J. C. C., Jaffee, K., & Snowden, L. (2003). Racial/Ethnic disparities in the use of mental health services in poverty areas. *American Journal of Public Health*, *93*(5), 792-797. https://doi.org/10.2105/AJPH.93.5.792

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284-290. https://doi.org/10.1037/1040-3590.6.4.284

Clay, R. A. (2010). Ensuring that practitioners are competent. Retrieved from http://www.apa.org/monitor/2010/06/practitioners

Comer, J. S., & Kendall, P. C. (2004). A symptom-level examination of parent-child agreement in the diagnosis of anxious youths. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*(7), 878-886. https://doi.org/10.1097/01.chi.0000125092.35109.c5

Cook, J. M., & Coyne, J. C. (2005). Re-envisioning the training and practice of clinical psychologists: Preserving science and research orientations in the face of change. *Journal of Clinical Psychology*, 61(9), 1191-1196. https://doi.org/10.1002/jclp.20163

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4*(1), 5-13. https://doi.org/10.1037/1040-3590.4.1.5

Cronbach, L. J. (1949). *Essentials of psychological testing*. New York, NY: Harper.

Croskerry, P. (2009). A universal model of diagnostic reasoning. *Academic Medicine*, *84*(8), 1022-1028.

Croskerry, P. (2015). Clinical decision making. In *Pediatric and congenital cardiac care: Volume 2: Quality improvement and patient safety*. https://doi.org/10.1007/978-1-4471-6566-8_33

Croskerry, P., & Norman, G. (2008). Overconfidence in clinical decision making. *American Journal of Medicine, 121*(5), S24-S29. https://doi.org/10.1016/j.amjmed.2008.02.001

Curet, M. J., Darosa, D., & Mennin, S. (1999). University and practice-based physicians' input on the content of a surgical curriculum. *American Journal of Surgery*, *178*(1), 78-84. https://doi.org/10.1016/S0002-9610(99)00119-1

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11(2013)*, 126. https://doi.org/10.1186/1741-7015-11-126

Dana, R. H. (1998). *Understanding cultural identity in intervention and assessment. Understanding cultural identity in intervention and assessment.* Thousand Oaks, CA: Sage Publications, Inc. https://doi.org/10.4135/9781483328225

Daryanani, R., Hindley, P., Evans, C., Fahy, P., & Turk, J. (2001). Ethnicity and use of a child and adolescent mental health service. *Child Psychology and Psychiatry Review*, *6*(3), 127-132. https://doi.org/10.1017/S1360641701002659

Davis, J., & Bauman, K. (2013). School enrollment in the United States: 2011. *United States Census Bureau*, (September), 1-14. Retreived https://www.census.gov/prod/2013pubs/p20-571.pdf

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*(4), 483-509. https://doi.org/10.1037/0033-2909.131.4.483

DeSilva, R., Aggarwal, N. K., & Lewis-Fernández, R. (2015). The DSM-5 cultural formulation interview and the evolution of cultural assessment in psychiatry. *Psychiatric Times*, *32*(6), 5-8.

Dhami, M. K., & Mumpower, J. L. (2018). Kenneth R. Hammond's contributions to the study of judgment and decision making. *Judgment and Decision Making*, *13*(1), 1-22.

Dodge, K., & Pettit, G. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology*, *39*(2), 349-371. https://doi.org/10.1037/0012-1649.39.2.349

Domino, G., & Domino, M. L. (2006). The history of psychological testing. In *Psychological testing: An introduction* (pp. 517-534). Cambridge: Cambridge University Press.

Ebesutani, C., Bernstein, A., Chorpita, B. F., & Weisz, J. R. (2012). A transportable assessment protocol for prescribing youth psychosocial treatments in real-world settings: Reducing assessment burden via self-report scales. *Psychological Assessment*, *24*(1), 141-155. https://doi.org/10.1037/a0025176

Eignor, D. R. (2013). The standards for educational and psychological testing. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbooks in psychology. APA handbook of testing and assessment in psychology, Vol. 1. Test theory and testing and assessment in industrial and organizational psychology* (pp. 245-250). American Psychological Association. https://doi.org/10.1037/14047-013

Eisman, E. J., Finn, S. E., Kay, G. G., Meyer, G. J., Dies, R. R., Eyde, L. D., … Moreland, K. L. (2000). Problems and limitations in using psychological assessment in the contemporary health care delivery system. *Professional Psychology: Research and Practice*, *31*(2), 131-140. https://doi.org/10.1037/0735-7028.31.2.131

Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., … Barry, M. (2012). Shared decision making: A model for clinical practice. *Journal of General Internal Medicine, 27*(2012), 1361-1367. https://doi.org/10.1007/s11606-012-2077-6

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, *196*(4286), 129-136. https://doi.org/10.1126/science.847460

Evans, F. B., & Finn, S. E. (2017). Training and consultation in psychological assessment with professional psychologists: Suggestions for enhancing the profession and individual practices. *Journal of Personality Assessment*, *99*(2), 175-185. https://doi.org/10.1080/00223891.2016.1187156

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: a systematic literature review. *European Child & Adolescent Psychiatry*, *22*(6), 329-340. https://doi.org/10.1007/s00787-013-0375-0

Falzer, P. R. (2018). Naturalistic decision making and the practice of health care. *Journal of Cognitive Engineering and Decision Making*, *12*(3), 178-193. https://doi.org/10.1177/1555343418773915

Fernández-Ballesteros, R., De Bruyn, E. E. J., Godoy, A., Hornke, L. F., Ter Laak, J., Vizcarro, C., … Zaccagnini, J. L. (2001). Guidelines for the assessment process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, *17*(3), 187-200. https://doi.org/10.1027//1015-5759.17.3.187

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 506-522. https://doi.org/10.1207/s15374424jccp3403_7

Fletcher, K. E. (2005). Encyclopedia of psychological assessment. *Psychiatric Services, 56*(5), 614-615. https://doi.org/10.1176/appi.ps.56.5.614-a

