Journal of Clinical Child Psychology
2000, Vol. 29, No. 4, 475–478

Laboratory and Performance-Based Measures of Childhood Disorders: Introduction to the Special Section

Paul J. Frick
Department of Psychology, University of New Orleans

Serves as an introduction to a special section of the journal on laboratory and performance-based measures of childhood disorders. The articles in the special section were part of the work of a task force established by Division 12 of the American Psychological Association on "Upgrading the Science and Technology of Assessment and Diagnosis." In this introduction, I raise a number of issues involved in the use of laboratory and performance-based measures for the assessment of childhood psychopathology that cut across the different disorders covered in the special section. Some of these issues are common to most techniques used in the assessment of childhood psychopathology; others are more specific to this particular method of assessment. However, by focusing on these issues related to the use of laboratory and performance-based measures, it will hopefully encourage a critical examination of all techniques currently being used in the assessment of psychopathology and highlight important issues involved in translating measures that were developed primarily for use in research into forms that are useful in clinical practice.

Over the past decade, Division 12 (Clinical Psychology) of the American Psychological Association has been instrumental in promoting a critical examination of the interventions used in the treatment of psychopathology through the documentation and dissemination of empirically supported approaches to treatment (Task Force on the Promotion and Dissemination of Psychological Procedures, Division of Clinical Psychology, 1995), a trend recently extended specifically to the treatment of childhood psychopathology (Lonigan, Elbert, & Johnson, 1998). This process has sparked a lively and spirited debate that has encouraged a critical appraisal of the current treatment technology in psychology, and this debate has helped to define important goals for the immediate future of the field (e.g., Borkovec & Castonguay, 1998; Kazdin & Kendall, 1998). As a result, even if a consensus as to the best method for advancing treatment technology has not emerged from this process, it has clearly discouraged a "business as usual" approach to psychological intervention by forcing the field to critically examine the methods by which it seeks to treat individuals with psychopathological conditions.

Requests for reprints should be sent to Paul J. Frick, Department of Psychology, University of New Orleans, 2001 Geology and Psychology Building, New Orleans, LA 70148. E-mail: pfrick@uno.edu

Task Force for Upgrading the Science and Technology of Assessment and Diagnosis

In October of 1998, Division 12 , under the leadership of Thomas Ollendick, formed a task force to begin a similar critical examination process in the area of assessment and diagnosis. The overarching goal of the task force was to spark a scholarly debate concerning the current practice of assessment and diagnosis in psychology and the best goals and methods for improving the science and technology of this practice. Such a debate seemed especially important in the practice of assessment and diagnosis for a number of reasons. First, the choice of methods used in the practice of psychological assessment has often been driven more by an allegiance to a theoretical orientation that underlies a particular method of assessment or, even more problematic, by an allegiance to a particular assessment technique rather than being based on the most current understanding of the psychological construct that is being assessed and selecting the best method of assessment based on this understanding (Kamphaus & Frick, 1996). Second, this method of practice has led to a dichotomy between the measures of psychopathology being used in research and the assessment techniques being used in clinical practice. Importantly, this dichotomy is not solely due to a reluctance to use novel techniques by practicing psychologists but also by a failure of researchers to develop measures in a manner that

make them useful in clinical assessments. Third, if the field is to improve its treatment technology by being guided by advances in basic research (see Kazdin & Kendall, 1998), it is critical to translate the measures that are being developed and used in research to study psychopathological conditions into forms that can also be used in practice.

As a result, promoting a scholarly debate on the current state and future of psychological assessment is an important goal. However, it is also a very broad goal. Therefore, to begin this process, the task force chose to focus on one type of assessment technology: the use of laboratory and performance-based measures in the assessment of childhood psychopathology. It is important to note that the choice of this type of assessment was not predicated on the assumption that such techniques necessarily are the most useful methods for enhancing our assessment technology. Instead, one of the main reasons for this choice was that these measures provide an excellent example of an assessment technique that is being increasingly used in research on childhood psychopathology but, with some notable exceptions, is rarely used in clinical practice. As a result, these methods provide a good illustration of the duality between research and practice in psychological assessment. Also, a critical evaluation of these techniques provides the basis for raising a number of important issues involved in translating assessment tools developed for the study of a psychological construct into forms that can be used widely in practice.

