Scholarly Literature Review: Efficacy of Psychological Interventions for Pediatric Chronic Illnesses

Ivan L. Beale, PhD
School of Psychology, University of New South Wales

Objective To review empirical studies of the efficacy of psychological interventions as adjuvant therapies for children with pediatric diabetes, cancer, cystic fibrosis, and sickle cell disease. **Methods** A search was conducted for qualifying studies published since 1980. Only studies meeting basic criteria for external and internal validity were included. Nineteen studies were identified, providing data on 62 outcome variables. Effect sizes (ESs) were analyzed by illness type, intervention type, and strength of internal and external validity of the research design. **Results** Overall, interventions were associated with large ESs, which were not significantly moderated by illness type or intervention type. However, larger ESs were associated with lower scores on validity of research design. **Conclusions** Adjuvant psychological interventions for pediatric chronic illnesses appear in general to be efficacious, associated with a large mean ES across a range of outcome variables. However, until more studies have been completed using stronger research designs, only tentative conclusions can be drawn.

Key words effect size; intervention; pediatric chronic illness; validity.

It is now widely recognized that when children have a chronic physical illness, such as diabetes or cancer, the developmental course of that illness is jointly influenced by physical, psychological, and social factors. Understanding of the interdependencies between these three sets of factors has been a major research focus in recent decades (Benedito-Monleon & Lopez-Andreu, 1994; Brown et al., 1993; Buckley, Vacek, & Cooper, 1990; Cohen, 1971; Creed, 1990; Gartstein, Short, Vannatta, & Noll, 1999; King & Hanson, 1986; Pendley et al., 2002). One important aim of that research has been the identification and development of psychological interventions that are efficacious in improving psychological and physical states associated with the illness (Anie & Green, 2002; Beale, Bradlyn, & Kato, 2003; Billings, Moos, Miller, & Gottlieb, 1987; Buckelew & Parker, 1989; Ceccoli, 1992; McQuaid & Nassau, 1999).

Although there is now a wide variety of psychological interventions described in the research literature on pediatric chronic illness, it is unclear just how useful particular interventions are, which interventions are better than others, and whom are they more efficacious for (McQuaid & Nassau, 1999). The main problem is that different studies of a particular type of intervention invariably differ in the specifics of the intervention used, how it is implemented, the research design, the instruments used to measure change, when and how measurements were taken, and how results were analyzed and interpreted. For example, consider how an intervention is described and implemented. Unless an intervention is manualized or automated (and possibly even if it is), there may be wide variation between studies in the actual detail of a particular type of intervention, even though some seemingly minor differences may be critical to the intervention outcome. For example, how the intervention is applied, and for how long it is applied, may well affect its efficacy. Additionally, where delivery of an intervention involves "live" therapists interacting extensively with child participants, differences in the quality of therapist–child relationships may have a substantial effect on the so-called "nonspecific" intervention effects that typically are a significant component of

All correspondence concerning this article should be addressed to Ivan L. Beale, School of Psychology, University of New South Wales, Sydney 2052, Australia. E-mail: ibeale@psy.unsw.edu.au

Journal of Pediatric Psychology 31(5) pp. 437–451, 2006
doi:10.1093/jpepsy/jsj079
Advance Access publication September 14, 2005

change in outcome measures (Horvath, 1988; Shepherd & Sartorius, 1989). The magnitude of these effects may vary substantially between studies in a way that makes comparisons between the results of different studies somewhat hazardous (Chambless & Ollendick, 2001; Rosenthal & DiMatteo, 2001).

Over recent decades, systematic procedures have been developed to facilitate valid comparisons between different studies of particular interventions. For example, as part of an initiative to facilitate the use of "evidence-based" interventions by clinical psychologists, the American Psychological Association commissioned task groups to analyze the intervention literature to identify empirically supported therapies. This resulted in the development of the Chambless criteria by which specific types of intervention are designated "well established," "probably efficacious," or "experimental." A related set of criteria subsequently developed for interventions with children included an additional category, "promising interventions" (Chambless & Hollon, 1998; McQuaid & Nassau, 1999), a state of affairs reflecting the relatively small number of pediatric intervention studies. However, the use of these criteria as a sole basis for making judgments about the evidential status of an intervention has been criticized as insufficiently reflecting the dearth of evidence that research findings can be applied to real clinical settings and to ordinary clients, many of whom would not meet the strict inclusionary criteria typically used in research studies (Chambless & Ollendick, 2001).

The Chambless criteria provide a basis for identifying those studies that do not meet minimal standards to qualify as evidence that an intervention could be considered "promising." A basic requirement is that the study design includes a condition that can control for nonspecific effects associated with the intervention process. Studies with designs that do not include a basic control condition can be excluded from further consideration because they muddy the waters. The Chambless criteria can be applied to the remaining studies to determine whether a type of intervention can be considered an empirically supported treatment (EST) and whether should be regarded as being experimental, promising, probably efficacious, or well established.

