1

Overview of Diagnostic and Behavioral Assessment

Bryce D. McLeod, Amanda Jensen-Doss, and Thomas H. Ollendick

"Assessment" is the process by which information is gathered, interpreted, and used to produce a clinical description of an individual (Hunsley, 2002). In clinical practice, assessment of the individual can help inform the treatment process from initial intake to progress during treatment until termination. Indeed, assessment techniques can be used for a variety of purposes in psychotherapy, which include (but are not limited to) diagnosing disorders, informing treatment planning, building a case conceptualization, and monitoring and evaluating treatment outcomes. When assessment is used to arrive at an accurate description of the individual, then the treatment itself is presumably made more efficient and effective. However, a clinician needs to use the "right" assessment tools during the different phases of treatment to produce an accurate description of the individual. No single assessment technique can be used to guide every phase of treatment. Instead, a clinician must rely upon a number of tools pulled from both the diagnostic and behavioral assessment traditions.

The time is ripe for a book focused on using assessment to inform treatment. The past two decades have witnessed the rise of the evidence-based assessment movement, which has focused attention upon assessment training and practice (Ollendick, 1999). "Evidence-based assessment" is defined as an approach to clinical evaluation that utilizes science and theory to guide the assessment process (Hunsley & Mash, 2007). A goal of this movement is to develop and promote a set of assessment guidelines to direct research, structure training, and inform clinical practice.

Assessment is a complex process, and the assessment literature is voluminous. In the past, the large number of instruments and practices has made it difficult to compare instruments and to select psychometrically sound assessment tools. To address this problem, Hunsley and Mash (2008) outlined psychometric criteria for judging assessment tools and organized the assessment literature around particular disorders. These efforts represent important advances for the field, as it is now easier for clinicians to identify psychometrically strong assessment tools for routine clinical use.

Despite these recent advances, research on how best to employ and interpret assessment tools to improve treatment is lacking. Very few studies have examined the validity of assessment tools across different populations, or have considered whether incorporating assessment into treatment can improve treatment outcomes (Hunsley & Mash, 2007; Youngstrom, 2008). This means that there is little empirical evidence to guide the assessment process during treatment.

The goal of this book is to provide readers with the knowledge and skills needed to guide the assessment process from intake to termination. Other resources are available to help readers identify particular assessment tools for specific disorders (see Hunsley & Mash, 2008). Our purpose here is to cover areas of knowledge—such as the basics of psychometric theory; the connection between assessment and treatment; and the assessment of target behaviors and other key factors (cognitive, affective, behavioral, contextual)—that are needed to use assessment tools effectively over the course of treatment (Krishnamurthy et al., 2004). We also cover key skill areas, such as target behavior identification, case conceptualization, and tool selection and interpretation, which are relevant to using assessment tools to inform treatment. To gain the right combination of knowledge and skills, clinicians are advised to take graduate-level courses in developmental psychopathology, developmental psychology, psychometric theory, child and family intervention, and culture and diversity, as well as in evidence-based assessment.

TRENDS IN ASSESSMENT TRAINING AND PRACTICE

Assessment in clinical practice falls clearly under the domain of psychology. However, assessment does not presently represent a central focus of the field (Youngstrom, 2008). The past two decades have witnessed an explosion of treatment research resulting from the evidence-based practice movement. These efforts have generated a number of treatments for a wide range of youth emotional and behavioral problems (Barrett & Ollendick, 2004; McLeod & Weisz, 2004; Weisz, Jensen-Doss, & Hawley, 2005). However, research on assessment has not kept pace. As a result, assessment and treatment have become disconnected processes (Youngstrom, 2008).

Unfortunately, in our opinion, this disconnection has adversely affected graduate training in clinical psychology. At present, few graduate programs provide courses on the interrelationship between assessment and treatment (Childs & Eyde, 2002). Most assessment training required by graduate programs focuses on intellectual and personality testing, and sometimes behavioral assessment; relatively few programs require courses on clinical assessment (Childs & Eyde, 2002). Assessment training has been criticized for not adjusting to the new assessment trends. Graduate programs differ in their coverage of the knowledge and skills related to using assessment to inform treatment (Carama, Nathan, & Puente, 2000; Krishnamurthy et al., 2004). As a result, not all graduate students learn how assessment can be used to tailor evidence-based treatments to meet the unique needs of a particular child and her or his family.

This volume is intended to illustrate how assessment can be used to inform treatment for children and adolescents. Our basic premise is that effective treatment depends upon accurate assessment rather than solely subjective decision making. To do this, clinicians must know *when* to utilize various diagnostic and behavioral assessment tools at the various phases of treatment, and *why* these assessment tools are helpful for treatment selection and delivery. Next, we turn to the role assessment plays in the different phases of treatment.

WHAT ROLE DOES ASSESSMENT PLAY IN TREATMENT?

Assessment should directly inform choices about treatment, and different assessment tools are more or less relevant at various treatment stages (see Table 1.1). At each stage, clinicians must select the appropriate "nomothetic" and "idiographic" assessment tools available to them. Nomothetic tools (e.g., diagnostic interviewing and parent rating scales) provide data about where an individual child falls relative to the larger population on a domain of interest. For example, a clinician trying to determine whether a child needs services might use an anxiety rating scale to establish whether that child falls above a clinical cutoff. If the child's score is above this cutoff, this would indicate that the child's level of anxiety exceeds what is considered typical for children of similar age and gender, and treatment may be warranted. On the other hand, idiographic tools (e.g., direct observation and self-monitoring) provide individualized information that is more useful for case conceptualization, treatment planning, and outcome monitoring. For example, a clinician seeking to design an exposure hierarchy for a given child might design a behavioral avoidance activity tailored to the child's specific fears, to determine the exact type of situations most likely to elicit anxiety in that child. As we note below in more detail, nomothetic and idiographic tools play important, and complementary, roles in the different phases of treatment.

TABLE 1.1. The Role Assessment Plays in Different Phases of Treatment

Treatment phase	Definition and purpose
Screening	Brief assessment designed to identify children who have a problem or are at risk for developing a problem without intervention. Screening is used to determine the need for treatment, generate a prognosis, and gather baseline data on symptom severity and potential causal factors.
Diagnosis	Determining whether a child meets formal criteria for a psychiatric disorder. Diagnosis is used to establish treatment need and facilitates case conceptualization, treatment planning, and treatment selection.
Prognosis	A prediction regarding the course of an illness or the likelihood of developing of a problem given the presence of specific risk factors. A prognosis helps determine the need for treatment and informs treatment planning.
Case conceptualization	A set of hypotheses about the causes, antecedents, and maintaining factors of a client's target behaviors. Case conceptualization is a critical component of treatment that informs treatment planning, treatment selection, outcome monitoring, and treatment evaluation.
Treatment planning and selection	Using the case conceptualization to identify therapeutic interventions designed to address produce change in the target behaviors.
Treatment monitoring	Ongoing assessment of core symptoms, causal factors, and maintaining processes, in order to monitor treatment response and to identify changes needed to the treatment plan.
Treatment evaluation	Assessment conducted at the end of treatment, in order to evaluate the impact of treatment.

Screening

Screening is primarily used to identify areas in need of more detailed assessment. It can also be used to determine the need for treatment, generate a prognosis, and gather baseline data on symptom severity and potential causal factors. Brief nomothetic measures can be used to screen for possible diagnoses. Such measures can be used to generate an estimate about the likelihood a child meets diagnostic criteria for one or more diagnoses, and/or to ascertain whether symptoms are above a clinical cutoff. Screening can also be used to gather data to formulate a prognosis. Noting the presence of certain risk factors (e.g., trauma, child abuse, poverty) can help a clinician determine whether a child is likely to develop future problems. Sometimes screening may indicate that treatment is not needed (e.g., the child's score is below a clinical cutoff); however, when screening indicates the need for treatment, then the assessment data can serve as baseline data as well as help identify targets for in-depth assessment.

Diagnosis

Generating an accurate diagnosis is a critical step in treatment. Meeting criteria for a disorder indicates a need for treatment and provides a starting point for identifying target behaviors. In many instances, the actual symptoms of disorders can be operationalized and selected for change. In addition, the developmental psychopathology, treatment, and assessment literatures are all organized around diagnoses, so an accurate diagnosis is frequently fundamental to case conceptualization and treatment planning.

