Introduction to Special Issue: Evidence-Based Assessment in Pediatric Psychology

Lindsey L. Cohen,[1] PhD, Annette M. La Greca,[2] PhD, Ronald L. Blount,[3] PhD, Anne E. Kazak,[4] PhD, Grayson N. Holmbeck,[5] PhD, and Kathleen L. Lemanek,[6] PhD

[1]Georgia State University, [2]University of Miami, [3]University of Georgia, [4]University of Pennsylvania School of Medicine, The Children's Hospital of Philadelphia, [5]Loyola University Chicago, and [6]Ohio State University College of Medicine

In the past decade, there has been considerable attention devoted to evaluating the scientific merit of psychological treatments. In 1995, Division 12 (Clinical Psychology) of the American Psychological Association (APA) embarked on an evaluation of evidence-based (formerly referred to as empirically supported) treatments for psychopathology (Task Force on Promotion and Dissemination of Psychological Procedures, 1995). A similar venture soon followed from Division 53 (Society of Child Clinical and Adolescent Psychology) to identify psychosocial interventions for children and adolescents, which were supported by research findings (Lonigan, Elbert, & Johnson, 1998). Division 54 (Society of Pediatric Psychology) of the APA conducted an evaluation yielding a series of articles on evidence-based treatments germane to pediatric psychology (Spirito, 1999).

The efforts to identify, describe, and disseminate information about treatments that are supported by sound scientific evidence are important endeavors for advancing scientist-practitioner approaches to therapy and research in psychology. Scientist-practitioners are equally concerned about the research base supporting their assessment instruments. Thus, in 1999, Division 12 conducted one of the first systematic evaluations of evidence-based assessments (Ollendick, 1999). A particular emphasis was placed on identifying the treatment utility of the measures, which refers to the extent that they provide information that helps guide the design of interventions (Hayes, Nelson, & Jarrett, 1987; Nelson-Gray, 2003). Thus, there was more attention devoted to measures of malleable state-like variables than measures of stable trait-like characteristics. In addition, this task force limited the scope of the review to laboratory and performance-based measures (e.g., observations under standardized conditions) of child clinical psychology (e.g., autism, anxiety, depression, and attention deficit hyperactivity disorder). In a summary of the findings, Frick (2000) concluded that the measures were useful in providing additional information about the disorders under study, but the instruments did little in guiding the design of treatment interventions. An explanation is that many of the laboratory-based measures reviewed were developed and employed to examine specific hypotheses about disorders rather than for clinical use.

In 2005, Division 53 published a special issue focusing on evidence-based assessment of child and adolescent disorders. In the lead article, Mash and Hunsley (2005) discussed some of the difficulties of conducting a review of evidence-based assessments, including the large number of measurement tools, the varied purposes of assessment, the rapid developmental changes in children, and the need for multiple sources of information (e.g., teachers and parents). This series detailed assessment approaches with children with anxiety, depression, bipolar disorder, attention deficit hyperactivity disorder, conduct problems, learning disabilities, and autism spectrum disorders.

Whereas there is a overlap in the concerns and issues in the fields of child psychopathology and pediatric psychology, there are also some distinctions that should be noted. The populations evaluated and foci of concern of pediatric psychology are distinct. For example, adherence to medical regimens, psychosocial issues related to having a chronic illness, and evaluations of medical pain management are issues more commonly addressed by pediatric psychologists than clinical child psychologists. Given that pediatric psychologists are often working with medical professionals and in busy

All correspondence concerning this article should be addressed to Dr Lindsey L. Cohen, Department of Psychology, Georgia State University, Atlanta, Georgia 30302-5010. E-mail: llcohen@gsu.edu.

Journal of Pediatric Psychology 33(9) pp. 911–915, 2008
doi:10.1093/jpepsy/jsj115
Advance Access publication May 4, 2006

medical care settings, pragmatics (e.g., cost-effective, time-efficient, and ease of use; Roberts & McNeal, 1995) and whether results lead to clear treatment implications (Blount, Bunke, & Zaff, 2000) might be prioritized over theory development and richness of information.