This idea of classifying interventions, using standard criteria for determining the level of evidence that an intervention is empirically supported as being efficacious, really addresses just one of the many important questions. Additionally, it is useful to know the expected effect sizes (ESs) for relevant outcome variables and whether moderating variables have been identified. Also, it can be valuable to know the clinical significance of expected ESs, where this can be extracted from the outcome data. Even ESs that are classified as "large" by standard criteria may not necessarily indicate that an intervention has resulted in a socially or clinically important change in the child (Jacobson & Truax, 1991).

A second approach for comparing and combining the results of studies of intervention efficacy is provided by meta-analysis. Meta-analysis allows the combining of numerical results from several studies, the accurate estimate of the magnitude of changes in outcome measures, and the explanation of inconsistencies in the findings of different studies. Meta-analysis may in some instances also be used to identify variables that moderate the effects of the intervention. According to its advocates, meta-analysis allows conclusions about interventions that are more accurate and more credible than can be drawn from any single study or from a nonquantitative, narrative review (Rosenthal & DiMatteo, 2001).

Meta-analysis has been criticized as not taking adequate account of the methodological quality of studies, mixing the good with the bad (Hunt, 1997). However, it is possible to "rate" the methodology of studies and then use the rating scores as a variable in the meta-analysis. This potentially can reveal whether the quality of the methodology has a moderating effect on the overall size of the intervention effect (Rosenthal & DiMatteo, 2001). Meta-analysis has also been criticized as combining results from studies that vary importantly in aspects of procedure and measurement, even though their aims are similar. Thus, it is seen as "comparing apples and oranges" (Hunt, 1997). As with the quality issue, this combining of different studies may be an advantage, if there are sufficient studies to allow methodological or measurement differences to be extracted as a moderating factor in the meta-analysis (Rosenthal & DiMatteo, 2001).

A limitation of meta-analysis is that it works best if there are plenty of studies of an intervention that are similar enough that they can be analyzed meaningfully as a group and if results are reported in a form where it is possible to extract an appropriate ES statistic (Rosenthal & Rosnow, 1991). In the case of psychological interventions for pediatric chronic illness, the number of studies of specific interventions for particular illnesses generally is small, but the conceptual similarities across interventions and illnesses in intervention methods and psychological constructs targeted makes it feasible to conduct analyses across illnesses and intervention types.

This article reviews empirical evidence of the efficacy of many psychological interventions that have been evaluated with pediatric diabetes, cancer, cystic fibrosis, and sickle cell disease. The first objective is to obtain an

estimate of the size of the effect associated with psychological interventions in general. A second objective, depending on the nature and number of qualifying studies, is to estimate the ES associated with different types of intervention, when applied to different illnesses. An attempt will be made to identify moderating variables, both independent and dependent, especially research design variables that impact on internal and external validity. Although both the theoretical antecedents of interventions (Zeltzer, 1999), and the basis of the methods of implementation (Clark & Valerio, 2003), may well be predictors of intervention efficacy, their possible influence is not analyzed in this review.

Method

Search Strategy

Relevant databases were searched for published empirical studies that evaluated psychosocial interventions for pediatric populations with chronic illnesses. Databases searched were PsycInfo, PsycScan, PubMed, and Cochrane Systematic Reviews. Searches were conducted in April, 2004, for the period 1980–2004 using combinations of the keywords: pediatric; treatment; intervention; psychosocial; psychological; psychoeducation; education; chronic illness; cancer; diabetes; sickle cell; cystic fibrosis. Additional studies were identified from the reference lists of the studies located in the initial search.

Review Process

Studies were included only if they met the following criteria: the sample studied includes adolescents or children with a chronic illness other than asthma; the research design includes a control condition that can be contrasted with the intervention being evaluated; the intervention being evaluated includes a psychological (psychoeducational or psychosocial) component; a psychological/educational component of the intervention is applied directly to the children or adolescents in the study (interventions directed only at parents; community resources or health providers are not included). Not all chronic illnesses were included, simply to keep the analysis and report to a manageable size. Asthma was excluded because a meta-analysis of asthma interventions had recently been published (Wolf, Guevara, Grum, Clark, & Cates, 2003). Interventions solely for pain management were also excluded, because they were considered sufficiently numerous to require a separate study.