Prognosis

Youngstrom (2008) defines "prognosis" as "the course of illness or the longitudinal outcomes that are likely for individuals affected by a particular condition or showing a particular marker or trait" (p. 46). To formulate a prognosis, a clinician must assess for the presence of risk factors that are associated with particular outcomes and use knowledge from the developmental psychopathology literature to determine the likely outcome for a child. A prognosis may indicate a need for treatment if it is determined that a child will probably develop a disorder without intervention. For example, a young child with an inhibited temperament may not meet diagnostic criteria for a disorder, but may be at increased risk for developing an anxiety disorder. Or a prognosis may guide treatment planning by helping to identify the most important treatment targets for a child experiencing multiple problems. For example, when a child presents with multiple problems, it is often important to give problems that have the greatest potential for adverse long-term outcomes (e.g., symptomatology related to a trauma) the highest priority in the treatment plan.

Case Conceptualization

"Case conceptualization" is defined as a set of hypotheses about the causes, antecedents, and maintaining factors of a client's emotional, interpersonal, and behavior problems (Eells, 2007; McLeod, Jensen-Doss, Wheat, & Becker, in press; Nezu, Nezu, Peacock, & Girdwood, 2004). Case conceptualization is a critical component of treatment, as its hypotheses guide assessment and treatment. Assessment related to case conceptualization includes a focus upon "mediators," "moderators," and "therapy processes." Mediators are factors (e.g., cognitions or physiological processes) that account for change in the target behavior. Moderators are factors (e.g., developmental level, gender, ethnicity, or socioeconomic status) that might influence the course of treatment and/or a target behavior. Therapy processes include client (motivation, involvement) and therapist (treatment integrity, competence) factors that influence the effectiveness of psychotherapy (see Shirk, Reyes, & Crisostomo, Chapter 14, this volume). All

assessment should inform case conceptualization and feed directly into treatment design/planning, outcome monitoring, and treatment evaluation.

Treatment Planning and Selection

Treatment planning and selection are important, and challenging, components of treatment. Children often present for treatment with multiple problems, and clinicians must determine which problems warrant treatment. Numerous evidence-based treatments designed to treat a variety of specific diagnoses (e.g., anxiety disorders, depression, conduct disorder, eating disorders) exist. Treatment planning starts with generating a diagnosis and identifying the treatments designed to treat that disorder. A clinician must then use the case conceptualization to select the "right" evidence-based treatment and then tailor the intervention to meet the needs of the individual and his or her family.

Treatment Monitoring

Once treatment begins, ongoing assessment of core symptoms, causal factors, and therapy processes can be used to monitor treatment response and to identify any changes that may be needed in the treatment plan. In adult psychotherapy, evidence indicates that continual assessment and feedback to the client can improve therapy outcomes (Lambert et al., 2003), and studies are beginning to support this benefit in child psychotherapy as well (Stein, Kogan, Hutchison, Magee, & Sorbero, 2010). Because assessment during treatment can become time-consuming, selective targeting of variables for treatment monitoring that will directly inform treatment planning is important. In some cases, this will require assessment to be focused upon the specific symptoms being treated (e.g., panic attacks), causal variables (e.g., anxiety sensitivity), or therapy processes (e.g., exposures to feared stimuli).

Treatment Evaluation

At the end of treatment, a thorough assessment is warranted. To evaluate the impact of treatment, it is important to determine whether the child still meets diagnostic criteria for the disorder that was the focus of treatment, as this represents an important indicator of clinically significant change. This assessment can also determine level of functioning, which represents another clinically meaningful category. The end-of-treatment evaluation should assess the need for further referrals or interventions as well. Clearly, it is important to select outcome assessment measures that have been demonstrated to be sensitive to change.

WHY FOCUS ON DIAGNOSTIC AND BEHAVIORAL ASSESSMENT?

We focus on both diagnostic and behavioral assessment in this volume, because the theory and tools that are part of each tradition inform different facets of treatment (see Table 1.2). These assessment traditions developed along separate paths. Each approach has different conceptual underpinnings and psychometric strengths. It is important to understand how tools from both approaches can be used to produce a complete picture of a child across treatment.

Diagnostic Assessment

For the purposes of this volume, diagnostic assessment is considered to include tools and techniques designed to generate diagnoses and classify behavior. Diagnostic assessment stems from the medical model of psychopathology, which posits that symptoms (and, by extension, diagnoses) fall into classifiable disorders that express themselves in somewhat uniform

TABLE 1.2. Diagnostic and Behavioral Assessment

	Diagnostic	Behavioral
Goals	Based upon nomothetic principles, which are concerned with the discovery of general laws as applied to a large number of individuals	Based upon idiographic principles, which are focused on mapping out the interactions among variables distinctively patterned in each individual
Uses	Classification and prediction	Monitoring target behaviors and/or antecedent, causal, and maintaining variables
Common methods	Interviews Rating scales	Functional interviews Rating scales Direct observation Self-monitoring
Treatment phases	Screening Prognosis Diagnosis Case conceptualization Treatment monitoring Treatment evaluation	Case conceptualization Treatment monitoring
Psychometric principles	Classical test theory Internal reliability Test–retest reliability Interrater reliability Construct validity Criterion validity	Generalizability theory Interrater reliability Accuracy Construct validity Criterion validity

fashion and are caused by identifiable factors that are internal to the child (e.g., genetics, biology) and external/contextual to the child (e.g., family factors, socioeconomic status, traumatic events). Consistent with this model, most diagnostic tools are based upon nomothetic principles, which are concerned with the discovery of general laws as they are applied to large numbers of individuals (Cone, 1986). The nomothetic approach is said to be variable-centered (i.e., deals with how particular characteristics or traits are distributed in the population). Measures and tools designed according to this tradition also include rating scales and interviews that provide global statements about how the behavior of a particular child compares to that of the larger population.

Nomothetic measures are used primarily for classification and prediction (Barrios & Hartmann, 1986) and are typically designed according to the tenets of classical test theory. This approach views scores on a measure as representative of an underlying construct that cannot be directly assessed. Because nomothetic measures are designed to assess individual differences, the meaning of a score produced by such a measure is derived by comparing it to norms from the general population. Indeed, diagnostic tools classify individuals along categories (diagnoses) that are posited to be consistent across time and situations (Bem & Allen, 1974). Variation in scores across situations, time, or items is considered error. Thus the development of nomothetic measures emphasizes stability, and the measures may not be most appropriate for repeated administration (e.g., weekly outcome monitoring; Foster & Cone, 1995).

Early in treatment, diagnostic tools are appropriate for screening, determining whether a child's behavior is normative, and formulating a prognosis. Once treatment begins, diagnostic tools can be used for treatment evaluation. Using diagnostic tools at the end of treatment allows a clinician to determine whether a child has experienced clinically significant change or returned to a normative developmental trajectory. Diagnostic tools therefore play a number of important roles in treatment.

Behavioral Assessment

The behavioral assessment approach is based upon idiographic principles. This person-centered approach focuses upon the uniqueness of a given individual (Cone, 1986; Ollendick & Hersen, 1984); unlike the nomothetic approach, it focuses upon mapping out the interactions among variables distinctively patterned in each individual. The point of comparison for idiographic measures is the child's own behavior across situations and/or time. As Mischel (1968, p. 190) observed over 40 years ago, "Behavioral assessment involves an exploration of the unique or idiosyncratic aspects of the single case, perhaps to a greater extent than any other approach."

Idiographic assessment focuses upon a target behavior or response class. A "response class" is defined as a group of behaviors that serve the

same function within a specific context (Jackson, 1999; Johnston & Pennybacker, 1993). "Experiential avoidance," defined as the avoidance of situations and conditions eliciting certain internal experiences (emotions, cognitions) that an individual finds intolerable, is an example of a response class (see Hayes, Wilson, Gifford, Follette, & Strosahl, 1996). The assessment of behavior is guided by two key concepts. The first of these is "situational specificity," or variance in a child's behavior as situational factors surrounding the child change. This means that assessment focuses upon those variables that elicit (antecedents) and maintain (consequences) a target behavior in a particular situation (Olweus, 1979). Situational specificity necessitates that assessment samples behavior across diverse settings and time points. Hence assessment of the child's behavior at home, in school, and on the playground is important, in addition to information obtained in the clinic setting. Furthermore, the information obtained from these various settings probably will not be, and in fact should not be, expected to be the same. For instance, the child may behave aggressively in school and on the playground, but not at home or in the clinic. The second concept is "temporal instability," or variance in a child's behavior over time. Such instability in behavior dictates that a child's behavior needs to be assessed at several points in time.

