Enhancing Reviews of Psychological Treatments With Pediatric Populations: Thoughts on Next Steps

Dennis Drotar, PhD
Case Western Reserve University School of Medicine

Objective: To critique the approach used by the Journal of Pediatric Psychology for reviews of empirical support for psychological treatments and provide recommendations to enhance the scientific review of interventions.

Methods: Application of the Chambless et al. (1996, 1998) criteria to select and describe empirically supported treatments was evaluated from the standpoint of method, theory, and clinical significance.

Results: While reviews of treatments based on the Chambless et al. (1996, 1998) criteria have provided valuable information, they are limited by reliance on statistical significance rather than effect sizes and also need to consider systematically the theoretical and clinical significance of intervention findings.

Conclusions: Reviews of psychological treatments conducted with pediatric populations would be enhanced by inclusion of information concerning effect sizes, theory specification, and clinical significance. In addition, broadening the focus of initial reviews to include meta-analyses, clinical significance, and theory-guided interventions would enhance the scientific knowledge base of psychological interventions with pediatric populations.

Key words: empirically supported treatments; Chambless criteria; psychological interventions; review.

In recent years, pediatric psychology researchers have underscored the central importance of developing a scientific knowledge base concerning psychological interventions with pediatric populations, especially for interventions that can be used in practice (Drotar, 1997; Kazak, 1998). To facilitate the development of empirically supported interventions and, in particular, to identify those that have been supported by scientific evidence, the Journal of Pediatric Psychology (JPP) has initiated a series of reviews to summarize the results of interventions with pediatric populations (Kazak, 1998) for the past several years (e.g., Holden, Deichman, & Levy, 1999; Janicke & Finney, 1999; McGrath, Mellon, & Murphy, 2000; McQuaid & Nassau, 1999; Mellon & McGrath, 2000; Powers, 1999; Walco, Sterling, Conte, & Engel, 1999). Based on the application of the Chambless et al. (1996, 1998) criteria that specify the evidence and methods to identify empirically supported interventions, publications in this series have advanced the science of psychological interventions with pediatric populations in several important respects. They have helped to sensitize the readership of JPP to the importance of empirically supported interventions with pediatric popu-
lations, identified interventions that have been supported by data, underscored methodological limitations of intervention studies, and suggested new research directions.

Yet, despite these advances, the approach that has guided the reviews of empirically supported intervention adopted by *JPP* has important but here-tofore unrecognized limitations as a strategy to advance the science of psychological interventions, especially the generalizability of intervention findings. Some of these limitations reflect the state of the art in research concerning interventions in the field of pediatric psychology. Others relate to policy and procedures concerning the reviews of empirically supported treatments in *JPP*. Written primarily to improve the reviews of empirically supported treatments, this critique also contains information of interest to the producers of such research.

This critique describes a central methodological problem: reliance on an approach that identifies and summarizes findings from interventions on the basis of statistical significance rather than effect sizes. A second problem is the failure to consider the theoretical significance of interventions in the approach to the reviews. A third, more pragmatic limitation is that published summaries have not considered critical questions related to the clinical significance of interventions. This commentary describes the implications of each of these limitations and offers recommendations to enhance the reviews of interventions with pediatric populations.


#### Methodological Limitations

The most significant methodological problem in the Chambless et al. (1996, 1998) criteria is the use of descriptive (Kazdin, 2000) or vote-counting procedure (Glass, McGraw, & Smith, 1981; Light & Pillemer, 1984) that relies heavily on the demonstration of statistical significance to select efficacious interventions and report the findings in the review. The fact that effect sizes of intervention findings were not described or considered in these reviews has several important limitations (Glass et al., 1981; Rosenthal, 1991). Larger samples typically produce greater statistically significant findings than small samples do, but this does not necessarily mean that such effects are more powerful or clinically significant. Very large samples have an abundance of power and can demonstrate statistically significant, albeit clinically trivial effects (Rosenthal, 1991). In general, treatments with larger effect sizes have a greater chance of affecting children in ways that make a meaningful difference to parents, teachers, medical care providers, and the children themselves (McCartney & Rosenthal, 2000). These criteria contain insufficient information and can be misleading, because readers cannot easily gauge the strength of intervention effects based on the information contained in reviews.

