

Advancing Assessment of Children and Adolescents: Commentary on Evidence-Based Assessment of Child and Adolescent Disorders

Thomas M. Achenbach

Research Center for Children, Youth, and Families
University of Vermont

This article addresses the following cross-cutting issues evident in the special section of this issue: (a) current diagnoses do not provide adequate validity criteria; (b) the heterogeneity and comorbidity of target problems raise taxonomic challenges; (c) accurate assessment requires integration of multisource data; (d) developmental variations must be accommodated; (e) appropriate norms are needed; and (f) categorical and quantitative approaches are not incompatible. Less evident in the special section articles but equally important are the need to cope with multicultural issues, avoid premature closure regarding diagnostic labels, assess caregivers, and standardize broad-spectrum assessment procedures. Studies and use of evidence-based treatment (EBT) should be linked to evidence-based assessment (EBA) to advance both EBT and EBA.

I am grateful to Eric Mash and John Hunsley for organizing this much-needed special section and for inviting me to comment on the issues it raises. Like Mash and Hunsley (2005), I use the term evidence-based assessment (EBA) to include both empirically based and evidence-based assessment, just as evidence-based treatment (EBT) includes both empirically based and evidence-based treatment. I agree wholeheartedly that EBA deserves greater attention in light of “the omission of assessment considerations in recent efforts to promote EBTs” (Mash & Hunsley, 2005). Without accurate identification and measurement of the problems to be treated and of outcomes following treatment, the potential benefits of EBT cannot be achieved. Although the need for EBA may have been neglected in the praiseworthy rush to advance EBT, a leading scholar of EBT has now endorsed EBA (Weisz & Addis, in press). Furthermore, biologically based investigative procedures, such as neuroimaging and genetic analyses, also require accurate EBA of the target problems.

In reading the articles comprising this special section, I was impressed by their scholarly literature reviews and detailed documentation of issues specific to their respective topics. Despite the differences among the kinds of problems addressed, I was also struck by the degree to which certain issues seemed to recur in

relation to very different problems. The issues that cut across different problems include the following:

1. The lack of definitive assessment procedures and diagnoses for identifying the problems that each article targets.
2. The heterogeneity of the target problems, such that they cannot be viewed in terms of either single observed phenotypes or single inferred genotypes or other constructs.
3. The pervasiveness of comorbidity.
4. The need to obtain assessment data from multiple sources.
5. The challenge of integrating conflicting assessment data obtained from different procedures and different sources.
6. The need to deal with developmental differences, continuities, and discontinuities.
7. The need for appropriately representative normative samples.
8. Dialectical interplays between categorical and quantitative (including but not limited to dimensional) constructs, assessment procedures, and decision making.
9. The need for multiple stages of assessment, including initial broad-spectrum assessment to identify strengths and problems, narrower spectrum assessment of targets for intervention, ongoing assessment during the course of interventions, and outcome assessment.
10. The need to meet psychometric standards for assessment.

I am grateful to the following people for their helpful comments: Levent Dumenci, James Hudziak, Masha Ivanova, Robert Krueger, Stephanie McConaughy, Leslie Rescorla, David Rettew, and John Weisz.

Requests for reprints should be sent to Thomas M. Achenbach, Research Center for Children, Youth, and Families, University of Vermont, 1 South Prospect St., Burlington, VT 05401-3456. E-mail: Thomas.Achenbach@uvm.edu

People who work with children may be acutely aware of these issues, but life-span perspectives can

help us avoid artificial dichotomies between assessment of children (including adolescents) and adults. Because assessment challenges do not change radically at a particular age, lessons learned from assessment of children may be applicable to adults and vice versa. For example, although assessment of children is known to require multi-informant data (Issue 4), meta-analyses have shown that cross-informant correlations between reports of adult problems (Achenbach, Krukowski, Dumenci, & Ivanova, 2005) are not materially better than those for child problems (Achenbach, McConaughy, & Howell, 1987). Coupled with large discrepancies between clinical conclusions drawn from different sources of data for adults (Meyer, 2002), the modest cross-informant correlations suggest that multi-informant data are needed for assessment of adults, as well as children. As another example, the dialectical interplays between categorical and quantitative approaches (Issue 8) that are so salient in child assessment are also receiving attention from the primarily adult psychiatrist planners of the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-V)*, who concluded that it is "important that consideration be given to advantages and disadvantages of basing part or all of *DSM-V* on dimensions rather than categories" (Rounsaville et al., 2002, p. 12).