Measures designed according to idiographic principles generally adhere to "generalizability theory" (Cronbach, Gleser, Nanda, & Rajaratnam, 1972). Generalizability theory is a statistical framework for investigating and designing idiographic tools that evaluates the performance of a measure across facets (i.e., sources of variation) relevant to different applications of the measure. Five facets—forms, items, observers, time, and situation—are typically considered (Barrios & Hartmann, 1986; Cronbach et al., 1972). In contrast to classical test theory, which views variability across these facets as error, generalizability theory views this variability as something to understand and something central to change. Variability in scores across the different facets is examined. If a facet is associated with significant variability, then this suggests that scores would not generalize from one condition of the facet to another. For example, if observer bias accounts for a significant proportion of the variance in direct observations of aggressive behavior, then the scores would not be considered reliable across these different observers. Generalizability theory is therefore consistent with the tenets of behavioral theory.

The underlying assumptions and psychometric strengths of idiographic tools make them uniquely suited for specific aspects of treatment. In essence, these assessment tools pick up where nomothetic tools leave off. After a diagnosis is assigned, idiographic tools are uniquely suited to assessing behavioral, cognitive, affective, and contextual variables that may serve to maintain the target behavior. These tools therefore play a critical role in generating the case conceptualization and determining the impact of treatment.

Combining Diagnostic and Behavioral Assessment

The importance of using both diagnostic and behavioral assessment tools is increasingly being recognized (Mash & Barkley, 2007). However, it is important that clinicians understand when to use diagnostic and behavioral tools during treatment. At one time, it was relatively easy to differentiate behavioral from diagnostic assessment on the basis of the methods employed. Direct observation was originally the sole assessment technique of behavioral assessment, whereas interviewing characterized diagnostic assessment. However, as both assessment traditions evolved to include a wider repertoire of methods, differentiating behavioral and diagnostic approaches simply on the use of specific methods became more difficult. Indeed, there is now considerable overlap in ongoing assessment practices between the two approaches.

Presently, the difference between the two approaches lies less in the methods employed than in the manner in which findings generated with these assessment tools can (and should) be interpreted and used. Measures designed according to nomothetic principles are designed to compare an individual child to the larger population on a domain of interest. Measures designed according to idiographic principles are designed to identify target behaviors (overt or covert), their controlling conditions, and the functions they serve for a particular individual. Because the line between diagnostic and behavioral assessment techniques has become blurred, clinicians must understand how to critically evaluate assessment tools to determine their most appropriate application. This means that they need to understand how to determine whether a measure was developed according to nomothetic or idiographic principles, and to know when assessment tools from each tradition should be used at different stages in treatment.

ASSESSMENT PRINCIPLES

As noted at the beginning of this chapter, child assessment requires knowledge and skills from several domains: developmental principles, child psychopathology, psychometric theory, diversity/cultural issues, and therapy process and outcome research. The knowledge and skills gained from these different areas are required to guide the selection and interpretation of assessment data throughout treatment. Though the field has accumulated data on a wide variety of measures, we currently lack research on how to combine and interpret findings from these various measures. In the absence of empirical findings to direct the assessment process, we recommend that clinicians adhere to a set of principles to guide this process. The following overarching principles that guide our approach to child and adolescent assessment are summarized in Table 1.3. We now turn to a more detailed description of each principle and how each applies to assessment.

TABLE 1.3. Six Assessment Principles

Number	Principle
Principle 1	Empirical evidence and developmental psychopathology theory are used in selecting the constructs to target in assessment, as well as the best methods and tools to use.
Principle 2	Assessment is an ongoing process that uses a hypothesis-testing approach to inform decision making. As such, emphasis is placed on assessment tools that inform screening, diagnosis, case conceptualization, treatment selection/planning, and evaluation of treatment progress and outcome.
Principle 3	Thorough child assessment requires a multimethod, multi-informant approach that utilizes both nomothetic and idiographic assessment tools, and that focuses on a child's behavior, cognitions, affect, and social context.
Principle 4	Selecting constructs for assessment, determining a method for gathering assessment data, and interpreting findings should be informed by knowledge of developmental norms associated with specific child and adolescent emotional and behavioral problems.
Principle 5	Selecting constructs for assessment, determining a method for gathering assessment data, and interpreting findings should be informed by knowledge of ways in which culture and diversity can influence the experience and expression of child and adolescent emotional and behavioral problems.
Principle 6	The choice of assessment tools should be based on the strength of the tools' psychometric support for the type of client being assessed and the goals of the assessment. Careful attention should also be given to the judgmental heuristics that guide the interpretation of findings.

Principle 1. Empirical evidence and developmental psychopathology theory are used in selecting the constructs to target in assessment, as well as the best methods and tools to use.

Child assessment is complex. It is challenging to choose the right methods and informants for the wide range of possible targets. To reduce the potential for bias, the selection of methods and informants needs to be guided by the most recent empirical evidence and theory. The developmental psychopathology perspective (e.g., Cicchetti & Cohen, 1995; Masten & Braswell, 1991) provides an organizational framework for understanding childhood psychopathology and identifying the mechanisms and processes implicated in the development, maintenance, and alleviation of these problems (McLeod et al., in press; Youngstrom, 2008). As the mechanisms and processes represent potential treatment targets, this literature is ideal for identifying constructs to target in assessment and treatment.

To appreciate how this research can inform child assessment, it is important to understand the foundational principles of developmental

psychopathology. According to this perspective, children grow and change within the context of larger systems that can exert an influence upon child development and the expression of psychopathology. Risk and protective factors that are external (e.g., familial, social/environmental) and internal (e.g., biological, cognitive) interact to determine whether a child successfully masters each developmental stage. Mastery of the skills associated with each stage tends to leave a child better equipped to handle subsequent challenges. Failure to master the skills associated with a developmental stage leaves a child unprepared to deal with the demands of successive stages. Protective factors promote adaptation and help a child successfully negotiate a particular developmental stage. Risk factors, in contrast, decrease the likelihood that a child will achieve developmental milestones (Cicchetti & Cohen, 1995). The longer a child goes without mastering the skills of a developmental stage, the harder it is for her or him to return to normality. Ultimately, the interplay of risk and protective factors determines child outcomes; the development of a problem is more likely when the number of risk factors outweigh the protective factors.

Though certain risk factors are implicated in the development of specific disorders, developmental psychopathology considers each child unique. Two concepts explain this perspective. The first is the concept of "multifinality," or the idea that a single risk factor may lead to a variety of outcomes, depending upon the context in which it occurs. Basically, this principle suggests that a process (i.e., impact upon the individual) of any one factor (e.g., genetics, environment) varies, depending upon the context (e.g., family system) in which the factor operates (Cicchetti & Cohen, 1995). For example, some research suggests that parenting style can buffer children with behaviorally inhibited temperaments against the development of anxiety (Fox, Henderson, Marshall, Nichols, & Ghera, 2005). The second concept is "equifinality," or the idea that any given outcome (e.g., an anxiety disorder, conduct disorder) can have multiple causes. This means that a single causal pathway does not universally account for the development of specific emotional or behavioral problems. For example, an inhibited temperament (i.e., genetic pathway) and traumatic events (i.e., classical conditioning) have both been linked to social phobia (Stemberger, Turner, Beidel, & Calhoun, 1995). Together, these concepts have important implications for child assessment. Specifically, it is important to assume that although research can tell us about general risk factors for specific disorders, each child will have a unique combination of risk factors implicated in the development and maintenance of a disorder; assessment therefore needs to be highly individualized and tailored to each child.

Principle 2. Assessment is an ongoing process that uses a hypothesis-testing approach to inform decision making. As such, emphasis is placed on assessment tools that inform screening, diagnosis, case conceptualization, treatment selection/planning, and evaluation of treatment progress and outcome.