The absence of effect sizes in reviews of treatments also makes it impossible to compare effect sizes across different studies. In areas of research in

### History of Reviews of Empirically Validated Therapies in Clinical and Pediatric Psychology

In 1996, the Division 12 Task Force on Promotion and Dissemination of Psychological Procedures developed criteria for empirically validated therapies and examples of evidence-based treatments that met these criteria (Chambless et al., 1996). The primary criteria used to classify empirically validated treatments into two major categories, well-established or probably efficacious, include the demonstration of statistical significance either in group designs or single case experiments, the development of treatment manuals, clear specification of client samples, and the demonstration of effects by at least two different investigators or investigatory teams (Chambless et al., 1996, 1998). Based on application of these criteria, in 1999 *JPP* published the first of a continuing series of reviews of “treatments that work,” or empirically supported interventions in pediatric psychology. This series was initiated to stimulate and improve research concerning psychological treatments, to help guide selection of intervention strategies in practice, and to promote an interactive dialogue among scientists concerning intervention research (Kazak, 1998). Although the series has provided structure and consistency in the reviews of treatment-related research and facilitated the publication of high quality reviews of interventions and commentaries, it also has limitations.
which empirical support for certain types of interventions has been established, the key scientific questions often shift to documenting the specific characteristics of interventions, methods, or samples that result in the most powerful interventions. Comparison of effect sizes across different studies allows readers to judge the strength of findings across different interventions, sample characteristics, methods, or outcomes by using the common metric of effect size (Weiss, Weiss, Alecke, & Klotz, 1987). In the absence of such information, precise estimates of factors that influence the power or consistency of intervention effects across different studies cannot be made (Glass et al., 1981). The resulting lack of precision poses a special problem for areas of pediatric psychology research where relatively large numbers of psychological intervention studies have been conducted, such as with diabetes or asthma (McQuaid & Nassau, 1999). In such areas, information concerning effect sizes is necessary to recommend refinements in interventions to improve the strength and clinical significance of the effects.

**Limited Attention to Clinical Relevance and Significance**

Though identification of clinically significant interventions was never the primary goal of the JPP reviews, the hope was that the data generated from such reviews would eventually guide selection of intervention strategies in practice settings (Kazak, 1998). In fact, information contained in the reviews does allow readers to identify broad categories of interventions identified as empirically supported (e.g., behavioral interventions to improve the disease-related symptoms of children with asthma or diabetes; McQuaid & Nassau, 1999; or the role of the urine alarm in treatment of enuresis; Mellon & McGrath, 2000).

On the other hand, while findings based on the recent reviews of interventions published in JPP have given scientists and practitioners new information, they do not translate directly into helping practitioners select and implement interventions in practice settings, as was intended. One problem is that statistical significance does not translate into clinical significance (Kazdin, 2000; Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999), and clinical significance of effects is not addressed in the reviews.

The complex documentation of clinical significance can be done in a number of ways (see Jacobson, Roberts, Berns, & McGlinchey, 1999; Kazdin, 1999; Kendall, Marrs-Garcia, Nath, Sheldrick, 1999). For example, Kendall et al. (1999) have described the importance of documenting whether an intervention returned once extreme or abnormal cases to within normative ranges on key dependent measures. These authors advocate complementing traditional statistical analysis with equivalency testing to test the hypothesis that two groups are equivalent within a given range of scores (Kendall et al., 1999).

Kazdin (1999, 2000) has also described the importance of other methods of documenting clinical significance: (1) comparison with pretreatment samples (i.e., demonstration that the score for an individual at the end of treatment has departed substantially from the score for an individual at the beginning of treatment); (2) use of psychiatric diagnostic criteria (i.e., demonstrating that individuals no longer meet diagnostic criteria for disorder or problem for which they were referred for treatment); (3) inclusion of subjective evaluations of the impact of intervention (i.e., evidence that others interact with the child or the child himself or herself perceives a change that makes a clear difference in perceived functioning); and (4) social impact (i.e., demonstrating that treatment has affected the child’s functioning in ways that are important to society, families, or those who value referrals for intervention).