The need for multiple stages of assessment (Issue 9) and the need to meet psychometric standards (Issue 10) pertain to all kinds of assessment and were thoroughly covered throughout the special section. I therefore focus first on Issues 1 through 8 and some important connections among them. I then raise another set of issues that were not explicitly addressed in the special section in any detail but that are of increasing importance throughout the world. These are multicultural issues, to which EBA may make especially beneficial contributions. Thereafter, I suggest additional ways to advance assessment.

Issues 1 Through 8 and Relations Among Them

Because several issues are intertwined, I first lay foundations on which to base consideration of the cross-cutting issues. Thereafter, I refer more specifically to Issues 1 through 8.

With the possible exception of pediatric bipolar disorder (Youngstrom, Findling, Youngstrom, & Calabrese, 2005), the problem areas addressed by articles in the special section have extensive literatures on assessment. Yet, despite the extensive literatures, progress in understanding the problems may be impeded by failure to improve taxonomy of the problems to be assessed. To clarify relations between assessment and taxonomy, it is helpful to think of them as different stages of a process wherein (a) *assessment* refers to the identification

of the distinguishing features of problems, disorders, or cases; and (b) *taxonomy* refers to the grouping of problems, disorders, or cases according to their distinguishing features. Good assessment procedures are thus needed to construct and then to apply good taxonomies. Conversely, good taxonomies are needed to provide targets and guidance for good assessment procedures. The term *taxonomy* is preferred to *diagnostic system* here because *diagnostic* implies disease models that may be inappropriate for some of the problems discussed in the special section, such as conduct problems (McMahon & Frick, 2005). Furthermore, confusion may arise from the multiple meanings of *diagnostic*, including "diagnostic processes," which refer to gathering diagnostic data, and "diagnostic formulations," which refer to integrative statements about multiple aspects of a case.

Although significant efforts have been devoted to taxonomic issues, such as trying to separate anxiety from depression, assessment is often targeted on constructs that are taken for granted and are implicitly assumed to represent separate entities. For example, when strong associations are found between measures of anxiety and depression, such associations tend to be interpreted as indicating comorbidity between distinct disorders of anxiety and depression. The notion of comorbidity may thus be invoked to "explain" the co-occurrence of problems that are assumed a priori to represent different disorders. To determine whether particular measures actually do represent different disorders, appeals are sometimes made to "diagnoses" as the ultimate validity criteria. However, comorbidity among diagnoses is as rampant as among psychometric measures of problems (e.g., Klein, Dougherty, & Olino, 2005; McMahon & Frick, 2005; Silverman & Ollendick, 2005; Youngstrom et al., 2005). Furthermore, there is only meager support for the reliability and validity of many *DSM* diagnoses of childhood disorders, as indicated by several articles in the special section.

Diagnoses as Validity Criteria for Assessment and Taxonomy

Other than IQ and achievement tests for diagnosing mental retardation and learning disorders, the *DSM* does not specify assessment operations for childhood disorders (Widiger & Clark, 2000). Several kinds of diagnostic interviews have therefore been developed to operationalize *DSM* criteria by formulating the criteria in terms of questions designed to yield yes/no conclusions about whether each criterion for each disorder is met. In particular, the Diagnostic Interview Schedule for Children (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000) has been heavily supported by the National Institute of Mental Health and has been widely used to operationalize *DSM* diagnoses of children.

However, meta-analyses have shown that diagnoses made from the Diagnostic Interview Schedule for Children and other diagnostic interviews often fail to agree with diagnoses made by other means, such as comprehensive clinical evaluations for adult as well as child disorders (Rettew, Achenbach, Doyle, Ivanova, & Dumenci, 2005). The failure of diagnostic interviews to agree well with diagnoses made by clinical evaluations suggests that neither current versions of the diagnostic interviews nor clinical diagnoses provide adequate validity criteria against which to test assessment or taxonomic procedures. Furthermore, diagnostic categories and criteria for childhood disorders are still in flux, as indicated by substantial changes from the third edition (*DSM-III*; American Psychiatric Association [APA], 1980), to the third edition revised (*DSM-III-R*; APA, 1987), to the fourth edition (*DSM-IV*; APA, 1994), to the fourth edition text revision (*DSM-IV-TR*; APA, 2000), with further changes likely in *DSM-V*.