Studies were analyzed to extract descriptive information in the following categories: chronic illness type (cancer, diabetes, etc.); primary research population [age range(s)]; country where the study was conducted; nature of intervention(s) being used (information video, CBT, hypnosis, etc.); numbers of participants included in study, and number in each group; gender distribution between groups; primary variables targeted in intervention; type of control condition (waitlist, attention, etc.); random or nonrandom allocation of participants to conditions; calculation of statistical power of design (reported or not); who was blind to participant allocation (not stated, researchers, participants, etc.); duration of intervention; duration of follow-up period included; details of participant attrition; measures of acceptability or credibility included; information on validity and reliability of measures; measures of treatment validity/integrity included; whether scorers of measures were blind to condition; whether relevant covariates were addressed in analyses; whether intent-to-treat analysis was included; ESs for variables showing significant effects of treatment.

Assessment of Quality of Research Design

The information tabulated for each study was used to derive two scores indexing the extent to which the study controlled for threats to (a) internal validity and (b) external validity (Cook & Campbell, 1979). A score of 1–3 (1, "poor or not reported"; 2, "adequate"; and 3, "good") was assigned for each feature of the research design that influenced either the internal or external validity of the study, as follows:

Internal validity: demonstrated reliability of measures; scorers of measures blind to experimental condition; participants blind to expected study outcome; random allocation to conditions (where relevant); specificity of control condition; control or measurement of treatment integrity. External validity: demonstrated validity of measures; measured acceptability/credibility of treatment; participant attrition recorded; intent-to-treat analysis provided; follow-up included; participant characteristics clearly defined; treatment duration measured; sufficient statistical power to detect large effect; range of dependent variables measured; relevant covariates controlled in analysis.

Scores (1–3) on each feature were summed to provide separate total scores for internal and external validity, as well as an overall validity score. Scoring of each study on these validity criteria was performed independently by two PhD-level psychologists with experience in clinical research. Inter-rater reliability (% agreement) was calculated separately for internal and external validity total scores. Inter-rater agreement scores for each study all exceeded 85%. Ratings by the principal investigator only are included in this report.

Calculation of ES

Because few of the studies reviewed reported estimates of ES, estimates were calculated from the reported data wherever possible, using standard procedures for between-groups, repeated measures, and mixed designs (Hedges & Olkin, 1985; Lipsey & Wilson, 2001) or small-N within-subjects designs (White, Rusch, Kazdin, & Hartmann, 1989). Although alternative methods have been proposed for deriving ESs from single-case designs (Carr, 2002), the preferred method allows direct comparison between different designs using the same statistic. In most cases, it was possible to calculate ESs directly as the standardized mean difference between scores of pre- and postintervention (Cohen's *d*). However, a few studies reported ES estimates as indices (*f*, *w*) that cannot be directly converted to values of *d*. In these cases, the reported value was converted into a value of *d* representing an effect of equivalent magnitude (small, medium, large), using published ES tables (Rosenthal & Rosnow, 1991). This process allowed comparison to be made across all studies by a single ES index (*d*). Where studies included multiple measures of a single construct, the mean ES for the group of measures was extracted as the appropriate index for the current meta-analysis, following the principle of one ES per concept (Lipsey & Wilson, 2001). Also, ESs were based only on the final set of measurements taken for each outcome measure, better to reflect the generalization of intervention effects over time. For example, in studies in which there were several times at which measurements were taken, either during or following intervention, ESs were based on the last follow-up measurement undertaken.

Because most studies did not report data necessary for calculation of standard error of pre–post ESs, it was not possible to calculate either the standard error or the inverse variance weight of the ES. This resulted in some limitations on the analyses that could legitimately be performed on the samples of ESs obtained in this study. For example, results of procedures, such as *t* tests and analysis of variance (ANOVA), must be interpreted with more than of the usual degree caution, because of increased probability of systematic variance within groups of ES values and the consequent likelihood of inflated type-II error (Hedges & Olkin, 1985).

Most intervention studies yield ESs for comparisons between groups (e.g., intervention vs. control) as well as for comparisons within groups (pre vs. post), and a decision was made in this review to use only ESs for pre–post comparisons. The effects represented by these ESs necessarily include both the specific effects of the intervention and the nonspecific effects shared by the intervention and control conditions, both of which typically contribute to outcomes of interventions applied in clinical settings. Therefore pre–post ESs are arguably the best predictors of the effectiveness of interventions, especially when used in a normal clinical context.

Analyses Across the Set of Studies

Analyses were conducted across the set of studies to provide summary statistics of ESs, validity scores, and associations between these measures. ESs associated with control conditions were not included in these analyses, the purpose being to examine the changes in dependent measures associated with the specific intervention being evaluated. In addition, simple descriptive and statistical analyses were conducted on the samples of ESs associated with study characteristics, different illness categories (cancer, diabetes, etc.), and different outcome measure categories (symptoms, psychological adjustment, self-care, etc.).