Once this topic was chosen, experts in the assessment of various forms of childhood psychopathology were invited to serve on the tasks force and begin reviewing the existing laboratory and performance-based measures within their area of expertise. The initial findings from these reviews were presented at a symposium sponsored by Division 12 at the 1999 national convention of the American Psychological Association. Based on feedback provided at this symposium, the articles for this special section were developed and underwent the peer review process prior to publication. As a result, the articles presented in this section were developed by experts in the assessment of childhood psychopathology and have already been subjected to review and comment by other authorities in the field. Because the focus of the task force is to encourage a dialogue over the current state and the future of assessment and diagnosis, I strongly encourage comments directed to me, as chair of the task force, and to the authors of the articles.

Critical Issues

In the course of putting together this special section, a number of important issues arose that provide an important context for the articles that are included in this special section. One issue involved defining the boundaries of what would be considered a laboratory and performance-based measure. Generally, a technique was considered for inclusion if the child's behavior was observed under standardized conditions, usually involving stimuli designed to evoke the specific behavior of interest. This definition led to the exclusion of naturalistic observational methodologies in which conditions are not standardized. Also, laboratory techniques designed primarily to elicit and observe physiological or neuropsychological responses were also not included. Even with this rather straightforward definition, there were a number of techniques that, depending how one viewed them, could or could not fit this definition. Final decisions on the inclusion of techniques were left with the authors of each article. An additional decision on content was to not only consider techniques that assess the primary symptoms of the disorder under consideration but to also include laboratory techniques that assess potentially important processes that are related to the disorder. This decision had the potential of substantially broadening the content of each article and, again, the judgment of which potentially important processes to include was left to the individual author.

In preparing the articles for the special section, it also became clear that there were a number of important theoretical, psychometric, and ethical issues related to the use of laboratory and performance-based measures that cut across the individual psychopathological conditions. Some of these issues are common to the development of any assessment technique for clinical use with children and adolescents. For example, the issue of whether or not the technique is appropriate across multiple developmental levels and for different ethnic groups is an issue that is critical to all assessment techniques used for children and adolescents, albeit an issue that too often is ignored. Similarly, the issue of whether the measure is tapping a process that is specific to a given disorder or one that is more generally related to psychopathological conditions in children is relevant to each article in this special section. Again, this issue of the specificity of the information provided by the assessments techniques is an important one for all measures used in the assessment of childhood psychopathology (Kamphaus & Frick, 1996).

However, in addition to these issues common to most, if not all, techniques used in the assessment of children and adolescents, the articles in this section highlight a number of issues that are more specific to the clinical use of laboratory and performance-based measures. First, most of the measures reviewed in this special section were developed to test very specific, theory-driven hypotheses about the underlying processes to childhood disorders. A good example are the measures of cognitive functioning used in the assessment of autism and other pervasive developmental dis-

orders that are reviewed by Klinger and Renner (this issue). This characteristic is important because it explicitly ties the assessment technique to the most current theoretical models being developed to understand a disorder. However, it also means that the measures are frequently modified across studies to test specific hypotheses, and as a result, any one form of the measure typically has limited data on its psychometric characteristics. Also, the type of data that is collected for the theory testing purposes for which the measures were developed are often correlational and typically involve comparisons of mean differences between a diagnostic and control group. Such group data provide only limited information for diagnostic purposes when one must interpret the scores of an individual child.

Second, each of these tasks were designed to conform to standardized procedures in which specific stimuli are presented to elicit a desired behavior in the child. Such procedures allow for systematic observation of the behavior of interest and give important information on factors that may be controlling such behavior. However, by design, the procedures are standardized and are only laboratory analogues to what a child might experience in his or her natural environment. As a result, there is always the question of the ecological validity of the behavior that is observed under these conditions. There are data on some of the techniques correlating the results from the laboratory measures with other, more ecologically valid measures, such as correlating the results of a laboratory measure of stealing with parent ratings of stealing in the home (see Frick & Loney, this issue). However, such checks of the ecological validity of the measures have not been uniformly done, and it leaves open the question of what information the laboratory and performance-based measures add to measures that may be more ecologically valid.