Frick, P. J. (2007). Providing the evidence for evidenced-based practice. *Journal of Clinical Child and Adolescent Psychology*, *36*(1), 2-7. https://doi.org/10.1080/15374410709336563

Garb, H. N. (1997). Race bias, social class bias, and gender bias in clinical judgment. *Clinical Psychology: Science and Practice*, *4*(2), 99-120. https://doi.org/10.1111/j.1468-2850.1997.tb00104.x

Garb, H. N. (1998). Improving psychological assessment. In H. N. Garb, *Studying the clinician: Judgment research and psychological assessment* (pp. 231-248). Washington, DC: American Psychological Association. https://doi.org/10.1037/10299-009

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology*, *1*(1), 67-89. https://doi.org/10.1146/annurev.clinpsy.1.102803.143810

Garb, H. N. (2013). Cognitive and social factors influencing clinical judgment in psychiatric practice. *World Psychiatry*, *12*(2), 108-110. https://doi.org/10.1002/wps.20045

Garb, H. N., Lilienfeld, S. O., Nezworski, M. T., Wood, J. M., & O'Donohue, W. T. (2009). Can quality improvement processes help psychological assessment meet the demands of evidence-based practice? *Scientific Review Mental Health Practice*, *7*(1), 17-25.

Geisinger, K. F. (2000). Psychological testing at the end of the millennium: A brief historical review. *Professional Psychology: Research and Practice, 31*(2), 117-118. https://doi.org/10.1037/0735-7028.31.2.117

Getz, G. E. (2011). Culture fair test. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of clinical neuropsychology* (pp. 755-756). New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-79948-3_1186

Gibby, R. E., & Zickar, M. J. (2008). A history of the early days of personality testing in American industry: An obsession with adjustment. *History of Psychology*, *11*(3), 164-184. https://doi.org/10.1037/a0013041

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*(2011), 451-482. https://doi.org/10.1146/annurev-psych-120709-145346

Goodenough, F. L. (1949). *Mental testing, its history, principles, and applications. Mental testing, its history, principles, and applications.* Oxford, England: Rinehart.

Gottfredson, L. J., & Saklofske, D. H. (2009). Intelligence: Foundations and issues in assessment. *Canadian Psychology, 50*(3), 183. https://doi.org/10.1037/a0016641

Gregory, R. J. (2004). *Psychological testing: history, principles, and applications* (4th ed.). Boston, MA: Pearson.

Gregory, R. J. (2015). *Psychological testing: history, principles, and applications* (7th ed.). Boston, MA: Pearson.

Gresham, F. M., & Elliott, S. N. (1984). Assessment and classification of children's social skills: A review of methods and issues. *School Psychology Review*, *13*(3), 292-301.

Hagbaghery, M. A., Salsali, M., & Ahmadi, F. (2004). The factors facilitating and inhibiting effective clinical decision-making in nursing: a qualitative study. *BMC Nursing*, *3*(1), 1-11. https://doi.org/10.1186/1472-6955-3-2

Hammond, J. S., Keeney, R. L., & Raiffa, H. (2006). The hidden traps in decision making. *Harvard Business Review,76*(5), 47-48.

Harrell, E. (2017). A brief history of personality tests. *Harvard Business Review*, *95*(2), 63.

Harvey, V. S. (2006). Variables affecting the clarity of psychological reports. *Journal of Clinical Psychology*, *62*(1), 5-18. https://doi.org/10.1002/jclp.20196

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychosocial assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238-247. http://dx.doi.org/10.1037/1040-3590.7.3.238

Heffer, R. W., Barry, T. D., & Garland, B. H. (2009). History, overview, and trends in child and adolescent psychological assessment. In J. L. Matson, F. Andrasik, & M. L. Matson (Eds.), *Assessing childhood psychopathology and developmental disabilities* (pp. 3-29). New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-09528-8_1

Hersen, M, & Reitman, D. (2008). *Handbook of psychological assessment, case conceptualization, and treatment, Volume 2: Children and adolescents*. (M. Hersen & D. Reitman, Eds.). New York, NY: John Wiley & Sons Inc.

Higgins, L. T., & Sun, C. H. (2001). The development of psychological testing in China. *International Journal of Psychology*, *37*(4), 246-254. https://doi.org/10.1080/00207590244000025

Hogarth, R. M. (2010). Intuition: A challenge for psychological research on decision making. *Psychological Inquiry*, *21*(4), 338-353. https://doi.org/10.1080/1047840X.2010.520260

Holm-Hansen, C. (2006). *Racial and ethnic disparities in children's mental health*. Saint Paul, MN: Wilder Research, 1-15. Retrieved from https://www.wilder.org/sites/default/files/imports/TroubledYouth_10-06.pdf

Hopwood, C. J., & Bornstein, R. F. (2014). *Multimethod clinical assessment*. (C. J. Hopwood & R. F. Bornstein, Eds.). New York, NY: Guilford Publications.

Horin, E. V., Hernandez, B., & Donoso, O. A. (2012). Behind closed doors: Assessing individuals from diverse backgrounds. *Journal of Vocational Rehabilitation*, *37*(2), 87-97. https://doi.org/10.3233/JVR-2012-0602

Hunsley, J. (2009). Advancing the role of assessment in evidence-based psychological practice. *Clinical Psychology: Science and Practice*, *16*(2), 202-205. https://doi.org/10.1111/j.1468-2850.2009.01159.x

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, *3*(1), 29-51. https://doi.org/10.1146/annurev.clinpsy.3.022806.091419

Hunsley, J., & Mash, E. J. (2008). Developing criteria for evidence-based assessment: An introduction to assessments that work. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (1st ed., pp. 3-16). New York, NY: Oxford University Press. https://doi.org/10.1093/med-psych/9780190492243.003.0001

Hunsley, J., & Mash, E. J. (2014). Evidence-based assessments. In D. H. Barlow (Ed.), *The Oxford handbook of clinical psychology* (pp. 76-97). New York, NY: Oxford University Press.

Hunsley, J., & Mash, E. J. (2018). *A guide to assessments that work* (2nd ed.). (J. Hunsley & E. J. Mash, Eds.). New York, NY: Oxford University Press.