Results

Nineteen studies were located which met the inclusion criteria, representing evaluations of interventions for sickle cell disease, cancer, diabetes (IDDM), and cystic fibrosis. For most studies, it was possible to extract ESs as standardized mean differences (pre–post) directly, though for two studies it was necessary to estimate the standardized mean difference from the reported values of partial eta-squared (*f*) or the chi-square-based index (*w*). Only one study reported results in a form where it was not possible to extract pre–post ES estimates. Because many studies reported results for several conceptually different outcome measures, a total sample of 94 ES values was obtained.

Summary data for each study, separated by illness type, are summarized in Tables I–III. In these tables, as well as in subsequent analyses, studies of diabetes and cystic fibrosis are grouped together in a single category reflecting the small number of cystic fibrosis studies and a joint emphasis on dietary control in the psychological interventions for these illnesses.

The right-most column in Tables I–III summarizes the category (1–5) designated for each outcome variable. This categorization was created to reduce the number of outcome variable types into five general categories that were conceptually distinct and contained samples large enough for differential analysis. They are (a) symptoms and physiological variables; (b) illness knowledge; (c) psychological adjustment; (d) self-care and coping behaviors; and (e) attitudinal variables.

Table I. Characteristics of Individual Studies of Interventions for Pediatric Sickle Cell Disease

Study	IV/18	EV/30	EST category	Intervention type	Participants	*N*	Outcome measures	ES (*d*)	Outcome category
Boroffice (1991)	11	16	3	1. Lecture education	Adolescents and young adults (13–22 years) (Nigeria)	60 (20, 20, 20) Gender numbers not reported	1. Attitude to services T1	0.48	5
				2. (1) + group discussion			T2	1.55	
				3. Standard care			2. Attitude to medication T1	1.6	5
							T2	1.51	
Thomas, Dixon and Milligan (1999)	13	21	3	1. CBT, relaxation, education	Adolescents and young adults (15–35 years) (UK) (worst-case sample)	97 started; 59 completers (19, 14, 26). Gender numbers not reported	1. Mental health	0.3	3
				2. Attention control			2. Coping	0.5	4
				3. Standard care			3. Pain S-E	0.7	1
							4. Pain severity	0.3	1
							5. Pain beliefs	0.5	1
Broome, Maikler, Kelber, Bailey, and Lea (2001)	12	24	3	Education classes and either CBT for pain (relaxation, distraction, and imagery), art therapy, or attention control	Children (75) and adolescents (46) (US)	121 started, 97 completers (F = 53, M = 44)	Number of coping strategies rated effective T1	0.47	4
							T2	1.1	4
							C	0.7	4
Kaslow et al. (2000)	13	22	3	Family psychoeducation program: coping with stress/pain; relaxation; imagery	7- to 16-year-olds (US)	47 [39 completers (20, 19)]. Gender numbers not reported	SCD knowledge (child) T	1.79	2

continued

Table I. *continued*

Study	IV/18	EV/30	EST category	Intervention type	Participants	*N*	Outcome measures	ES (*d*)	Outcome category
							C	0.06	2
							SCD knowledge (parent) T	1.64	2
							C	0.27	2
							Depression T	0.44	3
							C	0.37	3
							Internalizing problems T	0.37	3
							C	0.35	3
							Externalizing problems T	0.24	3
							C	0.44	3
Gil et al. (2001)	11	20	3	1. Child-coping practice: relaxation; imagery; self-talk	Children and adolescents (8–17 years) (US)	46 (26, 20) (F = 24, M = 22)	Self-reported coping	1.01	4
				2. Standard care					
Hazzard, Celano, Collins, and Markov (2002)	12	20	4	1. Starbright program	Children and adolescents (8–18 years) (US)	47 (18, 29). Gender numbers not reported	Knowledge T	0.49	2
				2. Standard education program			C	0.35	2
							Perceived social support T	0.16	4
							C	0.12	4
							Positive coping (ns) T	0	4
							(ns) C	0.09	4
							Negative coping T	1.1	4
							C	–0.18	4

C, control condition; ES (*d*), effect size (*d*); EST, category as empirically supported treatment (1, "well established"; 2, "probably efficacious"; 3, "promising"; 4, "experimental"); EV, external validity; *f*, estimated value of *d* equivalent to authors' ES (*f*); IV, internal validity; SCD, sickle cell disease; T, treatment condition.