Using the developmental psychopathology literature to guide the assessment process necessitates adopting a hypothesis-testing approach to assessment. Research has identified specific risk factors that are associated with the development and maintenance of each disorder. However, each child will have a unique combination of factors implicated in the development and maintenance of a disorder. Thus the literature can be used to generate hypotheses about the mechanisms and processes at play for a particular child, which then are tested through both assessment and treatment. In other words, assessment and treatment should focus on the research-supported risk factors associated with a particular disorder and should systematically work toward identifying the specific factor(s) that play a role for a particular child. By doing so, the clinician will be able to develop and test hypotheses about which factors need to be targeted as part of treatment.

Assessment designed to inform treatment is an ongoing process that constantly evolves over the course of treatment. To guide the clinical decision-making process, and to ensure that this entire process is objective and grounded in the research literature, we advocate that the clinician's hypothesis-testing approach be rooted in the empirical principles of the behavioral assessment tradition (Ollendick & Hersen, 1984). Again, using the research literature to select targets for assessment, the clinician can generate hypotheses about the functional relation among causal, maintaining, and target behaviors. These hypotheses form the foundation of the case conceptualization, which is designed to guide treatment selection and planning as well as evaluation of treatment progress and outcome. Specific behavioral assessment tools and strategies, such as functional analysis and single-case series design, can be used to test the hypotheses during treatment. Assessment during treatment is then used to test the hypotheses, and the resulting data are used to make adjustments to the case conceptualization and treatment plan.

Principle 3. Thorough child assessment requires a multimethod, multi-informant approach that utilizes both nomothetic assessment tools, and that focuses on a child's behavior, cognitions, affect, and social context.

A multimethod, multi-informant approach is an important part of child assessment. Child assessment aims to describe multiple target behaviors accurately, including overt behavior, affective states, cognitive processes, and information about the child's context (Barry, Frick, & Kamphaus, 2013). Multiple tools are needed to assess these different facets of child behavior, because most assessment tools are designed to characterize only one aspect of child behavior. For example, some measures catalog child symptoms (e.g., parent rating scales), but do not provide information about how contextual factors may influence the symptoms. Part of a multimethod approach is blending diagnostic (nomothetic) and behavioral (idiographic) methods, as these approaches play important, complementary roles in child assessment. Psychometric issues also necessitate the use of multiple

assessment tools. Every assessment tool has psychometric strengths and weaknesses. Clinicians must therefore pick measures with complementary areas of strength, so that appropriate tools are used to assess each aspect of child behavior (Barry et al., 2013). For these reasons, it is important to employ a multimethod approach to child assessment.

A multi-informant approach is also an important component of child assessment. In choosing informants, clinicians must consider a number of factors. In the following paragraphs, we discuss when and why different individuals might be asked to serve as informants in child assessment.

Children

When children present for treatment, it is important to get their perspective on the target behavior. Gathering information from children provides an opportunity to build an alliance as well as to arrive at treatment goals. Several factors must, however, be considered in collecting and interpreting data from child informants. Younger children may not provide accurate information or be able to report about certain symptoms (Kamphaus & Frick, 2005; Schroeder & Gordon, 2002). As children enter adolescence, they are able to report upon their behavior and may be better informants for certain types of problems (e.g., substance abuse, anxiety) and/or for problems that adult informants do not observe (e.g., covert behaviors, such as stealing). When clinicians are considering the accuracy of child report, however, social desirability must also be considered (see De Los Reyes & Kazdin, 2005). Some children may want to please a clinician and thus provide answers they believe the clinician wants to hear. Other children may wish to conceal certain behaviors and thus may not provide accurate answers about those. The attributions children make about the causes of their problems may also influence the information they provide (De Los Reyes & Kazdin, 2005). In particular, children may be more likely to attribute the cause of their problems to environmental factors (e.g., family relations), which may make them less likely to endorse specific symptoms of certain disorders (e.g., less likely to see their own behavior as oppositional or defiant).

Parents

Typically, parents are the primary informants in child assessment (Paikoff & Brooks-Gunn, 1991), especially if a child is young. Interviewing parents provides clinicians with an important opportunity to build an alliance and engage the parents in the clinical process (see Hawes & Dadds, Chapter 12, this volume). However, a number of factors can influence the accuracy of parent report. First, parental psychopathology (e.g., depression) can negatively influence the accuracy of parent report (Chi & Hinshaw, 2002; Richters, 1992). Second, parents may not have an accurate understanding of normative child behavior, especially compared to that of other adults (e.g., teachers; Barry et al., 2013). If so, parents may see particular child

behaviors as problematic when in fact the behaviors are part of a normative developmental process. Third, parents' attributions regarding child behavior must be considered (De Los Reyes & Kazdin, 2005). Finally, parents of adolescents may not have full access to all aspects of their children's lives and thus may not be accurate reporters on certain behavior (e.g., affect, stealing, drug use).

Teachers

For certain child problems, teachers can be important informants. Certain symptoms may be observed first at school or may be particularly problematic in this setting. In such cases, a teacher can offer an important viewpoint, especially if a parent does not have the same opportunity to view the behavior. Furthermore, because teachers may have more experience with children than parents have, they may have a more developmentally sensitive view of child behavior (Barry et al., 2013). This can be helpful in deciding whether a parent's report is accurate or not. As with other informants, however, several factors may influence the accuracy of teacher data. First, teachers often do not have access to all facets of child behavior, so they are generally more accurate when reporting upon behavior they have actually observed (Loeber, Green, Lahey, & Stouthamer-Loeber, 1991). Second, a clinician must consider the amount of contact a teacher has with a student when asking the teacher to provide information on a child. Teachers typically have less contact with children as they get older, so this must be considered for children of middle school age or above (Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985). Third, the attributions made by a teacher about a child's behavior must be considered in interpreting teacher report data.

Peers

Social disruptions and impairment are common problems in youth. Peer report provides a unique perspective on a child and his or her social functioning/status. However, asking peers to provide ratings raises some ethical issues (e.g., asking peers to provide ratings might violate a client's confidentiality). Thus a clinician must be careful not to be too intrusive when asking peers to report data.

Principle 4. Selecting constructs for assessment, determining a method for gathering assessment data, and interpreting findings should be informed by knowledge of developmental norms associated with specific child and adolescent emotional and behavioral problems.

The choice of assessment methods and the process of interpreting the findings should be influenced by knowledge of the developmental norms associated with specific disorders and symptoms (Holmbeck et al., 2008). Age-related constraints are numerous in child assessment and should be

considered in selecting specific methods of assessment. For example, interviews may be more difficult to conduct and self-report measures less reliable with younger children, whereas self-monitoring and direct observations may be more reactive at older ages (Ollendick & Hersen, 1984). The selection of assessment instruments should therefore be guided by knowledge of cognitive and socioemotional developmental processes.

The interpretation of assessment data must also be informed by normative guidelines. As part of the assessment process, clinicians must ascertain whether a child is exhibiting developmentally adaptive or maladaptive behavior. Symptoms and behaviors that are considered normative at one developmental stage may not be considered so at a later stage. For example, it is typical for fear of separation to develop in infancy, for fear to move on to social situations in childhood, and for fear to become more generalized in adolescence (Gullone, 1996). So an intense fear of separation from caregivers is not unusual in young children, but is not considered developmentally appropriate in school-age children (Gullone, 1996; Warren & Sroufe, 2004). As another example, tantrums are considered normative at a young age, but become less so for school-age children. It is therefore important to determine whether the expression of a particular symptom is congruent with a child's developmental level or likely to represent a symptom that is interfering with functioning (Silverman & Ollendick, 2005; Warren & Sroufe, 2004).

Another factor that influences the interpretation of assessment data is knowledge of how age differences can influence the expression of behaviors and syndromes. Young children cannot manifest certain symptoms, such as guilt, hopelessness, or worry, before they achieve certain developmental milestones. For example, worry requires insight, which may not fully develop until late childhood (e.g., Dadds, James, Barrett, & Verhulst, 2004). Understanding the relation between cognitive development and the experience of certain symptoms can help a clinician avoid misattributing reports of a child's behavior to symptoms that are not consistent with the child's developmental level.

In sum, it is important for clinicians to take developmental factors into consideration when using assessment data to drive the therapy process. Developmental factors determine what assessment tools (e.g., interview, questionnaire, and direct observation) will provide accurate and valid information. Clinicians must therefore select developmentally appropriate methods of assessment and interpret the resulting data from the perspective of developmental norms.