Hunt, W. A. (1946). The future of diagnostic testing in clinical psychology. *Journal of Clinical Psychology*, *2*(4), 311-317. https://doi.org/10.1021/es00094a612

Hurt S., & Tomoyasu N. (1995) Psychological testing and assessment. In G. K. Zammit & W. Hull (Eds.), *Guidebook for clinical psychology interns*, a*pplied clinical psychology* (pp. 121-134). Boston, MA: Springer. https://doi.org/10.1007/978-1-4899-0222-1_8

Hyman, S. E. (2011). Diagnosing the DSM. *Cerebrum: The Dana Forum on Brain Science*. Retrieved from https://dana.org/article/diagnosing-the-dsm/

Institute of Medicine. (2000). *To err is human: building a safer health system*. (L. T. Kohn, J.M.Corrigan, & M. S. Donaldson, Eds.). Washington, DC: National Academy Press. https://doi.org/10.17226/9728

Ivnik, R. J., Smith, G. E., Petersen, R. C., Boeve, B. F., Kokmen, E., & Tangalos, E. G. (2000). Diagnostic accuracy of four approaches to interpreting neuropsychological test data. *Neuropsychology, 14*(2), 163-177. https://doi.org/10.1037/0894-4105.14.2.163

Jacobs, V. R., & Fischer, T. (2012). A pragmatic guide on how physicians can take over financial control of their clinical practice. *Journal of the Society of Laparoendoscopic Surgeons*, *16*(4), 632-638. https://doi.org/10.4293/108680812X13517013316438

Jensen-Doss, A., & Hawley, K. M. (2010). Understanding barriers to evidence-based assessment: Clinician attitudes toward standardized assessment tools. *Journal of Clinical Child and Adolescent Psychology*, *39*(6), 885-896. https://doi.org/10.1080/15374416.2010.517169

Jensen-Doss, A., & Hawley, K. M. (2011). Understanding clinicians' diagnostic practices: Attitudes toward the utility of diagnosis and standardized diagnostic tools. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*(6), 476-485. https://doi.org/10.1007/s10488-011-0334-3

Jensen-Doss, A., & Weisz, J. R. (2008). Diagnostic agreement predicts treatment process and outcomes in youth mental health clinics. *Journal of Consulting and Clinical Psychology, 76*(5), 711-722. https://doi.org/10.1037/0022-006X.76.5.711

Kamp-Becker, I., Albertowski, K., Becker, J., Ghahreman, M., Langmann, A., Mingebach, T., … Stroth, S. (2018). Diagnostic accuracy of the ADOS and ADOS-2 in clinical practice. *European Child and Adolescent Psychiatry*, *27*(9), 1193-1207. https://doi.org/10.1007/s00787-018-1143-y

Kamphaus, R. W., Petoskey, M. D., & Rowe, E. W. (2000). Current trends in psychological testing of children. *Professional Psychology: Research and Practice*, *31*(2), 155-164. https://doi.org/10.1037/0735-7028.31.2.155

Kaplan, R. M., & Saccuzzo, D. P. (2008). *Psychological testing: Principles, applications, and issues* (7th ed.). San Francisco, CA: Cengage Learning.

Kaslow, N. J., Borden, K. A., Collins Jr, F. L., Forrest, L., Illfelder-Kaye, J., Nelson, P. D., … Willmuth, M. E. (2004). Competencies conference: Future directions in education and credentialing in professional psychology. *Journal of Clinical Psychology*, *60*(7), 699-712.

Kazdin, A. E. (2005). Evidence-based assessment for children and adolescents: Issues in measurement development and clinical application. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 548-558. https://doi.org/10.1207/s15374424jccp3403_10

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, *63*(3), 146-159. https://doi.org/10.1037/0003-066X.63.3.146

Kelley, M. F., Sexton, D., & Surbeck, E. (1990). Traditional psychometric assessment approaches. In *Comprehensive assessment in special education: Approaches, procedures and concerns* (pp. 5-28). Springfield, IL: Charles C. Thomas Publisher.

Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Üstün, T. B. (2007). Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry*, *20*(4), 359-364. https://doi.org/10.1097/YCO.0b013e32816ebc8c

Kicklighter, T. H., Geisler, P. R., Barnum, M., Heinerichs, S., & Martin, M. (2018). Exploration of factors perceived to influence development of diagnostic reasoning in athletic trainers and athletic training students. *Athletic Training Education Journal*, *13*(2), 120-130. https://doi.org/10.4085/1302120

Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., … Rahman, A. (2011). Child and adolescent mental health worldwide: Evidence for action. *The Lancet*, *378*(9801), 1515-1525. https://doi.org/10.1016/S0140-6736(11)60827-1

Klein, G. A. (1998). *Sources of power: How people make decisions* (20th anniversary ed.). Cambridge, Massachusetts; London, England: The MIT Press. https://doi.org/10.2307/j.ctt1v2xt08

Klein, G. (2008). Naturalistic decision making. *Human Factors*, *50*(3), 456-460. https://doi.org/10.1518/001872008X288385

Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. *Journal of Applied Research in Memory and Cognition*, *4*(3), 164-168. https://doi.org/10.1016/j.jarmac.2015.07.001

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, *4*(3), 186-209. https://doi.org/10.1518/155534310X12844000801203

Klein, G., Hintze, N., & Saab, D. (2013). Thinking inside the box: The ShadowBox method for cognitive skill development. *Proceedings of the 11th International Conference on Naturalistic Decision Making*, *Marseilles, France*, 2013. Retrieved from https://www.researchgate.net/publication/269335857_Thinking_Inside_the_Box_The_ShadowBox_Method_for_Cognitive_Skill_Development

Klein, G., & Hoffman, R. (1992). Seeing the invisible: Perceptual-cognitive aspects of expertise. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 203-226). Mahwah, NJ: Erlbaum.