Table II. Characteristics of Individual Studies of Interventions for Pediatric Cancer

Study	IV/18	EV/30	EST category	Intervention type	Participants	*N*	Outcome measures	ES (*d*)	Outcome category
Dragone, Bush, Jones, Bearison, and Kamani (2002)	13	19	4	1. Educational; interactive CD-ROM	Children with leukemia (US)	31 (14 4–6 years) (17 7–11 years). Group numbers not reported	Health locus of control T1	0.52	5
				2. Book (you and leukemia)			T2	–0.28	5
							Knowledge (pre–post; group effect)	0.29	2
Zeltzer, Dolgin, LeBaron, and LeBaron (1991)	12	22	3	1. Hypnosis/imagery	Children and adolescents (5–17 years) (US)	54 (17, 21, 16) (M = 26, F = 28)	Somatic score T1	0.64	1
			4	2. Support group—cognitive (distraction/relaxation)			T2	0.28	1
				3. Attention control			C	–0.89	1
							Functional score T1	0.5	4
							T2	–0.05	4
							C	–0.51	4
							Symptom score T1	0.71	1
							T2	0.28	1
							C	–0.69	1
							Anticipating vomiting and nausea T1	0.22	1
							T2	–0.06	1
							C	–0.18	1
Hinds et al. (2000)	12	23	3	1. Education on self-care and coping; rehearsal	Adolescents (12–21 years) (US)	78 (40, 38) (M = 42, F = 36)	Hopeful/hopelessness T	0.25	5
				2. Attention control			C	0.62	5
							Health locus of control T	0.23	5
							C	0.27	5
							Self-esteem T	0.17	5
							C	0.31	5

continued

Table II. *continued*

Study	IV/18	EV/30	EST category	Intervention type	Participants	*N*	Outcome measures	ES (*d*)	Outcome category
							Symptom distress T	0.3	1
							C	0.59	1
							Self-efficacy T	0.1	5
							C	0.36	5
							Toxicity T	0.42	1
							C	0.87	1
Favara-Scacco, Smirne, Schiliro, and Di Cataldo (2001)	7	14	4	1. Art therapy	Children (2–14 years) (Sicily)	32. Group numbers not reported	Proportion "good responders" (treatment vs. control)	0.54	4
				2. Standard care					
Kolko and Rickard-Figueroa (1985)	10	22	2	1. Video game distraction	Children and adolescents (11–17 years) (US)	3 (M = 3)	Anticipatory symptoms (treatment vs. control)	3.23	1
				2. Standard care					
							Behavioral distress (treatment vs. control)	2.76	1
Varni, Katz, Colegrove, and Dolgin (1993)	12	26	3	1. Social skills training	Children and adolescents (5–13 years) (US)	64 (33, 31) (M = 38, F = 36)	State anxiety T	0.79	3
				2. Standard school reintegration program			C	0.3	3
							Teacher social support T	0.47	4
							C	–0.23	4
							Behavior problems T	0.56	3
							C	0.03	3
							Internalizing problems T	0.32	3
							C	0.05	3
							Externalizing problems T	0.55	3
							C	0.05	3

C, control condition; ES (*d*), effect size (*d*); EST, category as empirically supported treatment (1, "well established"; 2, "probably efficacious"; 3, "promising"; 4, "experimental"); EV, external validity; IV, internal validity; T, treatment condition.

Table III. Characteristics of Individual Studies of Interventions for Pediatric Diabetes (IDDM) and Cystic Fibrosis (CF)

Study	IV/18	EV/30	EST category	Intervention type	Participants	*N*	Outcome measures	ES (*d*)	Outcome category
Brown et al. (1997) (IDDM)	10	18	3	1. Videogame "Packy & Marlon"—diabetes content	Children and adolescents	59 (31, 28). Gender numbers not reported	Self-efficacy (self-care) T	0.66	5
				2. Attention control game	8–16 years (US)		C	0.21	5
							Social support T	0.58	4
							C	–0.17	4
							Knowledge T	0.31	2
							C	0.13	2
							Self-care (parent report) T	0.29	4
							C	–0.42	4
							Urgent doctor visits T	0.46	1
							C	–0.04	1
Kaplan, Chadwick, and Schimmel (1985) (IDDM)	11	19	4	1. Social learning	Adolescents (13–18 years) (US)	21 (?, ?) (M = 8, F = 13)	HbA levels (ns)	No ES available	
				2. Information program			Knowledge (ns)		
							Behavior (ns)		
							Attitudes		
Boardway, Delamater, Tomakowsky, and Gutai(1993) (IDDM)	12	20	3	1. CBT stress management	Adolescents (12–17 years) (US)	19 (9, 10) (M = 9, F = 10)	HbA levels (ns)		
							Stress rating T	0.91	3
				2. Standard care			Coping responses (ns)		
							Self-efficacy (ns)		
Moore, Geffken, and Royal (1995) (IDDM)	9	21	3	1. CBT fear reduction	Adolescents (11–13 years) (US)	2. Gender numbers not reported	Time to complete injection (pre–post comparison)	1.79	4