In fact, some of the empirically supported treatments reviewed in JPP have demonstrated effects on functional outcomes relevant to practitioners and parents. These include Finney, Lemanek, Cataldo, Katz, and Fuqua’s (1989) demonstration that brief targeted therapy for recurrent abdominal pain not only relieved children’s pain but reduced school absences and medical visits; Olness’s (1981) finding that imagery and self-hypnosis reduced pain frequency as well as the number of medications required to manage pain in a child with cancer; and Varni, Gilbert and Dietrich’s (1987) demonstration that progressive muscle relaxation, meditative breathing, and guided imagery lessened pain with reduced medications and hospitalizations, and increased activity for a child with hemophilia. However, for the most part, treatment-related studies contained in JPP’s reviews did not include information concerning clinical significance, even though some of these studies included clinically relevant outcome measures. For example, a number of the studies considered by McQuaid and Nassau’s (1999)
review of interventions in asthma and diabetes included measures of physiologic health status, such as pulmonary functioning in asthma or control of diabetes as measured by the hemoglobin AIC. To enhance the clinical significance of findings, it would have been useful for investigators to describe and analyze the degree of change in such measures as a function of treatment and to document whether these measures had improved to levels of adequate or more optimal health status.

A second problem that relates to the question of clinical significance is that the children who have participated in many of the intervention studies described in reviews *JPP* published are not necessarily the same children seen by practicing pediatric psychologists. As has been the case for psychotherapy research with children (Kazdin, 2000), research participants in controlled intervention studies generally have less severe, as well as more homogeneous problems than those who are generally seen by practitioners in pediatric settings. Moreover, many of the participants in preventive psychological interventions conducted and evaluated for some populations (e.g., children with chronic health conditions) do not demonstrate clinically significant psychological problems, which obviously limits the implications of these studies for the practice of pediatric psychology.

On the other hand, it should be noted that several studies reviewed in the *JPP* series did include criteria that helped to establish the clinical significance of the problems that were treated. For example, Stark and her colleagues’ study of the impact of behavioral group treatment for encopresis included failure of previous medical treatment for this problem as a requirement for participation in their studies (Stark et al., 1997; Stark, Owens-Stively, Spirito, Lewis, & Guevremont, 1990). Moreover, Finney et al.’s (1989) aforementioned study of treatment of abdominal pain required study participants to have episodes of abdominal pain sufficient to affect activity. Children in these treatment studies are likely to resemble children referred to pediatric psychologists for management of such problems. Consequently, the findings from treatments conducted with such samples would be more likely to generalize to the children seen by pediatric psychologists in practice settings as compared with psychological treatment studies of unselected populations such as children with asthma selected from a clinic sample (Vazquez & Buceta, 1993).

Neglect of Theoretical Issues in Evaluating Treatments

The Chambless et al. (1996, 1998) criteria were crafted to judge empirical support for interventions, whether an intervention works, and do not address theoretical questions such as why an intervention works. Though the criteria have served the purpose for which they were intended, the neglect of theory has certain consequences. For example, by not including theory explicitly in reviews, this approach neglects hypothesized processes of change and specification of how theory informs the operational principles and implementation of tested treatments (Douglas-Kelly, Nixon, & Bickman, 2000). If a treatment is found to be successful but the theory that guides the intervention is not specified, then the processes by which the effect was achieved cannot be discerned. On the other hand, if a theory of treatment-related change is clearly specified and linked to a successful test, then the theory of change is also supported. Consequently, successful tests of treatments closely linked to and guided by theory are much more likely to advance the development of generalizable principles of change, which is an important long-range scientific and clinical goal for treatments (Bickman, 1987; Kazdin, 2000; Weiss, 1995). Other researchers have noted that theory-based research is especially important to the field of pediatric psychology in general (Wallander, 1992), as well as to research on psychological interventions with children with chronic illness in particular (Bauman, Drotar, Leventhal, Perrin, & Pless, 1997), but it has been underemphasized.

Systematic inclusion of theory in the reviews of treatment effects also has pragmatic advantages for identifying the specific features of a successful intervention model that are most effective, for whom, and under what circumstances (Douglas-Kelley et al., 2000). On the other hand, if the framework or theory that guided how an intervention program was implemented is not clearly specified, negative intervention effects may be difficult, if not impossible, to interpret because they may reflect either a problematic program or intervention theory or a failure to accurately test the program theory (Douglas-Kelley et al., 2000). For example, Gustafsson, Kiellman, and Lederblad (1985), cited in McQuaid and Nassau (1999), found no effects of family therapy on disease activity and morbidity of children with asthma. However, it was not clear
why these authors expected family therapy to change the medical outcomes of children with asthma. The family therapy protocol may have been implemented as expected, but the mechanism or theory linking family therapy to improvement in medical outcomes for asthma may not have been plausible. Because the theory that guided the treatment model was not articulated in detail, it is impossible for readers to know.