Klein, G., & Wright, C. (2016). Macrocognition: From theory to toolbox. *Frontiers in Psychology*, *7*(54), 1-5. https://doi.org/10.3389/fpsyg.2016.00054

Kinder, B. (1994). Where the action is in personality assessment. *Journal of Personality Assessment,* 62(3), 585-588.

Knauss, L. K. (2001). Ethical issues in psychological assessment in school settings. *Journal of Personality Assessment*, *77*(2), 231-241. https://doi.org/10.1207/S15327752JPA7702_06

Kolen, M. J., & Hendrickson, A. B. (2013). Scaling, norming, and equating. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbooks in psychology. APA handbook of testing and assessment in psychology, Vol. 1. Test theory and testing and assessment in industrial and organizational psychology* (pp. 201-222). Washington, DC: American Psychological Association. https://doi.org/10.1037/14047-011

Krishnamurthy, R., VandeCreek, L., Kaslow, N. J., Tazeau, Y. N., Miville, M. L., Kerns, R., … Benton, S. A. (2004). Achieving competency in psychological assessment: Directions for education and training. *Journal of Clinical Psychology*, *60*(7), 725-739. https://doi.org/10.1002/jclp.20010

Leach, M. M., & Oakland, T. (2007). Ethics standards impacting test development and use: A review of 31 ethics codes impacting practices in 35 countries. *International Journal of Testing*, *7*(1), 71-88, https://doi.org/10.1080/15305050709336859

Leadbeater, B., Thompson, K., & Gruppuso, V. (2012). Co-occurring trajectories of symptoms of anxiety, depression, and oppositional defiance from adolescence to young adulthood. *Journal of Clinical Child and Adolescent Psychology*, *41*(6), 719-730. https://doi.org/10.1080/15374416.2012.694608

Legare, F., Ratte, S., Gravel, K., & Graham, I. D. (2008). Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions. *Patient Education and Counseling*, *73*(3), 526-535. https://doi.org/10.1016/j.pec.2008.07.018

Lemberger, M. E., Morris, C. A. W., Clemens, E. V., & Smith, A. L. (2013). A qualitative investigation of the referral process from school counselors to mental health providers. *Journal of Chemical Information and Modeling*, *53*(9), 1689-1699. https://doi.org/10.1017/CBO9781107415324.004

Leong, F. T. L., Park, Y. S., & Leach, M. M. (2013). Ethics in psychological testing and assessment. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbooks in psychology. APA handbook of testing and assessment in psychology, Vol. 1. Test theory and testing and assessment in industrial and organizational psychology* (pp. 265-282). Washington, DC: American Psychological Association. https://doi.org/10.1037/14047-015

Liang, J., Matheson, B. E., & Douglas, J. M. (2016). Mental health diagnostic considerations in racial/ethnic minority youth. *Journal of Child and Family Studies*, *25*(6), 1926-1940. https://doi.org/10.1007/s10826-015-0351-z

Lilienfeld, S. O., & Lynn, S. J. (2015). Errors/Biases in clinical decision making. In R. L. Cautin & S. O. Lilienfeld. (Eds.), *The encyclopedia of clinical psychology* (1st ed., pp. 1-9). New York, NY: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118625392.wbecp567

Lin, L., & Stamm, K. (2016). Where are the highest concentrations of licensed psychologists? *Monitor on Psychology*, *47*(3). Retrieved https://www.apa.org/monitor/2016/03/datapoint

Lovitt, R. (1974). Deficits in the skill development of clinical psychologists. *Professional Psychology*, *5*(4), 415-420. https://doi.org/10.1037/h0021330

Loyens, S., Kirschner, P. A., & Paas, F. (2012). Problem-based learning. In K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, & H. L. Swanson (Eds.), *APA handbooks in psychology. APA educational psychology handbook, Vol. 3. Application to learning and teaching* (pp. 403-425). Washington, DC: American Psychological Association. https://doi.org/10.1037/13275-016

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., … Lord, C. (2009). The autism diagnostic observation schedule - toddler module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(9), 1305-1320. https://doi.org/10.1007/s10803-009-0746-z

Magnavita, J. J. (2016). *Clinical decision making in mental health practice*. American Psychological Association.

Majnemer, A. (1998). Benefits of early intervention for children with developmental disabilities. *Seminars in Pediatric Neurology*, *5*(1), 62-69. https://doi.org/https://doi.org/10.1016/S1071-9091(98)80020-X

Makary, M. A., & Daniel, M. (2016). Medical error - the third leading cause of death in the US. *BMJ*, *353*, i2139. https://doi.org/10.1136/bmj.i2139

Makoul, G., & Clayman, M. L. (2006). An integrative model of shared decision making in medical encounters. *Patient Education and Counseling*, *60*(3), 301-312. https://doi.org/10.1016/j.pec.2005.06.010

Mash, E. J., & Hunsley, J. (2005). Special section: Developing guidelines for the evidence-based assessment of child evidence-based assessment of child and adolescent disorders: Issues and challenges. *Joumal of Clinical Child and Adolescent Psychology*, *34*(3), 362-379. https://doi.org/10.1207/s15374424jccp3403

Masling, J. M. (1992). What does it all mean? In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness: Cognitive, clinical, and social perspectives* (pp. 259-276). New York, NY: Guilford Press.

McGinnis, P. Q., Hack, L. M., Nixon-Cave, K., & Michlovitz, S. L. (2009). Factors that influence the clinical decision making of physical therapists in choosing a balance assessment approach. *American Physical Therapy Association*, *89*(3), 233-247.

McIntire, S. A., & Miller, L. A. (2007). *Foundations of psychological testing: A practical approach* (2nd ed.). Newbury Park, CA: SAGE Publications.