Principle 5. Selecting constructs for assessment, determining a method for gathering assessment data, and interpreting findings should be informed by knowledge of ways in which culture and diversity can influence the experience and expression of child and adolescent emotional and behavioral problems.

"Culture" is defined as "an integrated pattern of human behavior that includes thought, language, action, and artifacts and depends on man's capacity for learning and transmitting knowledge to succeeding generations" (Frisby & Reynolds, 2005, p. 5). Culture can influence the experience and expression of distress, so an individual's nationality, ethnicity, acculturation level, socioeconomic status, and gender must all be considered during assessment. The failure to take culture into consideration when conducting an assessment can lead to interpretative errors (Edwards, 1982; Ridley & Kelly, 2007). First, a behavior may be labeled as pathological when it is in fact normative within a given culture. Second, the opposite may also occur: A child's behavior may be assumed to be explained by cultural factors when the behavior is pathological. In either case, serious errors can occur.

Therefore, it is important for clinicians to consider cultural and diversity factors when selecting assessment tools and interpreting assessment data (see, e.g., Friedberg & McClure, 2002; Ridley & Kelly, 2007). First, the available evidence suggests that the expression of psychological symptoms and/or distress may vary across cultures (Weisz, Sigman, Weiss, & Mosk, 1993). This variation may be due to value systems that find different symptoms more or less acceptable. For example, cultures that place a high value on deference to authority appear to have lower rates of externalizing problems (Weisz, Suwanlert, Chaiyasit, & Walter, 1987). As another example, some have hypothesized that the acceptability of medical symptoms (as opposed to psychological symptoms) in the Hispanic/Latino cultures explains why Hispanic/Latino children report more somatic symptoms than European American children (see, e.g., Pina & Silverman, 2004).

Second, cultural factors can influence reporting practices. Specifically, the acceptability of certain symptoms may influence what symptoms are reported as problematic. For example, families from inner-city communities may see aggressive behavior as adaptive and thus may not report aggressive behaviors to a clinician (Atkins, McKay, Talbot, & Arvanitis, 1996). This means that a clinician cannot assume that a particular symptom is present or absent just because a family does not report it as a problem. The accurate interpretation of assessment data depends, in part, upon gaining an understanding of particular families' values.

In sum, understanding how culture influences symptom expression and reporting practices is an important component of conducting a culturally sensitive assessment. In selecting assessment tools, it is important to determine whether the tools have demonstrated validity across different cultural groups (see Pina, Gonzales, Holly, Zerr, & Wynne, Chapter 13, this volume). It is also important to work with the child and family to understand whether cultural factors influence the interpretation of assessment data (e.g., by using cultural mapping techniques; Pina, Villalta, & Zerr, 2009). It is vital for clinicians to be able to distinguish between normal variations associated with culture and abnormal variations characteristic of psychopathology.

Principle 6. The choice of assessment tools should be based on the strength of the tools' psychometric support for the type of client being assessed and the goals of the assessment. Careful attention should also be given to the judgmental heuristics that guide the interpretation of findings.

Assessment procedures should not only be culturally sensitive and developmentally appropriate, but also psychometrically validated (Ollendick & Hersen, 1984). To date, the practice of child assessment has been marked by the use of assessment tools that are convenient, with far too little attention paid to the measures' psychometric properties (Hunsley & Mash, 2008; Youngstrom, 2008). However, recent commentators on assessment have argued for a technology of evidence-based assessment that includes efforts to identify important psychometric dimensions and a system for rating the quality of each metric (Hunsley & Mash, 2008). It is imperative that the selection of assessment instruments be informed by the psychometric properties of specific tools, the evidence supporting how best to interpret particular tools, and the methods used to integrate multiple sources of data. When a clinician is reviewing the data in support of a particular instrument, it is important to note that a given tool is only supported for particular types of clients and probably only for certain purposes (Hunsley & Mash, 2008). For example, a scale may be very reliable when completed by an adolescent, but much less so when completed by a younger child. Similarly, a scale may be very useful for screening, but may not be sensitive to change and therefore not useful for outcome monitoring. In other words, the selection of an assessment tool needs to be influenced by whether the psychometric properties of the tool are supported for the type of client being assessed and the goals of the assessment.

It is important to note that the psychometric concepts relevant to nomothetic tools do not directly apply to idiographic instruments (Foster & Cone, 1995). In fact, there have been debates about what psychometric categories are relevant to idiographic tools (Foster & Cone, 1995; Jackson, 1999). Perhaps for this reason, recent efforts to identify important psychometric dimensions have focused upon nomothetic tools. Hunsley and Mash (2008) recently introduced a framework for considering the psychometric properties of nomothetic tools. This framework focuses upon the following categories: standardization, norms, reliability, validity, and clinical utility. Below, we cover these domains; however, we also review domains relevant to idiographic instruments.

Standardization

"Standardization" refers to the extent to which an assessment technique is delivered in a consistent manner across various conditions of administration (Barrios & Hartmann, 1986). The goal of standardization is to

improve reliability by minimizing the influence of potential sources of error—child, clinician, context—on the scores produced by a particular instrument. When clinicians do not follow a predetermined set of questions or procedures, assessment techniques are susceptible to bias. The administration of diagnostic (e.g., diagnostic interviews) and behavioral (e.g., self-monitoring activities) tools can be standardized. Unstructured clinical interviews have been criticized for producing variable results (Angold & Fisher, 1999; Garb, 1998, 2005). Similarly, idiographic tools have been criticized because they lack standardized administration (Jackson, 1999). For this reason, it is important to ensure that when idiographic instruments are used, the instructions and items used with children and their families are held constant, so that changes in scores can be clearly interpreted.

Norms

For nomothetic tools, the availability of norms provides a concrete assessment of a child's behavior relative to other children. However, the quality of a measure's norms is important to consider. Ideally, a normative sample (1) should be representative of the population under study, (2) should be large enough to provide stable estimates of the population mean and standard deviation, and (3) should include clinical and nonclinical samples (Anastasi, 1988; Hunsley & Mash, 2008). To evaluate whether the norms for a measure fit a specific client, it is necessary to compare the composition of the normative sample to specific client characteristics. If a client is very different from a measure's normative sample on characteristics that might affect the meaning of the client's scale score, it is better to find a measure with more representative norms. Idiographic tools do not typically rely upon population norms, so this psychometric dimension is not relevant when the psychometric strength of idiographic tools is being considered.

Reliability

For nomothetic instruments, "reliability" refers to the consistency and dependability of a person's score on a measure. For example, a self-report measure is considered reliable if it provides the same score across repeated assessments. When clinicians are evaluating the quality of a nomothetic tool, it is important to consider internal consistency and test–retest reliability. "Internal consistency" assesses whether all questions in a measure contribute consistently to the overall measure score. Low internal consistency indicates that the questions may not all assess the same construct (e.g., depression). According to Hunsley and Mash (2008), the accumulated evidence for a measure should suggest that the internal validity for a measure is at least .70. "Test–retest" reliability assesses the stability of scores over multiple time points. This form of reliability is used when an instrument is designed to assess a construct that is purported to be stable over time (e.g.,

temperament). According to Hunsley and Mash (2008), test–retest coefficients are considered acceptable if equal to or greater than .70 over a short period of time (days or weeks) and excellent if over .70 for a long period of time (1 year or longer).

The concepts of reliability and accuracy also overlap with idiographic instruments (Cone, 1998; Jackson, 1999). "Accuracy" refers to the extent to which recorded data (self-report, self-monitoring, observational) provide a good representation of a target behavior (Cone, 1998; Jackson, 1999). To determine accuracy, an incontrovertible index is required that represents a "gold-standard" measure of the target behavior. As an incontrovertible index rarely exists for most behaviors, it is difficult to determine the true accuracy of specific observations. Reliability can also be assessed within the framework of generalizability theory. Facets that result in significant variability indicate that scores on a measure are not reliable for those facets. Decision studies can be used to determine how many observations are needed from a particular facet in order to produce a reliable estimate (Brennan, 2001).