The systematic evaluation of assessment instruments is essential to many areas of scientist-practitioner endeavors. When conducting therapy, professionals rely on the accuracy of measures for diagnostic and case conceptualization purposes, and decisions regarding the course and efficacy of treatment are often based on results from assessment tools. In scientific research, the assessment tools are essential in the accurate examination of the constructs of interest. Given that the quality of the assessment measures are crucial to whether a psychosocial treatment is deemed to be "empirically supported," it seems that the horse has been put before the cart, or perhaps the hitch has unwittingly come undone somewhere along the way in the movement to identify evidence-based treatments.

Evaluating pediatric psychology assessment measures presents a host of challenges (La Greca & Lemanek, 1996). For example, the instrument might be used across a variety of populations, and the validity and reliability might vary across these groups. In addition to the question of whether an assessment instrument is valid for a particular population, there is also the question of whether the instrument is valid for a particular purpose. For instance, the researcher might be using the instrument to broaden the understanding of a construct or of factors associated with a disorder, identify avenues for intervention, or assess treatment outcome. A given measure might prove valid for some of these purposes but not others. Because there are multiple purposes for which an assessment instrument could be used, the scope of articles reviewed is diverse and could encompass practically any type of empirical investigation. These various considerations might help explain why evidence-based assessment reviews have lagged behind evidence-based treatment reviews in the literature.

In 2002, Annette La Greca assembled a task force of Division 54 (Society of Pediatric Psychology) to identify, critique, and disseminate information regarding the assessment measures used in pediatric psychology. Valid and reliable assessment is essential to the efforts of health care professionals who work with children. Diagnostic and treatment decisions are made based largely on results from the assessment devices used with children in medical settings. Unfortunately, some of the measures used with pediatric populations have had questionable utility (La Greca, 1994). Specifically, many instruments commonly used in pediatric settings were initially developed for use with healthy children or those with psychopathology; thus, no relevant norms are available for use with pediatric populations. Furthermore, many measures included in pediatric investigations are the modifications of adult measures without adequate developmental considerations. Some instruments do not readily yield clear treatment implications (Quittner, 2000), and others are too long or complicated to use in many busy health care settings. Related, there appears to be a split between those measures used in research and those used in clinical practice, which is likely due to the failure of researchers to adequately develop and disseminate practical measures and also to clinicians for not selecting instruments based on research findings (Beutler, 2000; Frick, 2000). Thus, it is apparent that a systematic review of pediatric psychology measures is in warranted to examine the state of the assessment field and to help determine where attention should be focused.

The goal of this series of articles is to identify and systematically critique the assessment tools available to the child health care community. The reviews are designed to help guide child health professionals in identifying and selecting instruments for particular purposes, to focus researchers' efforts in developing and evaluating measurement tools, and to foster greater integration of science and practice.

Compilation of Measures

The collection of measures was conducted using a multistep process. First, the APA Division 54 (Society of Pediatric Psychology) assessment task force steering committee (SPP-ATF) identified the following eight broad areas of interest: quality of life, family functioning, psychosocial functioning and psychopathology, social support and peer relations, adherence, pain, stress and coping, and cognitive functioning. Second, the SPP-ATF selected experts in each of the eight topic areas to chair small work groups, who would conduct and write the reviews for the series. Third, the SPP-ATF, along with the work groups, identified measures falling into the eight topic areas. At this point in the process, quantity over quality in measures was the goal; the work group identified as many measures as possible that tapped the eight topic areas. Fourth, a survey was conducted that listed all 367 measures identified by the SPP-ATF and work groups. The number of measures per category ranged from 24 to 66 (M = 45.87, SD = 15.92). The survey respondents were instructed to indicate with

a check mark whether they used or considered using the measure in research or practice. In addition, respondents were asked to list any other measures that they use that were not included in the survey. The survey was distributed via the internet to the APA, Division 54 (SPP) listserv, which consists of approximately 325 subscribers. Eighty-seven people (27%) responded to the survey. Although this response rate is low, it should be noted that many subscribers might not have been able to open the large attachment. It is also likely that the time-consuming nature of the survey contributed to the low response rate. Of the respondents, 0–79 people endorsed any particular measure (M = 10.57, SD = 15.00). In addition, 56 measures not initially included in the survey were identified by respondents. Thus, 423 measures were compiled by the SPP Task Force. Despite the efforts to compile a comprehensive list of measures used in pediatric psychology, the SPP-ATF acknowledges the likelihood that many measures were unintentionally excluded from the final list.