McReynolds, P., & Ludwig, K. (1987). On the history of rating scales. *Personality and Individual Differences*, *8*(2), 281-283. https://doi.org/10.1016/0191-8869(87)90188-7

Mechanic, D., Angel, R. J., & Davies, L. (1992). Correlates of using mental health services: Implications of using alternative definitions. *American Journal of Public Health*, *82*(1), 74-78. https://doi.org/10.2105/AJPH.82.1.74

The Mental Health Foundation. (2016). *Fundamental facts about mental health 2016*. Mental Health Foundation. Retrieved from https://www.mentalhealth.org.uk/sites/default/files/fundamental-facts-about-mental-health-2016.pdf

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the national comorbidity survey replication-adolescent supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(10), 980-989. https://doi.org/10.1016/j.jaac.2010.05.017

Merten, E. C., Cwik, J. C., Margraf, J., & Schneider, S. (2017). Overdiagnosis of mental disorders in children and adolescents (in developed countries). *Child and Adolescent Psychiatry and Mental Health*, *11*(1), 1-11. https://doi.org/10.1186/s13034-016-0140-5

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749. https://doi.org/10.1037/0003-066X.50.9.741

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., … Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*(2), 128-165. https://doi.org/10.1037//OOO3-O66X.56.2.128

Mihura, J. L., Roy, M., & Graceffo, R. A. (2017). Psychological assessment training in clinical psychology doctoral programs. *Journal of Personality Assessment*, *99*(2), 153-164. https://doi.org/10.1080/00223891.2016.1201978

Miller, L. A., & Lovler, R. L. (2018). *Foundations of psychological testing: A practical approach* (6th ed.). Newbury Park, CA: SAGE Publications.

Miller, R. H., & Luft, H. S. (1997). Does managed care lead to better or worse quality of care? *Health Affairs*, *16*(5), 7-25. https://doi.org/10.1377/hlthaff.16.5.7

Minton, H. L. (1988). *The American social experience series, 11. Lewis M. Terman: Pioneer in psychological testing.* New York, NY: New York University Press

Mischel, W. (1968). *Personality and assessment.* New York, NY: John Wiley & Sons Inc.

Morey, L. C., & Benson, K. T. (2016). An investigation of adherence to diagnostic criteria, revisited: Clinical diagnosis of the DSM-IV/DSM-5 section II personality disorders. *Journal of Personality Disorders*, *30*(1), 130-144. https://doi.org/10.1521/pedi_2015_29_188

Mukolo, A., Heflinger, C. A., & Wallston, K. A. (2010). The stigma of childhood mental disorders: A conceptual framework. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(2), 92-103. https://doi.org/10.1097/00004583-201002000-00003

Mülberger, A. (2017). Mental association: Testing individual differences before Binet. *Journal of the History of the Behavioral Sciences*, *53*(2), 176-198.

Muntean, W. J. (2011). *Nursing clinical decision-making: A literature review*. Internal NCSBN Literature Review. Unpublished.

Naglieri, J. A. (2013). Psychological assessment by school psychologists: Opportunities and challenges of a changing landscape. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbooks in psychology. APA handbook of testing and assessment in psychology, Vol. 3. Testing and assessment in school psychology and education* (pp. 3-19). Washington, DC: American Psychological Association. https://doi.org/10.1037/14049-001

National Academies of Sciences, Engineering, and Medicine; Institute of Medicine; Board on Health Care Services; & Committee on Diagnostic Error in Health Care. (2015). *Improving diagnosis in health care*. (B. T. Balogh, J. R. Miller, & J. R. Ball, Eds.). Washington, DC: National Academies Press. https://doi.org/10.17226/21794.

National Alliance on Mental Health [NAMH]. (2019). Mental health by the numbers. Retrieved from https://www.nami.org/learn-more/mental-health-by-the-numbers%0Ahttps://www.nami.org/Learn-More/Mental-Health-By-the-Numbers

National Center for Education Statistics. (2019). The condition of education - elementary and secondary education - finance - public school expenditures - indicator May (2015). Retrieved June 12, 2019, from https://nces.ed.gov/programs/coe/indicator_cgg.asp

Nelson-Gray, R. O. (2003). Treatment utility of psychological assessment. *Psychological Assessment*, *15*(4), 521-531. https://doi.org/10.1037/1040-3590.15.4.521

Norcross, J. C. (1991). Prescriptive matching in psychotherapy: An introduction. *Psychotherapy*, *28*(3), 439-443. https://doi.org/10.1037/0033-3204.28.3.439

Oakland, T. (2004). Selected ethical issues relevant to test adaptations. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 77-104). London: Psychology Press.

Oakland, T., Poortinga, Y. H., Schlegel, J., & Hambleton, R. K. (2001). International Test Commission: Its history, current status, and future directions. *International Journal of Testing*, *1*(1), 3-32. https://doi.org/10.1207/S15327574IJT0101_2

Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J., Sanford, M., & Lipman, E. L. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(8), 1078-1085. https://doi.org/10.1097/00004583-199608000-00019

Owens, P. L., Hoagwood, K., Horwitz, S. M., Leaf, P. J., Poduska, J. M., Kellam, S. G., & Ialongo, N. S. (2002). Barriers to children's mental health services. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(6), 731-738. https://doi.org/10.1097/00004583-200206000-00013

Palmiter, D. J. (2004). A survey of the assessment practices of child and adolescent clinicians. *American Journal of Orthopsychiatry*, *74*(2), 122-128. https://doi.org/10.1037/0002-9432.74.2.122

Panter, J. E., & Bracken, B. A. (2013). Preschool assessment. In *APA handbook of testing and assessment in psychology, Vol. 3: Testing and assessment in school psychology and education.* https://doi.org/10.1037/14049-002

Pastor, P. N., & Reuben, C. A. (2005). Racial and ethnic differences in ADHD and LD in young school-age children: Parental reports in the national health interview survey. *Public Health Reports*, *120*(4), 383-392. https://doi.org/10.1177/003335490512000405

Pelco, L. E., Ward, S. B., Coleman, L., & Young, J. (2009). Teacher ratings of three psychological report styles. *Training and Education in Professional Psychology, 3*(1), 19-27. https://doi.org/10.1037/1931-3918.3.1.19

Perepletchikova, F., & Kazdin, A. E. (2005). Oppositional defiant disorder and conduct disorder. In K. Cheng & K. M. Myers (Eds.), *Child and adolescent psychiatry: The essentials* (pp. 73-88). Philadelphia, PA: Lippincott Williams & Wilkins.