continued

Table III. *continued*

Study	IV/18	EV/30	EST category	Intervention type	Participants	*N*	Outcome measures	ES (*d*)	Outcome category
				2. Observation only					
							Proportion trials with distress (pre–post comparison)	0.52	1
Rose, Firestone, Heick, and Faught (1983) (IDDM)	10	22	3	1. Anxiety management	Adolescents (15–18years) (Canada)	5 (F = 5)	Urine glucose % (baseline/treatment comparison)	1.42	1
				2. Attention control					
				3. Standard care					
Stark et al. (1996) (CF)	11	21	3	1. Parent/child education and behavior management training groups	Children (5–10 years) (US)	9 (5, 4). Gender numbers not reported	Total calories T	1.31	1
							C	0.66	1
				2. Waitlist			Standardized weight		
							T	0.62	1
							C	–0.07	1
							Standardized height		
							T	–0.14	1
							C	0.14	1
Stark et al. (1993) (CF)	9	22	3	1. Parent/child education and behavior management training groups	Children (3–8 years) (US)	3 (M = 1, F = 2)	Total calories (pre–post)	2.66	1
				2. Observation only			Proportion expected weight gain (pre–post)	1.63	1
							Proportion expected height gain (pre–post)	0.17	1

C, control condition; ES (*d*), effect size (*d*); EST, category as empirically supported treatment (1, "well established"; 2, "probably efficacious"; 3, "promising"; 4, "experimental"); EV, external validity; IV, internal validity; T, treatment condition.

Characteristics of Studies

Most of the qualifying studies were based in the United States, although many other countries were represented, including Canada, United Kingdom, Sicily, and Nigeria. Most interventions were psychoeducational in character, usually involving combinations of information- and skill-training modules. Most skill training embraced self-management or coping skills directed to a variety of issues, including distraction from pain or anxiety or social interactions. Specific therapeutic techniques included cognitive behavior therapy, biofeedback training, hypnosis, and interactive computer games or educational tutorials specific to the illness concerned. Parents were often included in training sessions either as observers or as participants. There was marked variability both in the content and procedural aspects of interventions, to the extent that it was difficult to come up with a few terms, such as "educational" or "CBT" or "biofeedback" that would really provide a valid characterization of many of the interventions for the present purposes. The descriptions of the interventions in the publications themselves were extremely varied with respect to details both of content and procedure. Because many of the studies did not report using measures of treatment integrity or other aspects of quality assurance, it is difficult for the reviewer to know much about how interventions really were implemented. This is problematic in as much as a given type of intervention might be implemented well or badly, potentially with very different results for the recipients.

Validity of Research Studies

It is notable that as a whole, the group of studies surveyed did not meet high standards of internal validity. For example, there appeared typically to be no consideration of the differential demand characteristics, between groups, associated with researcher or therapist expectations of change in participants. Given that, with psychological interventions, it usually is not possible to arrange that researchers or participants are blind to the intervention a participant is receiving, there is a potential risk of reporting or scoring bias that might mask a real intervention effect (Rosenthal & Rosnow, 1991). Such bias might be minimized by the use of procedures, such as blind scoring, but none of the studies reported using such safeguards. Other unaddressed validity threats common to most of the studies were associated with failure to compute the statistic power of the research design (and, therefore, the balance of probabilities between type-I and type-II errors), and failure to report an intent-to-treat analysis that would have indicated how the overall result was affected by participant attrition. In fact, some studies reported substantial attrition rates.

Scores on criteria for internal and external validity, as displayed in Tables I–III, varied over a range from 7 to 13 ($M = 11.05$, $SD = 1.58$) for internal validity and 14–26 ($M = 20.72$, $SD = 2.78$) for external validity. Note that all included studies also met a basic internal validity criterion of having a design that contrasted at least two different experimental conditions, usually a treatment condition and a control condition. For this reason, the internal validity score does not represent a complete measure of a study's overall internal validity and is intended only to provide a basis for comparing studies on the internal validity dimension. Scores on internal or external validity were not significantly different either across the types of illness to which interventions were applied, internal validity: $F(2, 11) = 0.925$, $p = .42$; external validity: $F(2, 11) = 0.49$, $p = .62$, or the type of intervention used, internal validity: $F(4, 13) = 1.31$, $p = .32$; external validity: $F(4, 13) = 1.67$, $p = .22$.

ES

The mean ES across all active intervention conditions was 0.71 (range = 0.28–3.23, $SD = 0.61$). The mean ES for control conditions was 0.12 (range = 0.82–0.87, $SD = 0.39$). All subsequent references to ES in this report refer to ES for treatment conditions only, unless stated otherwise. There was a significant negative correlation between ES and internal validity score ($N = 62$, $r = -.37$, $p = .003$), but no significant correlation between ES and external validity score ($N = 62$, $r = -.06$, $p = .65$). Examination of the variance in scores on the individual items in the internal validity scale indicated that negative correlation with ES was influenced primarily by scores on items 1, 4, and 6, that is, reliability of measures, randomness of allocation of participants to conditions, and control of treatment integrity.