Taxonomic Challenges

In the APA's *A Research Agenda for DSM-V*, Kupfer, First, and Regier (2002) pointed out that, since Robins and Guze (1970) proposed the taxonomic approach that engendered *DSM-III* and its successors, (a) "Not one laboratory marker has been found to be specific in identifying any of the *DSM*-defined syndromes," (b) "epidemiologic and clinical studies have shown extremely high rates of comorbidities among the disorders," (c) "epidemiologic studies have shown a high degree of short-term instability for many disorders," and (d) "lack of treatment specificity is the rule rather than the exception" (Kupfer et al., 2002, p. xviii). Consequently, to advance the official nosology, as well as assessment of the problems reviewed in the special section, it may be time to examine the prevailing taxonomic concepts and their relations to assessment.

The challenges raised by Issues 1 through 8 concern taxonomy as well as assessment, because they are unlikely to be mastered without taxonomic advances. For example, the lack of definitive assessment procedures and diagnoses for identifying the target problems suggests a need for improving taxonomic constructs on which to focus assessment.

Heterogeneity and comorbidity. The heterogeneity of the target problems (Issue 2) raises taxonomic questions about the boundaries between different kinds of problems, many of which occur together in varying combinations. The co-occurrence of different kinds of problems is often explained in terms of comorbidity (Issue 3), implying that the different kinds of problems are symptoms of different disorders that happen to occur together. However, the lack of definitive ways of identifying the target problems (Issue 1) and the heterogeneity of the target problems (Issue 2) cast doubt on comorbidity as an explanation for the co-occurrence of different kinds of problems.

ogeneity of the target problems (Issue 2) cast doubt on comorbidity as an explanation for the co-occurrence of different kinds of problems.

If we lack well-validated boundaries between taxa for child psychopathology, we cannot expect our assessment procedures to validly determine whether a child has (a) two or more distinct disorders; (b) problems that, like fever, may be nonspecific symptoms of different disorders; or (c) heterogeneous problems that do not necessarily signify different disorders. High rates of comorbidity are found between diagnoses such as anxiety disorders with depressive disorders, conduct disorder with oppositional defiant disorder, and attention deficit hyperactivity disorder (ADHD) with many other disorders (Klein et al., 2005; McMahon & Frick, 2005; Silverman & Ollendick, 2005). It may therefore be worth basing taxonomic constructs on empirically identified associations among problems rather than on assumptions about which problems are symptoms of different disorders.

Multisource data. Although the need for multisource data about children's problems (Issue 4) is now widely recognized, there is less recognition of the need for taxonomic constructs that take account of cross-source discrepancies and for assessment procedures that integrate often conflicting multisource data in taxonomic decisions (Issue 5). If no single source of data can provide a gold standard, and if different sources validly capture variations in children's functioning, then taxonomic constructs need to advance beyond yes/no definitions of criterial attributes. Cross-source discrepancies continue to sow confusion because of failures to consistently obtain multisource data in standardized ways and to rigorously integrate multisource data into well-validated conclusions about whether and for what a child needs professional help.

Developmental considerations. The need to deal with developmental differences, continuities, and discontinuities (Issue 6) raises taxonomic questions about which problems should be viewed as representing the same taxonomic constructs at different ages and which should be viewed as representing different taxonomic constructs. For example, longitudinal studies have provided evidence for developmental changes such as the following: (a) anxiety problems in early childhood are often followed by depressive problems (Kovacs, Gatsonis, Paulauskas, & Richards, 1989; Roza, Hofstra, van der Ende, & Verhulst, 2003), (b) hyperactivity and impulsivity often become less salient than attention problems (Pelham, Fabiano, & Massetti, 2005), (c) oppositional problems often precede conduct problems (McMahon & Frick, 2005), and (d) aggressive behavior is often followed by intrusive behavior that is not overtly aggressive (bragging, showing off, talking too much; Achenbach, Howell, McConaughy, & Stanger,

1995). Should these developmental changes in phenotypic patterns be interpreted as indicating categorical changes in disorders (e.g., from anxiety disorders to depressive disorders and from oppositional defiant disorder to conduct disorder)? Or are there taxonomic continuities in the genotypes that underlie the phenotypic changes? If there are taxonomic continuities, then assessment procedures should be geared to these continuities instead of assuming that the phenotypic changes necessarily reflect changes from one type of disorder to another.