Perversi, P., Yearwood, J., Bellucci, E., Stranieri, A., Warren, J., Burstein, F., … Wolff, A. (2018). Exploring reasoning mechanisms in ward rounds: A critical realist multiple case study. *BMC Health Services Research*, *18*(1), 1-11. https://doi.org/10.1186/s12913-018-3446-6

Phillips, S. E. (1993). *Legal implications of high-stakes assessment: What states should know* (pp. 63-89). Oak Brook, IL: North Central Regional Educational Laboratory.

Pikulski, J. J. (1990). The role of tests in a literacy assessment program. *The Reading Teacher*, *43*(9), 686-688.

Piotrowski, C. (1999). Assessment practices in the era of managed care: Current status and future directions. *Journal of Clinical Psychology*. New York, NY: John Wiley & Sons Inc. https://doi.org/10.1002/(SICI)1097-4679(199907)55:7<787::AID-JCLP2>3.0.CO;2-U

Piotrowski, C., Belter, R. W., & Keller, J. W. (1998). The impact of "managed care" on the practice of psychological testing: Preliminary findings. *Journal of Personality Assessment*, *70*(3), 441-447.

Plake, B. S., & Wise, L. L. (2014). What is the role and importance of the revised AERA, APA, NCME standards for educational and psychological testing? *Educational Measurement: Issues and Practice*, *33*(4), 4-12. https://doi.org/10.1111/emip.12045

Pliske, R., & Klein, G. (2003). The naturalistic decision-making perspective. In J. Shanteau & S. L. Schneider (Eds.), *Emerging perspectives on judgment and decision research* (pp. 559-585). Cambridge, United Kingdom: Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511609978.019

Podgorelec, V., Kokol, P., & Rozman, I. (2002). Decision trees: An overview and their use in medicine. *Journal of Medical Systems*, *26*(5), 445-463. https://doi.org/10.1023/A

Poland, S. F., Thurlow, M. L., Ysseldyke, J. E., & Mirkin, P. K. (1982). Current psychoeducational assessment and decision-making practices as reported by directors of special education. *Journal of School Psychology*, *20*(3), 171-179. https://doi.org/10.1016/0022-4405(82)90046-2

Pope, K., & Vaquez, M. J. T. (2016). *Ethics of psychotherapy and counseling: A practical guide* (5th ed.). Hoboken, NJ: John Wiley & Sons Inc.

Posada, M. M. (2004). Ethical issues in assessments with infants and children. *Graduate Student Journal of Psychology 2004*, *6*, 42-48.

Prinstein, M. J., Youngstrom, E. A., Mash, E. J., & Barkley, R. A. (2019). *Treatment of disorders in childhood and adolescence* (4th ed.). New York, NY: Guilford Publications.

Psihogios, A. M., Gutiérrez-Colina, A. M., Iskander, J. M., Wasserman, R. M., & Ramsey, R. R. (2019). Adherence knowledge and education needs among pediatric psychology trainees: A needs assessment and recommendations. *Clinical Practice in Pediatric Psychology*, *8*(1), 45-55. https://doi.org/10.1037/cpp0000306

Ready, R. E., & Veague, H. B. (2014). Training in psychological assessment: Current practices of clinical psychology programs. *Professional Psychology: Research and Practice*, *45*(4), 278-282. https://doi.org/10.1037/a0037439

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, *2*(3), 1-18. https://doi.org/10.1177/2332858416657343

Redelmeier, D. A., Ferris, L. E., Tu, J. V., Hux, J. E., & Schull, M. J. (2001). Problems for clinical judgement: Introducing cognitive psychology as one more basic science. *CMAJ: Canadian Medical Association Journal, 164*(3), 358-360.

Redman, B. K. (2006). Review of measurement instruments in clinical and research ethics, 1999-2003. *Journal of Medical Ethics, 2006*(32), 153-156. https://doi.org/10.1136/jme.2005.012567

Regier, D. A, Narrow, W. E., Kuhl, E. A., & Kupfer, D. J. (2011). *The conceptual evolution of DSM-5*. (American Psychopathological Association, Ed.). Washington, DC: American Psychiatric Publishing.

Reyna, V. F. (2008). Theories of medical decision making and health: An evidence-based approach. *Medical Decision Making:An International Journal of the Society for Medical Decision Making*, *28*(6), 829-833. https://doi.org/10.1177/0272989X08327069

Reynolds, C. R., & Suzuki, L.A. (2012). Bias in psychological assessment. In Weiner, I., Graham, J.R., & Naglieri, J.A. (Eds.), *Handbook of psychology* (2nd ed.). New York, NY: John Wiley & Sons, Inc. https://doi.org/doi:10.1002/9781118133880.hop210004

Riccio, C. A., & Rodriguez, O. L. (2007). Integration of psychological assessment approaches in school psychology. *Psychology in the Schools*, *44*(3), 243-255. https://doi.org/10.1002/pits.20220

Richardson, R., Trépel, D., Perry, A., Ali, S., Duffy, S., Gabe, R., … McMillan, D. (2015). Screening for psychological and mental health difficulties in young people who offend: A systematic review and decision model. *Health Technology Assessment*, *19*(1), 1-158. https://doi.org/10.3310/hta19010

Riggin, B. M., & Lack, C. W. (2018). Ethical decision-making models across mental health treatment: A review and expansion. *Current Psychiatry Reviews*, *14*(3), 171-177. https://doi.org/10.2174/1573400514666180816112109

Roysircar, G. (2005). Culturally sensitive assessment, diagnosis, and guidelines. In *Strategies for building multicultural competence in mental health and educational settings* (pp. 19-38). Hoboken, NJ: John Wiley & Sons Inc.