Third, this question of incremental validity is part of a broader issue of establishing the clinical utility of these measures. As mentioned previously, the development of most of these measures was not guided by their potential clinical uses, but most were developed to test very specific scientific hypotheses. As a result, establishing their clinical utility has not been a major focus of their development. The article by Vasey and Lonigan (this issue) on laboratory measures used in the assessment of anxiety disorders provides an excellent discussion of the many different types of clinical utility that are possible. Many of the measures reviewed in the special section have shown indirect clinical utility by providing information that advances our understanding and conceptualization of a particular disorder or construct, and hopefully, this more advanced understanding can lead to improvements in treatment.

Unfortunately, evidence for the direct clinical utility of these measures is uniformly minimal. In most cases, there is no evidence that the use of these measures enhances treatment outcome, such as by designating important processes that should be targets of intervention. Even if the direct clinical utility has not been established for many of the measures, a number of techniques reviewed in the special section show clear promise for such utility in the immediate future. However, because of the cost associated with the use of many of the techniques and because of some of the ethical issues involved in their clinical use, which are discussed later, it is also important that these measures show incremental clinical utility. That is, it is important to determine if these measures provide information that advance treatment above and beyond the information provided by other techniques. None of the measures reviewed in this section has established this type of utility.

Fourth, (a) performance-based measures often elicit behavior that is clearly maladaptive or unacceptable, and (b) to elicit such behavior, some level of deception is often involved in administering the techniques. In many cases the stimuli and situations to which the child is exposed are quite innocuous and are similar to those to which the child encounters in daily life, such as the social interactions involved in the Autism Diagnostic Observation Schedule reviewed by Klinger and Renner (this issue). However, the development and use of these measures raise a number of important ethical issues that need to be considered, especially when they are administered in the context of a therapeutic relationship. These important ethical issues are discussed in more detail in the article by Frick and Loney (this issue) on the use of performance-based measures of conduct disorders. These disorders provide a good context for raising these ethical issues because the primary symptoms of the disorders are antisocial and aggressive behavior that are clearly unacceptable. As a result, one must ensure that by placing a child in situations in which such behaviors are elicited, one does not inadvertently condone, encourage, or in any other way increase the likelihood that such behaviors will be displayed by the child outside of the assessment context.

Concluding Comments

As a result of the issues raised previously and others that are specific to the individual disorders that are reviewed in the separate articles, many of the authors conclude their articles with some very conservative and cautious recommendations for the appropriate clinical use of laboratory and performance-based measures. Does this mean that the choice for the initial focus of the Task Force on Upgrading the Science and Technology of Assessment and Diagnosis was misguided? My answer is *no* for several reasons. First, as mentioned previously, many of the issues raised in this introduction and throughout the special section are not specific to the use of laboratory and performance-based measures, such as the specificity of findings to a certain type of problem

and the failure to establish direct clinical utility for many measures. Hopefully, this critical evaluation of one type of assessment technique will spark similar evaluations of other techniques as well. Second, many of these issues raised in these articles are inherent in translating any type of assessment measure used in research into a form that is useful for clinical practice. Rather than discouraging further attempts to develop measures useful for both research and practice, hopefully these articles will clarify the many important issues involved in this process so that they can be addressed explicitly in the development of new measures. As such, these articles should spark a critical examination of current assessment practices and provide guidance for setting goals and selecting methods for advancing the science and technology of psychological assessment in the immediate future. These have been the two primary goals of the task force from its inception.

References

Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology, 66,* 136–142.

Kamphaus, R. W., & Frick, P. J. (1996). *Clinical assessment of children's personality and behavior.* New York: Allyn & Bacon.

Kazdin, A. E., & Kendall, P. C. (1998). Current progress and future plans for developing effective treatments: Comments and perspectives. *Journal of Clinical Child Psychology, 27,* 217–226.

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.

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