Exploratory analysis of associations between ES and illness type showed no evidence that larger ESs might occur with some illness types. The mean ES for each illness category was sickle cell disease, 0.77 ($SD = 0.56$); cancer, 0.56 ($SD = 0.77$); diabetes and cystic fibrosis, 0.88 ($SD = 0.75$). Also, there was no significant difference between the ESs based on outcome measurement obtained either directly following intervention ($M = 0.70$, $SD = 0.73$) or after a follow-up period ($M = 0.67$, $SD = 0.60$), $t(57) = 0.16$, $p = .48$. Analysis of ES by type of intervention was considered problematic, because of the wide diversity of intervention methods and content, together with the range of illnesses and issues addressed and the small number of qualifying

studies in any particular category. However, for what it is worth, the information provided in Tables I–III does allow easy eyeballing of the ESs associated with particular interventions used in individual studies, together with validity ratings for the study designs. As noted previously, it may well be that the detail about how well interventions were implemented, especially scores on treatment integrity, would be more likely to be associated with ES than would a summary descriptor of the intervention type, such as "CBT" or "biofeedback." Analysis of ES by the duration of intervention was considered impractical, because some studies did not report the duration and/or number of intervention sessions, and many interventions included homework requirements that involved things like skill rehearsal for undefined durations. Duration of interventions delivered by computer, such as health games, was often reported of access time rather than actual use. Nevertheless, it is notable that intervention durations varied between as little as 20-min computer access during one session and as much as 10 or more 1-h sessions of therapist-conducted coping skills training spread over several months.

Tables I–III summarize estimates of the EST category for each intervention, based on an overall evaluation using the modified Chambless criteria outlined earlier in this article. None of the interventions were considered to have met the criteria for category 1, and few met even category 2, because one-off evaluations were the rule, and independent supporting evidence was rare. Most studies fell into category 3, interventions showing promise.

Discussion

The main objective of this review was to obtain a broad perspective on the efficacy of psychological interventions across several of the more-prevalent types of chronic pediatric illness. Although there are many psychological issues common to different pediatric chronic illnesses, there are no previous meta-analyses of relevant intervention studies across different illness types. It was thought that a broad analysis might reveal possible differences between illness types either in the efficacy of interventions or in the characteristics of the intervention studies.

A main finding was that, in general, few intervention studies meeting basic validity criteria were located for most of the illness types. There were too few qualifying studies of any single illness type to provide a valid database for standard meta-analysis. This finding supports this strategy of combining illness types for meta-analysis. Using widely adopted criteria for interpreting ESs (Cohen, 1988), the obtained mean ES (0.71) across illness types and outcome measures is interpreted as indicating an overall large intervention effect. In other words, a typical intervention of the type reviewed is associated with an improvement on outcome measures for about 80% of participants. The statistic analyses conducted did not show differences between illness types on ES, although, as noted previously, the statistical tests used theoretically were biased toward type-II error. These findings support the view that psychological interventions exert a substantial positive influence on a range of outcomes when used as adjuvant therapies to medical treatments in the pediatric illnesses reviewed.

The overall mean ES obtained in this review (0.71) is considerably larger than the overall mean ES (0.30) previously reported in a meta-analysis of interventions with pediatric asthma (Wolf et al., 2003). However, the ESs reported in the review of asthma interventions are based on the difference between postintervention means for the treatment group versus the control group, whereas the ESs in this review are based on difference between pre- and postintervention means in the treatment group. As previously noted, the pre–post ES was chosen, because it better reflects the total change associated with treatment, including nonspecific effects shared with the control condition. If the nonspecific effects of treatment are estimated by the pooled mean ES for the control conditions across all the studies in Tables I–III ($M = 0.18$), it would appear that nonspecific effects accounted for less than one-third of the pooled mean pre–post ES (0.71). This estimate is within the range of proportions of nonspecific effects reported for various psychological interventions (Horvath, 1988).

A previous meta-analysis of psychological interventions for pediatric illnesses (Kibby, Tyc, & Mulhern, 1998) found an overall mean ES of 0.87 for treatment versus control comparisons and 1.4 for pre–post comparisons, this last figure being somewhat larger than was found in this analysis. It is notable, though, that the two types of ES were found not to be statistically different. Both are classified "large," as is the pooled ES reported in this analysis.