Critique of Measures

In an effort to conduct a systematic critique of measures, the SPP-ATF developed specific criteria to be used in the analyses of measures. The criteria closely parallel those used in the Special Series on Empirically Supported Treatments in Pediatric Psychology (Spirito, 1999). Chambless and Ollendick (2001) provide an overview of these and similar criteria used to evaluate psychological interventions. The three-level hierarchy for this series consisted of the following categories: well-established assessment, approaching well-established assessment, and promising assessment (Table I). These categories were based on criteria, such as validity and reliability, described as essential in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). The SPP-ATF debated the criteria appreciating that any system that results in hierarchical categorizations might be controversial if not problematic. A consensus was reached via the discussion with the common goal of advancing pediatric care by identifying and disseminating recommendations for measurement.

Charge to the Reviewers

The SPP-ATF provided the work group chairs with directives to describe, critique, and classify the measures in their domain. The work group chairs were provided with the survey results, which detailed the number of respondents who endorsed each of the initial 367 measures and the additional 56 measures generated by respondents. The SPP-ATF encouraged work groups to focus their evaluations on the most widely used and studied measures rather than attempting to be exhaustive in their reviews. To select the measures for inclusion in the review, SPP-ATF gave the work groups were given latitude to use a combination of the results of the survey, the extant literature, and their expertise. The work groups were provided with the criteria to use for critiquing and classifying the measures, which contained mutually exclusive indices (Table I). Furthermore, detailed

Table I. Criteria for Evidence-Based Assessment

Category	Criteria
Well-established assessment	The measure must have been presented in at least two peer-reviewed articles by different investigators or investigatory teams
	Sufficient detail about the measure to allow critical evaluation and replication (e.g., measure and manual provided or available upon request)
	Detailed (e.g., statistics presented) information indicating good validity and reliability in at least one peer-reviewed article
Approaching well-established assessment	The measure must have been presented in at least two peer-reviewed articles, which might be by the same investigator or investigatory team
	Sufficient detail about the measure to allow critical evaluation and replication (e.g., measure and manual provided or available upon request)
	Validity and reliability information presented in either vague terms (e.g., no statistics presented) or moderate values
Promising assessment	The measure must have been presented in at least one peer-reviewed article
	Sufficient detail about the measure to allow critical evaluation and replication (e.g., measure and manual provided or available upon request)
	Validity and reliability information presented in either vague terms (e.g., no statistics presented) or moderate values

descriptions and references were provided to assist the work groups in evaluating the psychometric properties of instruments. In addition, the SPP-ATF indicated that the work groups should pay special attention to several issues. First, the work groups were to evaluate the clinical and research utility of the measures, especially in terms of whether the tools were best suited for clinical work, research endeavors, or both. Second, the SPP-ATF stressed the importance of identifying whether the measure that was applicable for use with linguistic minorities, as well as racial or ethnic populations. A final area of emphasis was on evaluating whether results from the measure led to clear treatment implications. The SPP-ATF argues that assessment tools should ideally serve the dual purposes of extending the understanding of the patient or pathology being studied and also lead to clear recommendations for developing and implementing treatment (Blount et al., 2000; Hayes et al., 1987; Nelson-Gray, 2003; Ollendick, 2003).

To ensure that sufficient practical information was included in the review, the SPP-ATF encouraged the work groups to provide an appendix or table with detailed information about the measures. Specifically, the appendix should include the following: the name of the measure, the central references, a description of the measure (e.g., purpose, age range, population, format, administration, and scoring), how to obtain the measure, psychometric properties (i.e., reliability and validity), the clinical utility of the measure, evaluation of the measure (i.e., strengths and weaknesses), and the Task Force Criteria Rating (Table I).