Appropriately representative normative samples. The need for appropriate norms is related to developmental issues in that norms are needed for each developmental period to determine whether and in what way children's functioning differs from that of their peers. For example, to determine whether a child is deviant with respect to anxiety problems at one age and then becomes deviant with respect to depressive problems at a later age, we need to compare the child with norms for large, representative samples of peers at each age. Because there may be important gender, age, and informant differences in reported problems, the norms should be specific for each gender within particular age periods, as assessed by different informants such as parents and teachers, as well as by self-reports. The prevailing *DSM* nosology lacks normative reference points. Instead, it uses similar diagnostic criteria and clinical cut points for both genders and diverse ages, regardless of the source and type of assessment data.

Some psychometric instruments have "norms" based on convenience samples that were not carefully selected to be representative of relevant populations. Norms are also dubious if they are not based on high completion rates among respondents randomly selected for the normative samples. For some instruments, normative samples have been "poststratified" by selectively omitting or weighting cases to duplicate particular demographic distributions. However, poststratification cannot compensate for (a) failure to draw samples from explicit sampling frames, (b) failure to use probability sampling to select participants who are representative of particular populations, or (c) high attrition from the target samples.

Categorical and quantitative approaches. I referred to "dialectical interplays" between categorical and quantitative constructs, assessment procedures, and decision making (Issue 8) to emphasize that categorical and quantitative approaches need not be mutually exclusive (e.g., Pickles & Angold, 2003). Instead, they can coexist and enrich each other in various ways. For example, if a disorder is proven to be caused by a particular allele of a particular gene, the presence of that allele can be viewed as a categorical cause of the disorder. However, even single allele conditions may

be manifest in various ways and degrees in different individuals and at different points in the development of a disorder. Furthermore, many kinds of psychopathology are apt to involve multiple interacting environmental and genetic factors, as demonstrated by interactions between life stress and genetic vulnerability to depression (Caspi et al., 2003). Developmentally appropriate quantitative assessment may therefore be needed for research on the course, treatment, and outcome of even single allele disorders, as well as for clinical assessment to evaluate a particular child's needs for help.

Quantitative methods and models can contribute to exploratory and confirmatory research, as well as to theory, assessment, and clinical decisions. This would be true even if the custodians of the official nosology continue to require that diagnoses be verbally defined in terms of categorical, yes/no decisions. An advantage of quantitative assessment is that it can be used to assign individuals to categories by using cut points and probabilities while still preserving information about quantitative gradations in functioning. EBA can be especially effective in providing both categorical and quantitative information if assessment data are not limited to a priori categorical models. Despite tendencies to view categorical and quantitative approaches as mutually exclusive, EBA can include beneficial aspects of both approaches.

Although "dimensional" approaches are often viewed as the main alternatives to categorical approaches (e.g., Rounsaville et al., 2002), dimensional approaches constitute only one subset of ways to use quantification in assessment and taxonomy. In addition to quantitative dimensions, quantification can also be used to construct typologies of problems via cluster analyses of profiles of scale scores. After the typologies have been constructed, quantification can be used to assess the degree to which a child's profile of scale scores matches each of the profile types identified for that child's age and gender. For example, intraclass correlations can be computed between the child's scale scores and the scores that define each profile type. The child can then be assigned to the type for which the intraclass correlation is highest, as detailed elsewhere (Achenbach, 1993). Additional ways to use quantification include construction of typologies via latent class and latent profile analyses, followed by assignment of cases according to the probabilities of their membership in each class.

Multicultural Issues

Considering that the articles in the special section focus intensively on particular categories of clinical problems, it is understandable that overarching multicultural issues are not addressed in any detail. How-

ever, multicultural issues are relevant to all kinds of problems for several reasons. One reason is that many mental health professionals serve children from backgrounds very different from their own. These children include refugees, immigrants, and native-born minority children. A second reason is that mental health professionals may be called on to assist in cultures that lack sufficient indigenous personnel and to communicate with and train people from other cultures. A third reason is that mental health research is becoming increasingly multicultural as more researchers communicate and collaborate with colleagues from other cultures.