The concept of "interrater reliability" applies to both idiographic and nomothetic tools; it refers to the differences in obtained results among raters who are using the same instrument. This type of reliability estimate is useful when clinicians are using interviews (e.g., standardized interviews such as the Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions; Silverman & Albano, 1996) or direct observation (e.g., the Autism Diagnostic Observation Schedule; Lord et al., 2000). According to Hunsley and Mash (2008), for categorical data, acceptable interrater reliability (kappa) falls between .70 and .79 and is preferred to be above .85; an acceptable Pearson/intraclass correlation ranges from .70 to .79 and is preferred to be above .90.

Validity

"Validity" refers to whether or not an instrument assesses what it purports to measure. Foster and Cone (1995) draw a distinction between *representative* and *elaborative* validity. "Representative validity" focuses upon establishing whether a tool assesses the theoretical domain or response class it is designed to assess (e.g., depression, experiential avoidance). "Elaborative validity" refers to whether a tool has utility for measuring a construct or response class. Several different validity dimensions exist. Whether a validity dimension is relevant to a particular nomothetic or idiographic tool depends in part upon what the instrument is designed to do.

"Content validity" is an important component of representative validity for nomothetic and idiographic instruments; it means that items capture all aspects of a given domain (e.g., depression). To establish content validity, researchers must clearly define the target domain and then demonstrate that the items represent all facets of that domain. For nomothetic tools,

content validity is established by demonstrating that the items on an instrument tap into all aspects of the purported construct. Ideally, test developers of nomothetic instruments should clearly define the domain of interest and have expert judges rate the fit of each item on a quantitative scale (McLeod et al., in press). For idiographic tools that focus on a response class, content validity involves ensuring that all facets of the response class are defined. Content validity for idiographic tools also includes the sampling plan. To ensure content validity, a behavior must be sampled enough times to ensure the data will generalize across time and situations (Jackson, 1999).

Accuracy is a key validity dimension for idiographic instruments. As noted above, "accuracy" is defined as the extent to which scores represent the cognitive, affective, and behavioral components of the behavior (Foster & Cone, 1995). Establishing accuracy requires the existence of an independent, incontrovertible index of the target behavior. Three different ways of assessing accuracy have been proposed. First, scores produced by an idiographic instrument can be compared to physical evidence of a target behavior, such as mechanical recordings (Foster & Cone, 1995; Johnston & Pennypacker, 1993). Second, scores produced by an idiographic instrument can be compared to direct observation of a target behavior within the natural environment (Foster & Cone, 1995; Suen & Ary, 1989). And finally, controlled stimuli to which an idiographic instrument should be sensitive can be introduced to determine whether the tool captures the manipulation (Foster & Cone, 1999). For idiographic instruments that assess covert events and/or rely upon indirect methods (self-report), it is challenging to demonstrate accuracy. For this reason, these instruments are often evaluated in terms of convergent and discriminant validity (Jackson, 1999).

"Construct validity" provides evidence that an instrument taps into the theoretical concept that it was designed to assess (Foster & Cone, 1995; Hill & Lambert, 2004). There are multiple forms of construct validity (i.e., "convergent," "discriminant," "predictive," and "concurrent"). Two important categories of construct validity are convergent and discriminant evidence, which help to establish the representative validity of an instrument. This evidence is concerned with whether an instrument converges with measures of similar constructs and diverges from measures of different constructs. Traditionally, these categories were used for nomothetic instruments. However, as idiographic tools began to assess covert behaviors and employ indirect methods, these validity dimensions were applied to idiographic instruments.

Another important category of construct validity is "test–criterion relationships" (related to the traditional concepts of "concurrent" and "predictive" validities), which helps evaluate the elaborative validity of an instrument. This evidence applies to both nomothetic and idiographic tools, and indicates whether a measure is related to some present or future outcome that is thought to be meaningfully related to the construct the measure is supposed to be assessing. In the case of tools used to inform treatment,

an important test–criterion relationship might be whether people who are assigned diagnoses by the instrument also show high levels of functional impairment that would be anticipated to result from the disorder. For an instrument to demonstrate construct validity, the majority of the data collected on the measure should support the different facets of construct validity. A dimension related to test–criterion relationships is "validity generalization." This dimension relates to elaborative validity and assesses the extent to which a measure's test–criterion relationships generalize across settings and populations different from the ones in which the instrument was originally validated. Validity generalization addresses the important question of whether the tool can be used across multiple contexts (home, school, clinic) and/or populations (age, gender, ethnicity).

The final validity dimension is "treatment sensitivity," which is also related to elaborative validity. This dimension is only relevant to instruments used for treatment monitoring and evaluation. Evidence for treatment sensitivity is demonstrated when a measure evidences some sensitivity to change over the course of treatment. This is a relatively new validity dimension, so clear guidelines for assessing the strength of evidence do not exist. Important issues to consider in evaluating evidence for treatment sensitivity include whether a measure is responsive to change across different types of treatment and how often the measure can be administered (whether the measure can be administered weekly, monthly, or at longer intervals).

In sum, validity is extremely important, because it indicates whether a measure is assessing the construct or behavior of interest. Multiple studies are needed to demonstrate the different forms of validity evidence, so it can sometimes be challenging to review validity evidence for particular measures.

Clinical Utility

In order to meet evidence-based standards for assessments, measures should also provide some indicator of "clinical utility" (Hunsley & Mash, 2007; Nelson-Gray, 2003; Vasey & Lonigan, 2000). Clinical utility is a relatively new validity dimension and has not received much empirical attention. At present, it can include "diagnostic utility" (a measure's ability to lead to a correct diagnostic conclusion), "incremental utility" (what information a particular measure can provide that cannot be provided by other instruments), "treatment utility" (a measure's beneficial contribution to treatment outcome), and "feasibility" (the ease with which the measure can be integrated into clinical practice). Essentially, for an instrument to have clinical utility, empirical studies must demonstrate that the use of the tool improves the accuracy, outcome, and/or efficiency of clinical activities (Hunsley & Mash, 2007).

The practicality and cost of an instrument should also be considered

in the choice of an assessment tool. A number of factors can influence cost, including the amount of time required to administer the measure, the financial cost of administering and scoring it, the amount of time spent scoring and interpreting it, and required equipment (e.g., computers to score the measure; Jensen-Doss, 2005; Yates & Taub, 2003). Cost can also include training costs and/or the level of training required to administer and score the instrument (i.e., whether a trained clinician must administer the measure). In sum, balancing practical considerations with psychometric quality can be challenging, but it is an important aspect of deciding what measures to use to inform treatment.

COMBINING DATA ACROSS INFORMANTS

As noted earlier, a multimethod, multi-informant approach is recommended in child assessment. However, very few empirical data exist about how best to combine findings across measures and informants to produce a picture of an individual child. Questions persist about when and how certain measures should be used, whether tools need to be differentially weighted in the clinical decision-making process, and how to resolve discrepant reports. Obviously, it is important to use psychometrically strong instruments; however, combining the data also represents a critical step. Unfortunately, no research exists to guide this process. In this section, we provide some general issues to consider in combining findings across informants.

Once assessment data are collected, the clinician must integrate the data and produce a treatment plan. The low rates of agreement across informants commonly seen in child assessment can make this a challenging endeavor (De Los Reyes & Kazdin, 2005). Until recently, the field offered very little guidance on how to conceptualize or address such discrepancies. Fortunately, a new model has emerged that provides a framework for understanding this important aspect of assessment with children.

De Los Reyes and Kazdin (2005) have proposed the "attributions bias context" (ABC) model. The ABC model posits three factors that may influence informant discrepancies in child assessment. First, *informant's attributions* about the causes of problems may influence their reports. Children may be more likely to view their problems as contextual (e.g., "I am being bullied"), whereas others may view them as dispositional (e.g., "He is aggressive"). These differences in attributions are then related to differences in *informant's perspectives* regarding the nature of the problem and the need for treatment. For example, children may view their problems as lying within a specific situational context (e.g., "I get into fights because I am being bullied at school") and therefore not needing treatment, whereas others may perceive a need for treatment and may be more likely to report problems to support this view (e.g., "He is aggressive"). Finally, the *goal of*

the clinical process—which is often perceived as collecting negative information about the child—probably contributes to discrepancies, as children are less likely to want to provide this type of information than other informants. In addition to these perceptual differences, the model also posits that differences between informants arise from the circumstances under which they observe a child's behavior. A recent study supported this last point by demonstrating that children behaved differently in lab-based interactions with examiners than they did in interactions with their parents; their behavior with the examiners was strongly correlated with their teachers' reports of their behavior, whereas their behavior with their parents was correlated with the parents' reports (De Los Reyes, Henry, Tolan, & Wakschlag, 2009).