Although the SPP-ATF provided guidelines in how to systematically critique the measures and what to include in the appendix, the work groups were given latitude in the structure of the review articles. For instance, some workgroups working with fewer measures might present information on all available instruments, whereas other reviews might highlight representative assessments. The SPP-ATF also recognized that subjective appraisal would be involved, and thus, the reviews should be regarded as guidelines and recommendations rather than as prescriptive rules.

Concluding Comments

The overarching goals of this series are to update the field on the state of assessment in pediatric psychology, identify and critique the best and most widely used and studied measures, provide practical suggestions and guidelines for the selection of instruments, detail directions for future work, and help integrate science and practice. The SPP-ATF acknowledges some of the pitfalls in such an endeavor. For instance, measures might have been inadvertently missed in the compilation or selection process, and there is subjectivity inherent in critiquing and reviewing the selected measures. In closing, a critical examination of assessment measures used in pediatric psychology is essential, and the articles in this series represent a crucial first step in identifying and disseminating scientifically sound assessment tools for use with pediatric populations.

Acknowledgments

The authors thank the listserv members of Division 54 (Society of Pediatric Psychology) of the American Psychological Association for their assistance in compiling and rating the measures that will be reviewed in this series.

Received January 21, 2004; revisions received November 9, 2005 and February 14, 2006; accepted March 20, 2006

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Beutler, L. E. (2000). Empirically based decision making in clinical practice. *Prevention & Treatment, 3*, 27. Retrieved August 15, 2003, from http://journals.apa.org/prevention/volume3/pre0030027a.html

Blount, R. L., Bunke, V. L., & Zaff, J. F. (2000). The integration of basic research, treatment research, and clinical practice in pediatric psychology. In D. Drotar (Ed.), *Handbook of research in pediatric and child clinical psychology: Practical strategies and methods* (pp. 491–510). New York: Kluwer Academic/Plenum Publishers.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, *52*, 685–716.

Frick, P. J. (2000). Laboratory and performance-based measures of childhood disorders: Introduction to the special section. *Journal of Clinical Child Psychology*, *29*, 475–478.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, *42*, 963–974.

La Greca, A. M. (1994). Editorial: Assessment in pediatric psychology: What is a researcher to do? *Journal of Pediatric Psychology*, *19*, 283–290.

La Greca, A. M., & Lemanek, K. L. (1996). Assessment as a process in pediatric psychology. *Journal of Pediatric Psychology*, *21*, 137–151.

Lonigan, C. J., Elbert, J. C., & Johnson, S. B. (1998). Empirically supported psychosocial interventions for children: An overview. *Journal of Clinical Child Psychology*, *27*, 138–145.

Mash, E. J., & Hunsley, J. (2005). Special section: Developing guidelines for the evident-based assessment of child and adolescent disorders. *Journal of Clinical Child and Adolescent Psychology*, *34*, 362–379.

Nelson-Gray, R. O. (2003). Treatment utility of psychological assessment. *Psychological Assessment*, *15*, 521–531.

Ollendick, T. H. (1999, Spring). President's column: Empirically supported assessment for clinical practice: Is it possible? Is it desirable? *Clinical Psychologist*, *52*, 1–2.

Ollendick, T. H. (2003, Summer). President's message: The role of assessment in evidence-based practice. *Clinical Child and Adolescent Psychology Newsletter*, *18*, 1–2.

Quittner, A. L. (2000). Improving assessment in child clinical and pediatric psychology. In D. Drotar (Ed.), *Handbook of research in pediatric and child clinical psychology: Practical strategies and methods* (pp. 119–143). New York: Kluwer Academic/ Plenum Publishers.

Roberts, M. C., & McNeal, R. E. (1995). Historical and conceptual foundations of pediatric psychology. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (pp. 3–18). New York: The Guilford Press.

Spirito, A. (1999). Introduction to special series on empirically supported treatments in pediatric psychology. *Journal of Pediatric Psychology*, *24*, 87–90.

Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, *48*, 3–23.