To advance both mental health knowledge and services, assessment needs to take account of possible cultural variations in the manifestations and patterning of problems, in the feasibility of using particular procedures, and in etiologies, courses, and outcomes of disorders. Until assessment instruments, diagnostic criteria, and taxonomic constructs are tested and supported in multiple cultures, we will not know whether they are valid for children in general or only for children in a single culture. By systematically obtaining and comparing data for children from different cultures, EBA can help to determine the extent to which particular constructs, norms, and procedures are generalizable across different cultural groups. When this has been done, considerable similarity has been found in both the patterning and prevalence of problems among children from many very different cultures (Ivanova et al., 2005; Rescorla et al., 2005). However, if major differences are found between particular groups, EBA can be more readily adapted to take account of the differences than can assessment procedures that apply the same a priori constructs and cut points to everyone.

Additional Ways to Advance Assessment

Because assessment of psychopathology requires multiple sources of data and multiple ways to obtain and combine the data, no single method is sufficient for all kinds of problems. However, as with my suggestions for dealing with the issues already discussed, some additional suggestions may help to advance assessment by capitalizing on particular strengths offered by EBA, as discussed in the following sections.

Avoid Premature Closure on Diagnostic Labels

Perhaps owing to the explicit diagnostic criteria that have become dominant since the publication of *DSM-III* (APA, 1980), referrals often include diagnostic labels, such as ADHD, depressive disorder, conduct disorder, anxiety disorder, or oppositional defiant disorder. Furthermore, the emergence of numerous focalized treatments and of services that specialize in

particular disorders, such as ADHD or anxiety disorders, may reinforce the tendency to impose diagnostic labels on children even before assessment begins. Another contributor to this tendency may be third-party payer requirements for *DSM* diagnostic labels to justify reimbursement. Reimbursement algorithms may also shape practitioners' preferences for particular diagnostic labels.

If we want assessment to provide the most accurate possible picture of each child for research or clinical purposes, we need to avoid premature closure with respect to diagnostic labels for the child's problems. It is especially important to avoid premature closure when certain diagnostic labels are prepotent, owing to their use in referral complaints, their role in selecting cases for specialty clinics, and their effects on reimbursement contingencies. For example, if a child is referred for ADHD, we should resist the temptation to use only assessment procedures for ADHD. Even if the referral agents accurately report that a child fails to pay attention, instruments that assess only ADHD are not adequate for evaluating the child's functioning or for making a diagnosis. Many children who fail to pay attention have other important problems that are not fully subsumed by ADHD criteria. Not only may ADHD problems be comorbid with other problems, but the ADHD problems may be secondary to depression, anxiety, or other problems. To reap the benefits of EBA, we should always assess a broad spectrum of problems and competencies, in addition to the problems and diagnostic labels that are most salient in referrals. As illustrated by Pelham et al. (2005) with respect to ADHD, "the diagnosis per se has not been demonstrated to have treatment utility."

Assess Caregivers

An additional way to capitalize on the strengths of EBA is to include assessment of a child's parents and other primary caregivers, such as grandparents. There are many ways in which clear documentation of caregiver characteristics can contribute to understanding and helping children. To document important characteristics of caregivers, EBA instruments for adults can be used to assess children's caregivers. For example, caregivers can be asked to complete adult self-report instruments that assess the caregivers' own problems and adaptive functioning. To provide additional perspectives, parallel collateral-report instruments can be completed by each caregiver about the other. Versions of EBA adult self-report and collateral-report instruments that parallel EBA instruments for children can be used to identify similarities and differences between patterns of functioning found for children and their caregivers (Achenbach, Newhouse, & Rescorla, 2004; Achenbach & Rescorla, 2003). By obtaining parallel EBA data for children and their caregivers, mental

health professionals can plan interventions more effectively than when focusing only on the children. In some cases, for example, assessment of caregivers may tell us that interventions should focus more on them than on the child who is the identified patient.