**Recommendations**

Given the success of the reviews of empirically supported treatments in *JPP*, there is no reason to throw the baby out with the bath water by making substantial changes in their structure and format. Nevertheless, this analysis suggests that the scientific and pragmatic value of *JPP* reviews of empirically supported treatment would be enhanced by expanding the information contained in the reviews as well as broadening types of invited and published reviews.

**Expanding Information Contained in Reviews**

As shown in Table I, summaries of individual studies would be improved by including effect sizes to allow readers to compare the strength of treatment-related effects across different studies, populations, and treatment approaches. Information concerning effect sizes would not only improve the specificity of the conclusions and recommendations for future intervention research but provide data that can be translated into information of primary interest to policy makers (see McCartney & Rosenthal, 2000).

Readers’ appreciation of the clinical significance of empirically supported treatments would be facilitated by including the following information in reviews: (1) clinically relevant features of the samples that were studied (e.g., whether these children were referred for treatment, what these problems were, presence or absence of comorbid clinical problems or documenting clinical relevance of the problems studied); (2) explicit reference to the clinical significance (or not) of findings; and (3) clinical significance of findings obtained based on specific criteria (see Kazdin, 2000; Kendall et al., 1999).

Investigators who conduct treatments with pediatric populations can consider any number of options to describe the clinical relevance of a particular sample of children who participate in treatment. For example, studies of interventions to promote adherence to treatment for diabetes can include information about whether problems with adherence are judged to be interfering with medical management causing persistent hyperglycemia or episodes of ketoacidosis. Other options include description that establishes the clinical relevance of samples included in treatment studies, including cutoffs that describe clinically significant problems on measures of functioning (e.g., participation in activities or distress) and report changes as a function of treatment. Explicit reference to the clinical significance or potential clinical significance of findings obtained in treatment-related studies and in reviews would help readers to assess the relevance of the findings to their practice or research concerning treatment. However, reviewers will not be able to make explicit statements about the clinical significance of findings from reviews of specific studies unless researchers have focused their analyses on this question. For this reason, investigators who conduct treatment-related research should consider applications of methods to describe clinical significance discussed earlier (Kazdin, 2000; Kendall et al., 1999).

Finally, the value of reviews would be improved by specifying the theoretical basis of the tested interventions, including whether the intervention was informed by a specific theory, the framework or theory that guided implementation of the intervention, and to what extent this theory was supported

---

**Table I. Recommendations for Information Included in Reviews of Empirically Supported Treatments**

<table>
<thead>
<tr>
<th>Action and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Include effect sizes for treatment studies</td>
</tr>
<tr>
<td>A. Describe patterns of effect sizes for sampling characteristics, type of intervention, etc.</td>
</tr>
<tr>
<td>2. Specify whether the theory or conceptual model guided the implementation of the treatment</td>
</tr>
<tr>
<td>A. Specify the specific theory or conceptual model that was tested by the treatment</td>
</tr>
<tr>
<td>3. Indicate clinically relevant features of samples</td>
</tr>
<tr>
<td>A. Description of clinical problems</td>
</tr>
<tr>
<td>B. Description of comorbid problems</td>
</tr>
<tr>
<td>4. State potential clinical significance of the treatment effects</td>
</tr>
<tr>
<td>5. Specify implications of findings for practice and/or research in clinical settings</td>
</tr>
</tbody>
</table>
(Mellon & McGrath, 2000). It is difficult describe the theoretical underpinnings of treatments if in fact the original articles did not contain explicit reference to theory or tests of theory. Nevertheless, reviewers can pose relevant questions about theory that relate to the type and nature of the interventions studied.

One example of this approach was contained in McQuaid and Nassau’s (1999) review of empirically supported treatments of disease-related symptoms in asthma, diabetes, and cancer. In discussing interventions for childhood cancer, McQuaid and Nassau provided an instructive discussion of alternative mechanisms by which nausea and vomiting of patients who were undergoing chemotherapy for cancer might be alleviated by imagery-based or cognitive relaxation treatments. These included (1) distraction from sensations associated with chemotherapy by focusing attention on more pleasurable imagery; (2) reduction of distress or arousal or relaxation associated with specific images or activities may affect children’s symptom perception or physiologic processes such as gut motility or gastric acid production directly; (3) enhancement of children’s perceptions of control over their chemotherapy side effects by helping them engage in activities that have a positive effect. Including similar discussion of theoretical issues in other reviews would serve to stimulate refinements and tests of theory.