Russo, S. (2015). Foundations of behavioral, social, and clinical assessment of children (6 ed.; including resource guide). *Journal of Psychologists and Counsellors in Schools*, *25*(1), 128-132. https://doi.org/http://dx.doi.org/10.1017/jgc.2015.11

Rylander, M., & Guerrasio, J. (2015). Heuristic errors in clinical reasoning. *The Clinical Teacher*, *13*(2), 287-290. https://doi.org/https://doi.org/10.1111/tct.12444

Sallis, H., Szekely, E., Neumann, A., Jolicoeur-Martineau, A., van IJzendoorn, M., Hillegers, M., … Evans, J. (2019). General psychopathology, internalising and externalising in children and functional outcomes in late adolescence. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *60*(11), 1183-1190. https://doi.org/10.1111/jcpp.13067

Sanders, J. D., & Katz, S. (2013). The overuse and misuse of psychological testing: Why less is more. *American Journal of Family Law*, *24*(2), 213-219.

Sattler, J. M. (1982). *Assessment of children's intelligence and special abilities* (2nd ed.). Boston, MA: Allyn and Bacon

Sattler, J. M. (1988). *Assessment of children*. *Assessment of children* (3rd ed.). San Diego, CA: Jerome M. Sattler Publisher Inc.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Jerome M. Sattler Publisher Inc.

Sattler, J. M., & Hoge, R. D. (2006). *Assessment of children: Behavioral, social, and clinical foundations* (5th ed.). San Diego, CA: Jerome M. Sattler Publisher Inc.

Scarr, S. (1994). Culture fair and culture free test. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 322-328). New York, NY: Macmillan.

Schniering, C. A., Hudson, J. L., & Rapee, R. M. (2000). Issues in the diagnosis and assessment of anxiety disorders in children and adolescents. *Clinical Psychology Review*, *20*(4), 453-478. https://doi.org/10.1016/S0272-7358(99)00037-9

Shaban, R. Z. (2005). Theories of clinical judgment and decision-making: A review of the theoretical literature. *Journal of Emergency Primary Health Care*, *3*(1-2). https://doi.org/10.33151/ajp.3.1.308

Shilkofski, N., & Hunt, E. A. (2015). Identification of barriers to pediatric care in limited-resource settings: A simulation study. *Pediatrics*, *136*(6), e1569-e1575. https://doi.org/10.1542/peds.2015-2677

Simon, H. A. (1959). Theories of decision-making and behavioral science. *The American Economic Review*, *49*(3), 253-283.

Sladek, M. R., Umaña-Taylor, A. J., McDermott, E. R., Rivas-Drake, D., & Martinez-Fuentes, S. (2020). Testing invariance of ethnic-racial discrimination and identity measures for adolescents across ethnic-racial groups and contexts. *Psychological Assessment, 32*(6), 509-526. https://doi.org/10.1037/pas0000805

Smith, D. (2003, January). 10 ways practitioners can avoid frequent ethical pitfalls. *Monitor on Psychology*, *34*(1), 50.

Smith, M., Higgs, J., & Ellis, E. (2008). Factors influencing clinical decision making. In *Clinical reasoning in the health professions* (3rd ed., pp. 89-100). Edinburgh, London: Elsevier Health Sciences. https://doi.org/10.4103/0378-6323.45131

Spengler, P. M., Dixon, D. N., Strohmer, D. C., & Shivy, V. A. (1995). A scientist-practitioner model of psychological assessment: Implications for training, practice and research. *The Counseling Psychologist*. https://doi.org/10.1177/0011000095233009

Splinter, R. (2010). Medical decision making. In *Handbook of physics in medicine and biology*. https://doi.org/10.1201/9781420075250

Stiffman, A. R., Pescosolido, B., & Cabassa, L. J. (2004). Building a model to understand youth service access: the gateway provider model. *Mental Health Services Research*, *6*(4), 189-198. https://doi.org/10.1023/b:mhsr.0000044745.09952.33

Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., & Weiss, B. H. (1999). Can autism be diagnosed accurately in children under 3 years? *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(2), 219-226. https://doi.org/DOI: 10.1111/1469-7610.00435

Suhr, J. A. (2015). *Psychological assessment: A problem-solving approach.* New York, NY, US: Guilford Press.

Sundberg, N. D. (1961). The practice of psychological testing in clinical services in the United States. *American Psychologist*, *16*(2), 79-83. https://doi.org/10.1037/h0040647

Swerdlik, M. E., & Cohen, R. J. (2013). *Psychological testing and assessment: An introduction to tests and measurement* (9th ed.). McGraw-Hill Education. https://doi.org/13: 9780767421577

Teresi, J. A., & Jones, R. N. (2013). Bias in psychological assessment and other measures. In *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology* (pp. 139-164). https://doi.org/10.1037/14047-008

Thorndike, R. M., & Thorndike-Christ, T. M. (2010). *Measurement and evaluation in psychology and education* (8th ed.). Upper Saddle River, NJ: Pearson.

Tricare. (2018). TRICARE ABA benefits overview. Retrieved from https://www.tricare.mil/Plans/SpecialPrograms/ACD/QandA

Tuma, J. M., & Pratt, J. M. (1982). Clinical child psychology practice and training: A survey. *Journal of Clinical Child Psychology*, *11*(1), 27-34. https://doi.org/10.1207/s15374424jccp1101_5

The United Nations. (1989). Convention on the rights of the child. *Treaty Series*, *1577*, 3. Retrieved from https://www.refworld.org/docid/3ae6b38f0.html

United States Census Bureau. (2019). Census Bureau reports nearly 77 million students enrolled in U.S. schools. Retrieved from https://www.census.gov/newsroom/press-releases/2019/school-enrollment.html

U.S. Centers for Medicare and Medicaid Services. (2019). Historical. Retrieved from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical

U.S. Department of Education. (2018). U.S. Department of Education, Office of Special Education Programs, Individuals with Disabilities Education Act (IDEA) Section 618 Data Products: State Level Data Files. Retrieved from https://www2.ed.gov/programs/osepidea/618-data/state-level-data-files/index.html

van Meter, A., Youngstrom, E., Youngstrom, J. K., Ollendick, T., Demeter, C., & Findling, R. L. (2014). Clinical decision making about child and adolescent anxiety disorders using the Achenbach system of empirically based assessment. *Journal of Clinical Child and Adolescent Psychology*, *43*(4), 552-565. https://doi.org/10.1080/15374416.2014.883930

Van Ree, S. (2014). Review of foundations of behavioural, social, and clinical assessment of children. *The Australian Educational and Developmental Psychologist*. https://doi.org/10.1017/edp.2014.17

Versi, E. (1992). "Gold standard" is an appropriate term. *BMJ (Clinical Research Ed.)*, *305*(6846), 187. https://doi.org/10.1136/bmj.305.6846.187-b

Wasserman, J. D. (2018). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 3-55). The Guilford Press.