Standardize Initial Broad-Spectrum Assessment Procedures

Standardized tests are routinely used to assess children's ability and achievement. Although different tests of a particular cognitive construct are not totally interchangeable with each other, several tests of each construct correlate well enough and function similarly enough to become accepted as alternative ways of operationalizing ability and achievement. Standardized assessment of child psychopathology, however, has a shorter history. Numerous rating forms, questionnaires, tests, interviews, and observational procedures have been developed, but they are less routinely used to standardize assessment of child psychopathology than are tests of ability and achievement. Because psychopathology encompasses such a broad array of possibilities, routine standardized assessment of psychopathology may seem less practical than standardized assessment of ability and achievement. However, widespread use of EBA could help to promote greater standardization of broad-spectrum assessment of psychopathology to establish common baselines for research and clinical assessment of most children.

Establishment of common baselines for broad-spectrum assessment would not preclude more customized use of specialized procedures in the subsequent, narrower-spectrum phases of assessment. However, standardization of the broad-spectrum phase of assessment could help to prevent premature closure regarding diagnoses and to ensure that multiple areas of functioning are routinely assessed in ways that can be readily documented, considered, and communicated to others before more specialized assessment is done.

To protect the integrity of both broad-spectrum and narrower spectrum EBA, it is essential for assessment instruments to be used in their standard versions. Otherwise, the data will not be comparable to the reliability, validity, and other data obtained in research with the standard versions. Results obtained with abbreviated, altered, corrupted, and other "pirate" (Kollins, Epstein, & Conners, 2004) versions of instruments are apt to be quite different from results obtained with the standard versions. Users of pirate instruments may then wrongly imply that their results mean the same as if they were obtained with the standard instruments.

Summary and Conclusions

The EBT movement has contributed greatly to advancing the science and practice of interventions for psychopathology. However, for EBT to be appropriately applied and evaluated, EBA is needed to identify and measure the problems to be treated and the outcomes following treatment.

Mash and Hunsley (2005) have rendered a great service by organizing this special section on EBA and by setting the stage for scholarly articles on EBA in relation to particular categories of problems. The articles by leading specialists have abundantly documented needs for applying EBA to those problems.

In this article, I have identified issues that are evident in multiple categories of problems. As summarized in Table 1, I discuss these issues in terms of (a) the inadequacy of current diagnoses as validity criteria, (b) taxonomic challenges that are raised by the heterogeneity and comorbidity of the target problems, (c) the need to obtain and integrate multisource data, (d) the need to take account of developmental variations, (e) the need for appropriate normative samples, and (f) the dialectical interplay between categorical and quantitative approaches. I also address multicultural issues, which deserve far more attention, and additional ways

Table 1. *Main Points Pertinent to Advancing Child and Adolescent Assessment*

-
1. *DSM* Diagnoses Provide Inadequate Validity Criteria.
 - a. *DSM* does not operationally define diagnoses in terms of assessment operations, other than IQ and achievement tests for mental retardation and learning disorders.
 - b. Interviews that operationalize *DSM* diagnoses do not correlate well with diagnoses from clinical evaluations and other procedures.
 2. Taxonomic Challenges.
 - a. Heterogeneity and comorbidity of target problems.
 - b. Integrating multisource data.
 - c. Developmental considerations.
 - d. Appropriately representative normative samples.
 - e. Categorical and quantitative approaches.
 3. Multicultural Issues.
 4. Additional Ways to Advance Assessment.
 - a. Avoid premature closure regarding diagnoses.
 - b. Assess caregivers.
 - c. Standardize broad-spectrum assessment procedures.
-

Note: *DSM* = Diagnostic and Statistical Manual of Mental Disorders.

to advance assessment by avoiding premature closure regarding diagnoses, by assessing caregivers, and by standardizing broad-spectrum assessment procedures.

EBA and EBT are both needed to advance our ways of understanding, preventing, and ameliorating child psychopathology. EBA is especially vital for ensuring that EBT is properly used and evaluated. Without EBA, EBT may be like a magnificent house with no foundation. Studies and applications of EBT need to be linked to EBA to improve and extend both EBT and EBA.