### Table II. Recommendations for Treatment Research and Invited Reviews in Different Phases of Research

<table>
<thead>
<tr>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Document whether treatments are effective</td>
<td>1) Use meta-analyses to document strength of findings and analyze pattern of effect sizes based on:</td>
<td>1) Test theories of how treatments work</td>
</tr>
<tr>
<td>2) Summarize and synthesize above findings</td>
<td>a) Settings in which interventions are conducted</td>
<td>a) Test empirical support of theory</td>
</tr>
<tr>
<td>a) Include effect sizes</td>
<td>b) Characteristics of population/sample</td>
<td>b) Test efficacy of treatments based on alternative theories</td>
</tr>
<tr>
<td>b) State whether theory guided treatment</td>
<td>c) Treatment methods</td>
<td>c) Test generalizability of theories of treatment process across settings and populations</td>
</tr>
<tr>
<td>c) Indicate clinically relevant features</td>
<td>d) Outcomes</td>
<td></td>
</tr>
<tr>
<td>d) State potential clinical significance of findings and implications of practice</td>
<td>2) Provide statistical documentation of clinical significance of treatment effects:</td>
<td>2) Summarize and synthesize findings</td>
</tr>
<tr>
<td>2) Provide statistical documentation of clinical significance of treatment effects:</td>
<td>a) Changes in health or mental health status from normal to abnormal</td>
<td></td>
</tr>
<tr>
<td>a) Changes in health or mental health status from normal to abnormal</td>
<td>b) Changes in parent or child appraisal of impact</td>
<td></td>
</tr>
<tr>
<td>b) Changes in parent or child appraisal of impact</td>
<td>c) Changes in functioning</td>
<td></td>
</tr>
<tr>
<td>3) Apply effective treatment models to clinical populations</td>
<td>3) Apply effective treatment models to clinical populations</td>
<td></td>
</tr>
<tr>
<td>4) Summarize and synthesize findings</td>
<td>4) Summarize and synthesize findings</td>
<td></td>
</tr>
</tbody>
</table>

### Expanding the Types of Invited Reviews

While these recommendations will enhance the quality and comprehensiveness of the information contained in reviews of empirically supported treatments, they also need to be tempered by realism. In treatment research, especially new areas, there may be insufficient information to conduct meta-analyses or to articulate explicit theories of why an intervention works. In such instances, the current or traditional approach to empirically supported reviews will be most useful, though the recommendations to improve the reviews by reporting effect sizes, articulating theory, and describing clinical significance are still relevant.

In other areas of work, the research concerning particular treatments may have developed to a point where other types of reviews would be more likely to extend scientific knowledge and should be invited. Table II describes the types of reviews that might be valuable at different phases of intervention research. For example, for treatments for which empirical support has not been established (Phase 1), the first scientific priority is to document whether the treatment is effective and to synthesize these findings. This is the purpose of the reviews of the empirically supported interventions published in JPP thus far. The description of Phase 1 reviews would be enhanced by inclusion of the information.
noted in Table I whenever such information is available from the original studies.

In Phase 2, studies of intervention research and reviews of this work have matured to the point where the primary scientific question is no longer to document whether an intervention works but to ask more refined questions, such as how strong are the treatment effects. What factors are associated with greater or lesser effects? Such questions are best addressed by using meta-analysis. Meta-analyses of treatments identified by reviews as established or efficacious would provide a comprehensive analysis of how effect sizes vary as a function of clinically relevant factors, such as sample characteristics (age, presence or absence of associated problems, such as compliance problems, type of intervention, etc.). Such information would not only document differences in the strength of treatment effects across different studies but would suggest next steps in treatment research that has been largely untapped in the field of pediatric psychology (for an exception, see Kibby, Tyc, & Mulhern, 1998).