Watkins, C. E. (1991). What have surveys taught us about the teaching and practice of psychological assessment? *Journal of Personality Assessment*, *56*(3), 426-437. https://doi.org/10.1207/s15327752jpa5603_5

Watkins, M. W. (2009). Errors in diagnostic decision making and clinical judgment. In *Handbook of school psychology* (4th ed., pp. 210-229). New York, NY: John Wiley & Sons.

Weiss, L. G., Saklofske, D. H., Holdnack, J. A., Prifitera, A., Weiss, L. G., Saklofske, D. H., … Prifitera, A. (2016). WISC-V: Advances in the assessment of intelligence. *WISC-V Assessment and Interpretation*, 3-23. https://doi.org/10.1016/B978-0-12-404697-9.00001-7

Whaley, A. L. (2001). Cultural mistrust: An important psychological construct for diagnosis and treatment of African Americans. *Professional Psychology: Research and Practice*, *32*(6), 555-562. https://doi.org/10.1037/0735-7028.32.6.555

Wiener, J., & Costaris, L. (2012). Teaching psychological report writing: Content and process. *Canadian Journal of School Psychology*, *27*(2), 119-135. https://doi.org/10.1177/0829573511418484

Williams, J., Palmes, G., Klinepeter, K., Pulley, A., & Foy, J. M. (2005). Referral by pediatricians of children with behavioral health disorders. *Clinical Pediatrics*, *44*(4), 343-349. https://doi.org/10.1177/000992280504400410

Wills, C. E., & Holmes-Rovner, M. (2006). Integrating decision making and mental health interventions research: research directions. *Clinical psychology: A Publication of the Division of Clinical Psychology of the American Psychological Association*, *13*(1), 9-25. https://doi.org/10.1111/j.1468-2850.2006.00002.x

Wolraich, M. L., Hagan, J. F., Allan, C., Chan, E., Davison, D., Earls, M., … Zurhellen, W. (2019). Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *American Academy of Pediatrics*, *144*(4). https://doi.org/10.1542/peds.2019-2528

Wright, C. V., Beattie, S. G., Galper, D. I., Church, A. S., Bufka, L. F., Brabender, V. M., & Smith, B. L. (2017). Assessment practices of professional psychologists: Results of a National survey. *Professional Psychology: Research and Practice*, *48*(2), 73-78. https://doi.org/10.1037/pro0000086

Yoakum, C. S., & Yerkes, R. M. (Eds.). (1920). *Army mental tests.* Henry Holt and Company. https://doi.org/10.1037/11054-000

Young, A. S., & Rabiner, D. (2015). Racial/ethnic differences in parent-reported barriers to accessing children's health services. *Psychological Services*, *12*(3), 267-273. https://doi.org/10.1037/a0038701

Youngstrom, E. A. (2008). Evidence-based strategies for the assessment of developmental psychopathology: Measuring prediction, prescription, and processs. In W. E. Craighead, D. J. Miklowitz, & L. W. Craighead (Eds.), *Psychopathology: History, diagnosis, and empirical foundations* (pp. 34-77). John Wiley & Sons.

Youngstrom, E. A., Choukas-Bradley, S., Calhoun, C. D., & Jensen-Doss, A. (2015). Clinical guide to the evidence-based assessment approach to diagnosis and treatment. *Cognitive and Behavioral Practice*, *22*(1), 20-35. https://doi.org/10.1016/j.cbpra.2013.12.005

Youngstrom, E. A., Freeman, A. J., & Jenkins, M. M. (2008). The assessment of bipolar disorder in children and adolescents. *Child and Adolescent Psychiatric Clinics of North America*, *23*(1), 1-7. https://doi.org/10.1038/jid.2014.371

Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M. L., & Youngstrom, J. K. (2017). Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment. *Clinical Psychology: Science and Practice*, *24*(4), 331-363. https://doi.org/10.1111/cpsp.12207

Zero to Three. (2016). *DC:0-5: Diagnostic classification of mental health and developmental disorders of infancy and early childhood*: *Revised edition* (DC: 0-5). Washington, DC: ZERO TO THREE.

Appendix

Example Guide for Semi-Structured Interview

Descriptive Questions

- Tell me about/describe a child assessment that you have conducted.
- Please describe all the assessment methods you might use when considering an evaluation with a child.
- Are there any methods that you are aware of and do not/never use? Why?
    - Describe
    - After listing methods, then organize into categories
        1. Methods used most often
        2. Methods used sometimes
        3. Methods used rarely
        4. Methods never used
    - Please describe how you decided to organize the methods in this way?
- Describe the difficulties/limitations you come across when trying to perform assessments with children.
    - What strategies do you use to adapt to these challenges?
- Please describe how each of the following factors influence conducting assessments with children.
    - Referral
    - Time
    - Economics
    - Patient factors
    - Environment
    - Empirical data
    - Ethical/Legal considerations
    - Personal factors
    - Governing Bodies
    - Education

Structural Questions

These questions are to be asked at *decision points*. Decision points can be defined as the point when a cognizant process occurs for selecting one option from several that were considered (Klein et al., 2010).

- When you decided to (*decision point*), what did you base your decision on?
- Describe what influenced your choices?
- What made you decide to implement (*decision point*)?
- If aids, resources, guides, or supplemental materials are utilized….
    - When did you use that resource? What happened before?
- Did you consider other possibilities? (Potential *Decision point* – Yes? No?)
    - If Yes- What other possibilities did you consider?
- What made you select that particular (*decision point*)? What was your most important reason for choosing that approach?