The ABC model provides a new lens through which to consider data gathered from multiple informants. Despite these recent advances, however, there is precious little empirical evidence to guide the combination of data across informants. As a result, it is possible for biased clinical judgment to have a negative influence on this process. This is another reason why clinicians need to take an empirical, hypothesis-testing approach: to minimize the impact of clinician bias on the assessment process. Building upon the recommendations of De Los Reyes and Kazdin (2005), along with others (Barry et al., 2013), we offer the following two recommendations for dealing with multiple informants.

First, while clinicians are gathering assessment data, it may be helpful to gather information related to the ABC model that might help explain informant discrepancies (De Los Reyes & Kazdin, 2005). For example, it is helpful to ask informants what their attributions for the causes of a child's behavior are, and whether they think the child's behavior warrants treatment. In addition, given that informants are likely to vary in their views about whether the child's behavior is contextual or dispositional, it is important to strike a balance between general questions about a child's behavior (e.g., "Is your child anxious about talking to people she does not know?") and context-specific questions (e.g., "Does your child experience anxiety about talking to people she does not know when she is at parties with other children?"; example from De Los Reyes & Kazdin, 2005). Finally, gathering data that might help clarify contextual influences on informant behaviors can help interviewers interpret assessment data. Behavioral assessment methods can be very useful in this regard; indeed, they are based upon the idea that behavior is context-specific. For example, direct observation of a child's behavior in the school and in the home can help a clinician understand whether teachers and parents are observing and reporting on the same types of behaviors.

Second, once the assessment data have been gathered, a clinician should determine which informants have reported clinically significant behaviors and whether there is convergence across informants. The clinician may have more confidence in reports that converge across informants,

but variation does not mean that any of the reports are incorrect. The clinician must therefore generate theory-driven hypotheses about what factors, such as context, culture, perceptual differences, or development, might account for differences across informants. For example, a given child may exhibit behavioral problems at home because those problems are reinforced by the parents, but does not exhibit those problems at school because the consequences for doing so are consistently negative. In this case, discrepancies between parent and teacher reports of behavior problems are not only to be expected, but are helpful for treatment planning, as they suggest that interventions targeting the home environment might be more useful than interventions targeting the school. Viewed through the lens of a theoretical model like the ABC model, informant discrepancies can be considered important sources of clinical data, rather than "noise" to be removed from the clinical picture. Multiple informants can therefore help in problem identification, case conceptualization, and treatment planning.

ETHICS AND STANDARDS OF CHILD DIAGNOSTIC AND BEHAVIORAL ASSESSMENT

A number of ethical issues arise in diagnostic and behavioral assessment with children and adolescents. Many of these issues cut across areas of practice and are not unique to assessment. For example, the American Psychological Association's (2002) "Ethical Principles of Psychologists and Code of Conduct" outlines guidelines for protecting client confidentiality, setting appropriate professional boundaries, maintaining records and billing, and other general areas of professional behavior. Standard 2 of this code of conduct also discusses at length issues related to competence. When applied to assessment, Standard 2 specifies that someone who conducts assessment should stay within the boundaries of his or her education, training, and experience. For example, it would not be considered ethical for a psychologist who has never been trained in the assessment of autism spectrum disorders to conduct an assessment to determine whether a child meets criteria for one of those disorders. As such, it is important to obtain training in the assessment strategies relevant to one's clinical practice, and to refer cases with assessment questions falling outside of one's training and expertise to other clinicians.

In addition to these general ethical principles, Standard 9 of the code details ethical issues specific to assessment. Here we highlight those most relevant to the assessment of youth psychopathology. First, the recommendations stemming from an assessment should be based on sufficient data to support those recommendations. The assessment principles described in this chapter, including the use of multimethod, multi-informant assessment, can help ensure that this is the case. Second, these data must be interpreted in a way that takes into account the test-taking and personal characteristics

of the client that might influence the interpretation. Grounding an assessment in the developmental psychopathology literature and being sensitive to multicultural issues can help ensure that assessment data are interpreted in an appropriate manner.

The code also specifies that the tools used in an assessment should be used in a manner supported by research and should have established reliability and validity for the members of the population tested. As discussed by Pina et al. in Chapter 13, this often presents a challenge for clinicians working with ethnic minority youth, because few instruments have been well tested with these populations. In those cases, the ethical code indicates that clinicians must be clear about the strengths and limitations of their testing approach.

Standard 9 also details principles to guide the process of informed consent for assessment, indicating that before clients consent to an assessment, they must receive information about the nature and purpose of the assessment, the cost of the assessment, and limits to confidentiality. For child assessments, unique issues arise in relation to informed consent. Legally, only parental consent is often required for child services, including assessment. However, often it is ethical to obtain assent from child clients as well. The type of information that is developmentally appropriate to provide to children, and the correct timing of informed consent (e.g., is it ethical to conduct a behavioral observation before obtaining assent?), are among issues that are not clearly addressed in the ethical code.

Finally, Standard 9 states that the results of an assessment must be clearly explained to the individual or a "delegated representative." In the case of child assessment, this means that the results must be clearly explained to parents. Together with the parents, the clinician should also decide how much information is appropriate to provide to the child, taking into account the nature of the assessment feedback and the child's developmental level. A thorough discussion of all ethical issues related to assessment is beyond the scope of this volume. However, a detailed understanding of the entire ethical code is essential for both psychology trainees and licensed psychologists.

SUMMARY

The past two decades have witnessed exciting advances in assessment. The rise of the evidence-based assessment movement has focused attention on assessment training and practice. At the heart of the evidence-based assessment movement is the principle that science and theory should guide and inform the assessment process. Increasingly, the field is moving toward establishing a set of guidelines for assessment practice. Despite these advances, critical gaps still exist. Most notably, very little research is available to guide assessment practices during treatment. It is our hope that the

evidence-based assessment movement will inspire more research that will help fill these knowledge gaps in the coming years.

Our approach to assessment is informed by this movement. The principles presented in this chapter are consistent with evidence-based assessment and are designed to help guide the assessment process from intake to termination. Both diagnostic and behavioral assessment tools are needed to inform the treatment process, and we have covered areas of knowledge (such as psychometric theory) that are needed to apply assessment tools over the course of treatment. With an overview of our assessment approach thus presented, we now turn to more detailed coverage of the knowledge and skills needed for diagnostic and behavioral assessment.

REFERENCES

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060–1073.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Angold, A., & Fisher, P. W. (1999). Interviewer-based interviews. In D. Shaffer, C. P. Lucas, & J. E. Richters (Eds.), *Diagnostic assessment in child and adolescent psychopathology* (pp. 34–64). New York: Guilford Press.

Atkins, M. S., McKay, M. M., Talbot, E., & Arvanitis, P. (1996). DSM-IV diagnosis of conduct disorder and oppositional defiant disorder: Implications and guidelines for school mental health teams. *School Psychology Review, 25*, 274–283.

Barrett, P. M., & Ollendick, T. H. (Eds.). (2004). *Handbook of interventions that work with children and adolescents*. Chichester, UK: Wiley.

Barrios, B., & Hartmann, D. P. (1986). The contributions of traditional assessment: Concepts, issues, and methodologies. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 81–110). New York: Guilford Press.

Barry, C. T., Frick, P. J., & Kamphaus, R. W. (2013). Psychological assessment in child mental health settings. In B. Bracken, J. Carlson, J. Hansen, N. Kucel, S. Reise, & M. Rodrequez (Eds.), *APA handbook of testing and assessment in psychology*. Washington, DC: American Psychological Association.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review, 81*, 506–520.

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

Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Evaluation, 31*, 141–154.

Chi, T. C., & Hinshaw, S. P. (2002). Mother–child relationships of children with ADHD: The role of maternal depressive symptoms and depression-related distortions. *Journal of Abnormal Child Psychology, 30*, 387–400.

Childs, R. A., & Eyde, L. D. (2002). Assessment training in clinical psychology doctoral programs: What should we teach? What do we teach? *Journal of Personality Assessment, 78*, 130–144.

Cicchetti, D., & Cohen, D. J. (1995). *Developmental psychopathology: Vol. 1. Theory and methods*. New York: Wiley.