Another potential focus of Phase 2 intervention research and the reviews of such work is to go beyond statistical significance to provide a formal documentation of clinical significance using appropriate statistical methods (e.g., measures that include cutoffs reflecting clinically significant change and indicators of improved functioning) (Kazdin, 1999; Kendall et al., 1999). (See previous section for examples.) Phase 2 reviews of intervention research might also focus explicitly on documenting other areas of clinical significance, cost-effectiveness, or social validity of interventions. Such information is of primary importance to practitioners, families, as well as insurers (Drotar & Zagorski, 2001; Walders & Drotar, 1999). Phase 2 research might also involve application of empirically supported treatments to a range of clinical samples and eventual reviews of this evidence.

In the third phase of intervention research, the primary scientific interest is in documenting why treatment works. This can be accomplished by testing the efficacy of interventions guided by specific theoretical models, determining which specific components of treatments are most critical to obtain effects, and testing generalizability of intervention models across settings. One of the most compelling reasons to conduct reviews of Phase 3 research relates to the need to generalize the principles of treatment models across clinical settings. The generalizability of treatments may ultimately depend on the degree to which theoretical principles of treatment models can be clearly specified and used to develop treatments sufficiently powerful to be effective in multiple settings (Kazdin, 2000). Consider some examples of such research. The theoretical underpinnings of multisystemic therapy (MST) has enhanced the operationalization of key principles of this model intervention and are critical ingredients in the generalizability of this approach to a range of settings (Henggeler & Randall, 2000). Similarly, the success of programmatic research on interventions to promote children’s coping during painful medical procedures relates in no small measure to attention to theory in the context of clinical application (Blount, Bunke, & Zaff, 2000a, 2000b).

The contributions of Phase 3 reviews are to identify and specify the theoretical models and framework that guide psychological treatments, describe the nature and level of empirical support of theoretical models based on empirical data, refine and develop these models based on treatment findings, and document the generalizability of theories of treatments for different clinical problems and pediatric populations.

Caveats and Final Recommendations

Readers should recognize several limitations of this critique and recommendations. With respect to the specific recommendations for broadening the types of reviews published, reviewers need to rely on available scientific evidence to construct their reviews. Consequently, it is difficult to separate what is contained in reviews of treatments from the nature of the treatment-related research reviewed. Meta-analysis is by no means a panacea for the methodological and pragmatic problems of conducting and reviewing intervention research and has several methodological limitations (DerSimonian & Levine, 1999; Durlak & Lipsey, 1991; Eggers & Smith, 1997). Moreover, the comprehensiveness and value of meta-analysis depends on the data generated concerning a particular treatment. In some areas of treatment research, small numbers of studies or sample sizes limit applications of meta-analysis, although statistical methods to manage smaller samples are available (Dunlap, Cortina, Vas-
low, & Burke, 1996; Levin & Wampold, 1999). Nevertheless, the success of meta-analytic approaches in describing scientific progress and suggesting new directions in child psychotherapy research (Kazdin, 2000) is a powerful argument for including them in the scientific agenda for reviews of empirically supported treatments in pediatric psychology.

Recommendations to solicit and publish reviews that focus on clinical and theoretical significance of interventions depend on the extensiveness and quality of information related to these topics available in original research reports of treatment research. My belief is that clinical relevance of science concerning treatments in pediatric psychology will be enhanced if investigators address clinical relevance and theoretical significance in designing their research. To accomplish this difficult task, investigators can build on available precedents for clinically relevant research concerning treatments (Pinto & Hollandsworth, 1989; Wysocki et al., 1997) and theory-guided intervention models (Ireys, Sills, Kolodner, & Walsh, 1996), heed cogent recommendations to conduct treatment research in clinical settings (Weisz, 2000), and include outcome measures that reflect clinical relevance (Kazdin, 2000) and social validity (Foster & Mash, 1999). The continued growth of a clinically relevant science of intervention research and the reviews that result from such research may depend on the flexibility and ingenuity of researchers in using multiple methods to document the impact of interventions with a wide range of populations. Case studies and series, program evaluations, and randomized controlled trials are all needed to develop a clinically relevant and generalizable science of intervention research. Reviews of empirically supported treatments should reflect this multifaceted research agenda.

Acknowledgement

The hard and extraordinary work of Susan Wood in processing this manuscript is gratefully acknowledged.

Received September 1, 2000; accepted February 26, 2001

References


groups or repeated measures designs. *Psychological Methods, 1*, 170–177.