References

- Achenbach, T. M. (1993). *Empirically based taxonomy: How to use syndromes and profile types derived from the CBCL/4-18, TRF, and YSR*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., Howell, C. T., McConaughy, S. H., & Stanger, C. (1995). Six-year predictors of problems in a national sample: III. Transitions to young adult syndromes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*, 658-669.
- Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin*, *131*, 361-382.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213-232.
- Achenbach, T. M., Newhouse, P. A., & Rescorla, L. A. (2004). *Manual for the ASEBA older adult forms and profiles*. Burlington: University of Vermont, Research Center for Children, Youth, & Families.
- Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA adult forms and profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. rev.). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text rev.). Washington, DC: Author.
- Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., et al. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science*, *301*, 386-389.
- Ivanova, M. Y., Achenbach, T. M., Dumenci, L., Rescorla, L. A., Bilenburg, N., Domuta, A., et al. (2005). *Testing the configural invariance of the Child Behavior Checklist syndromes in 29 cultures*. Manuscript submitted for publication.
- Klein, D. N., Dougherty, L. R., & Olino, T. M. (2005). Toward guidelines for evidence-based assessment of depression in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 412-432.
- Kollins, S. H., Epstein, J. N., & Conners, C. K. (2004). Conners' Rating Scales-Revised. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment: Vol. 2. Instruments for children and adolescents* (3rd ed., pp. 215-233). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Kovacs, M., Gatsonis, C., Paulauskas, S. L., & Richards, C. (1989). Depressive disorders in childhood: IV. A longitudinal study of comorbidity with the risk for anxiety disorders. *Archives of General Psychiatry*, *46*, 776-782.
- Kupfer, D. J., First, M. B., & Regier, D. E. (2002). Introduction. In D. J. Kupfer, M. B. First, & D. E. Regier (Eds.), *A research agenda for DSM-IV* (pp. xv-xxiii). Washington, DC: American Psychiatric Association.
- Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology*, *34*, 362-379.
- McMahon, R. J., & Frick, P. J. (2005). Evidence-based assessment of conduct problems in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 477-505.
- Meyer, G. J. (2002). Implications of information gathering methods for a refined taxonomy of psychopathology. In L. E. Beutler & M. L. Malik (Eds.), *Rethinking the DSM: A psychological perspective* (pp. 69-105). Washington, DC: American Psychological Association.
- Pelham, W. E., Jr., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 449-476.
- Pickles, A., & Angold, A. (2003). Natural categories or fundamental dimensions: On carving nature at the joints and the rearticulation of psychopathology. *Development and Psychopathology*, *15*, 529-551.
- Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Dumenci, L., Bilenburg, N., Domuta, A., et al. (2005). *Problems reported by parents of children ages 6 to 16 in 30 cultures*. Manuscript submitted for publication.
- Retten, D. C., Achenbach, T. M., Doyle, A. C., Ivanova, M. Y., & Dumenci, L. (2005). *Meta-analysis of agreement between standardized diagnostic interviews and clinical diagnoses*. Manuscript in preparation, University of Vermont, Department of Psychiatry.
- Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry*, *126*, 983-986.
- Rounsaville, B. J., Alarcon, R. D., Andrews, G., Jackson, J. S., Kendell, R. E., & Kendler, K. (2002). Basic nomenclature issues for DSM-V. In D. J. Kupfer, M. B. First, & D. E. Regier (Eds.), *A research agenda for DSM-V* (pp. 1-29). Washington, DC: American Psychiatric Association.
- Roza, S. J., Hofstra, M. B., van der Ende, J., & Verhulst, F. C. (2003). Stable prediction of mood and anxiety disorders based on behavioral and emotional problems in childhood: A 14-year follow-up during childhood, adolescence and young adulthood. *American Journal of Psychiatry*, *160*, 2116-2121.
- Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 28-38.
- 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.
- Weisz, J. R., & Addis, M. E. (in press). The research-practice tango and other choreographic challenges: Using and testing evidence-based psychotherapies in clinical care settings. In C. D. Goodheart, A. E. Kazdin, & R. J. Sternberg (Eds.), *The evidence for psychotherapy*. Washington, DC: American Psychological Association.
- Widiger, T. A., & Clark, L. A. (2000). Toward DSM-V and the classification of psychopathology. *Psychological Bulletin*, *126*, 946-963.
- Youngstrom, E. A., Findling, R. L., Youngstrom, J. K., & Calabrese, J. R. (2005). Toward an evidence-based assessment of pediatric bipolar disorder. *Journal of Clinical Child and Adolescent Psychology*, *34*, 433-448.

Received March 15, 2005

Accepted March 25, 2005

Copyright of *Journal of Clinical Child & Adolescent Psychology* is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.