Cone, J. D. (1986). Idiographic, nomothetic, and related perspectives in behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 111–128). New York: Guilford Press.

Cone, J. D. (1998). Psychometric considerations: Concepts, contents, and methods. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed., pp. 22–46). Boston: Allyn & Bacon.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles*. New York: Wiley.

Dadds, M. R., James, R. C., Barrett, P. M., & Verhulst, F. C. (2004). Diagnostic issues. In T. H. Ollendick & J. S. March (Eds.), *Phobic and anxiety disorders in children and adolescents: A clinician's guide to effective psychosocial and pharmacological interventions* (pp. 3–33). New York: Oxford University Press.

De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology, 37*, 637–652.

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*, 483–509.

Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child Development, 56*, 265–275.

Edwards, A. W. (1982). The consequences of error in selecting treatment for blacks. *Social Casework, 63*, 429–433.

Eells, T. D. (2007). *Handbook of psychotherapy case formulation* (2nd ed.). New York: Guilford Press.

Foster, S. L., & Cone, J. D. (1995). Validity issues in clinical assessment. *Psychological Assessment, 7*, 248–260.

Fox, N. A., Henderson, H. A., Marshall, P. J., Nichols, K. E., & Ghera, M. M. (2005). Behavioral inhibition: Linking biology and behavior within a developmental framework. *Annual Review of Psychology, 56*, 235–262.

Friedberg, R., & McClure, J. (2002). Review of clinical practice of cognitive therapy with children and adolescents. *Journal of Developmental and Behavioral Pediatrics, 23*, 457–458.

Frisby, C. L., & Reynolds, C. R. (2005). *Comprehensive handbook of multicultural school psychology*. Hoboken, NJ: Wiley.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association.

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology, 1*, 67–89.

Gullone, E. (1996). Normal fear in people with a physical or intellectual disability. *Clinical Psychology Review, 16*, 689–706.

Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology, 64*, 1152–1168.

Hill, C. E., & Lambert, M. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's*

handbook of psychotherapy and behavior change (5th ed., pp. 84–136). New York: Wiley.

Holmbeck, G. N., Thill, A. W., Bachanas, P., Garber, J., Miller, K. B., Abad, M., Zuckerman, J. (2008). Evidence-based assessment in pediatric psychology: Measures of psychosocial adjustment and psychopathology. *Journal of Pediatric Psychology, 33*, 958–980.

Hunsley, J. (2002). Psychological testing and psychological assessment: A closer examination. *American Psychologist, 57*, 139–140.

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology, 3*, 29–51.

Hunsley, J., & Mash, E. J. (Eds.). (2008). *A guide to assessments that work*. New York: Oxford University Press.

Jackson, J. L. (1999). Psychometric considerations in self-monitoring assessment. *Psychological Assessment, 11*, 439–447.

Jensen-Doss, A. (2005). Evidence-based diagnosis: Incorporating diagnostic instruments into clinical practice. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 947–952.

Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Kamphaus, R. W., & Frick, P. J. (2005). *Clinical assessment of child and adolescent personality and behavior*. New York: Springer.

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*, 725–739.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome?: A meta-analysis. *Clinical Psychology: Science and Practice, 10*, 288–301.

Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1991). Differences and similarities between children, mothers, and teachers as informants on childhood psychopathology. *Journal of Abnormal Child Psychology, 19*, 75–95.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205–223.

Mash, E. J., & Barkley, R. A. (2007). *Assessment of childhood disorders* (4th ed.). New York: Guilford Press.

Masten, A. S., & Braswell, L. (1991). Developmental psychopathology: An integrative framework. In P. R. Martin (Ed.), *Pergamon general psychology series: Vol. 164. Handbook of behavior therapy and psychological science: An integrative approach* (pp. 35–56). New York: Pergamon Press.

McLeod, B. D., Jensen-Doss, A., Wheat, E., & Becker, E. M. (in press). Evidence-based assessment and case formulation for child anxiety. In C. Essau & T. H. Ollendick (Eds.), *Treatment of childhood and adolescent anxiety*.

McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias and child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology, 72*, 235–251.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
Nelson-Gray, R. O. (2003). Treatment utility of psychological assessment. *Psychological Assessment, 15*, 521–531.
Nezu, A. M., Nezu, C. M., Peacock, M. A., & Girdwood, C. P. (2004). Case formulation in cognitive-behavior therapy. In M. Hersen (Series Ed.) & S. N. Haynes & E. M. Heiby (Vol. Eds.), *Comprehensive handbook of psychological assessment: Vol. 3. Behavioral assessment* (pp. 402–426). Hoboken, NJ: Wiley.
Ollendick, T. H. (1999). Empirically supported assessment for clinical practice: Is it possible? Is it desirable? *The Clinical Psychologist, 52*, 1–2.
Ollendick, T. H., & Hersen, M. (Eds.). (1984). *Child behavioral assessment: Principles and procedures*. New York: Pergamon Press.
Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. *Psychological Bulletin, 86*, 852–875.
Paikoff, R. L., & Brooks-Gunn, J. (1991). Do parent–child relationships change during puberty? *Psychological Bulletin, 110*, 47–66.
Pina, A. A., & Silverman, W. K. (2004). Clinical phenomenology, somatic symptoms, and distress in Hispanic/Latino and European American youths with anxiety disorders. *Journal of Clinical Child and Adolescent Psychology, 33*, 227–236.
Pina, A. A., Villalta, I. K., & Zerr, A. A. (2009). Exposure-based cognitive behavioral treatment of anxiety in youth: A culturally-prescriptive framework. *Behavioral Psychology, 17*, 111–135.
Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112*, 485–499.
Ridley, C. R., & Kelly, S. M. (2007). Multicultural considerations in case formulation. In T. D. Eells (Ed.), *Handbook of psychotherapy case formulation* (2nd ed., pp. 33–64). New York: Guilford Press.
Schroeder, C. S., & Gordon, B. N. (2002). *Assessment and treatment of childhood problems* (2nd ed.). New York: Guilford Press.
Silverman, W. K., & Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions*. San Antonio, TX: Psychological Corporation.
Silverman, W. K., & Ollendick, T. H. (2005). Evidence-based assessment of anxiety and its disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*, 380–411.
Stein, B. D., Kogan, J. N., Hutchison, S. L., Magee, E. A., & Sorbero, M. J. (2010). Use of outcomes information in child mental health treatment: Results from a pilot study. *Psychiatric Services, 61*, 1211–1216.
Stemberger, R. T., Turner, S. M., Beidel, D. C., & Calhoun, K. S. (1995). Social phobia: An analysis of possible developmental factors. *Journal of Abnormal Psychology, 104*, 526–531.
Suen, H. K., & Ary, D. (1989). *Analyzing quantitative behavioral observation data*. Hillsdale, NJ: Erlbaum.
Vasey, M. W., & Lonigan, C. J. (2000). Considering the clinical utility of performance-based measures of childhood anxiety. *Journal of Clinical Child Psychology, 29*, 493–508.
Warren, S. L., & Sroufe, L. A. (2004). Developmental issues. In T. H. Ollendick & J. S. March (Eds.), *Phobic and anxiety disorders in children and adolescents:*

A clinician's guide to effective psychosocial and pharmacological interventions (pp. 92–115). New York: Oxford University Press.

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review and critique of the literature. *Annual Review of Psychology, 56*, 337–363.

Weisz, J. R., Sigman, M., Weiss, B., & Mosk, J. (1993). Behavioral and emotional problems among Embu children in Kenya: Comparisons with African-American, Caucasian, and Thai children. *Child Development, 64*, 98–109.

Weisz, J. R., Suwanlert, S., Chaiyasit, W., & Walter, B. R. (1987). Over- and undercontrolled referral problems among children and adolescents from Thailand and the United States: The *wat* and *wai* of cultural differences. *Journal of Clinical and Consulting Psychology, 55*, 719–726.

Yates, B. T., & Taub, J. (2003). Assessing the costs, benefits, cost-effectiveness, and cost-benefit of psychological assessment: We should, we can, and here's how. *Psychological Assessment, 15*, 478–495.

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

Guilford Publications
72 Spring Street
New York, NY 10012
212-431-9800
800-365-7006
www.guilford.com