It has been noted elsewhere that ESs are not a sufficient basis for estimating whether or not the effects of a treatment are "clinically significant" as opposed to merely efficacious (Kazdin & Weisz, 1998). Unfortunately, scores on outcome variables that could be used to estimate clinical significance of pre–post changes, such as *T* scores (Jacobson & Truax, 1991), were not available from most of the studies reviewed. With regard to the Chambless criteria for ESTs, these observations and interpretations concerning the ESs associated with this set of psychological treatments might be taken to indicate that the treatments

were in general very efficacious in producing desirable changes in a range of outcome variables. However, this conclusion must be tempered by the evidence that larger ESs tended to be associated with studies that scored lower on internal validity. As noted previously, lower internal validity scores tended to be associated with designs in which reliability of outcome measures was unknown, allocation of participants to conditions was not random, and there was no reported control of treatment integrity. Under these conditions, especially without the precautionary measure of blind scoring, both participants' and researchers' expectations of the treatment under study could create a bias toward obtaining spuriously large ESs. This tendency for weaker research designs to be associated with larger ESs has been noted previously (Garrett, 1985). At present, most of the interventions reviewed should best be regarded either as promising or probably efficacious.

An additional objective of this review was to attempt to identify intervention approaches that tended to be associated with larger ESs. This proved not to be achievable both because there could be found no workable method for reducing the large range of interventions into just a few conceptually valid categories suitable for analysis, and because the total number of qualifying studies was insufficient for this type of analysis. In particular, the notion of creating a few general categories for intervention approaches proved to be inconsistent with the multicomponent nature of most of the interventions studied. Moreover, relevant information about how well interventions were implemented, for example, treatment integrity scores, whereas clearly important for the purposes of this review, was generally lacking in the published reports reviewed. As in this analysis, Kibby et al. (1998) found no significant effect on ES of either the type of intervention or the type of outcome measure.

Other reviewers of the research literature on efficacy of psychological interventions for pediatric chronic illnesses have commented on the lack of good studies, and in particular, the lack of evidence regarding what sort of intervention is suitable for whom (La Greca & Varni, 1993; Zeltzer, 1999). Although the children and adolescents participating in the studies reviewed here presumably differed from one another on relevant variables, such as gender, developmental level, cognitive skills, or illness perceptions, there were no analyses reported of the interactions of such variables with the intervention process, nor of the impact of these variables on intervention outcome measures. If this review of ESs is taken as evidence that psychological interventions in general possibly are efficacious, it would make sense for future studies to consider examining more closely the issue of how interventions should best be matched to the characteristics of the potential recipient.

The findings of this review support some general observations about research on the efficacy of psychological interventions for pediatric chronic illnesses. Based on an overview of the published research on several chronic illnesses, using designs that meet basic validity criteria, there is evidence that, in general, psychological interventions possibly are efficacious. For the illnesses addressed in this review, there are in general too few studies to indicate whether particular treatment approaches are reliably efficacious for particular illnesses and outcome measures. Independent replications will be required to establish whether any particular intervention can be considered well established as being efficacious for treating a given aspect of illness. Based on the published reports, many studies could benefit from better research designs that provide better control for known threats to both internal and external validity. Particular issues are control or monitoring of treatment integrity and control of bias owing to participant and researcher expectations. It is desirable that studies evaluating particular interventions should include measures of treatment integrity, even if the intervention is manualized. Otherwise, it is impossible to be certain exactly what the intervention was that is being evaluated. Future evaluation research needs to be mindful of the need to use designs that will ensure that findings are not unduly biased by researchers' expectations, for example, by using blind scoring of outcome measures. Improvement in the quality of research designs will do much to lift confidence in the emerging belief in the efficacy of psychological interventions for pediatric chronic illnesses.

Acknowledgments

Ivan L. Beale was formerly at HopeLab, Palo Alto, where this research was initiated and supported.

Received February 28, 2005; revision received June 20, 2005; accepted August 23, 2005

References

Anie, K. A., & Green, J. (2002). Psychological therapies for sickle cell disease and pain. *Cochrane Database Systematic Reviews, 2*, CD001916.

Beale, I. L., Bradlyn, A. S., & Kato, P. M. (2003). Psychoeducational interventions with pediatric cancer patients: Part II. Effects of information and skills training on health-related outcomes. *Journal of Child and Family Studies, 12*, 385–397.

Benedito-Monleon, C., & Lopez-Andreu, J. A. (1994). Psychological factors in childhood asthma. *Behavioural and Cognitive Psychotherapy, 22*, 153–161.

Billings, A. G., Moos, R. H., Miller, J. J., 3rd, & Gottlieb, J. E. (1987). Psychosocial adaptation in juvenile rheumatic disease: A controlled evaluation. *Health Psychology, 6*, 343–359.

Boardway, R. H., Delamater, A. M., Tomakowsky, J., & Gutai, J. P. (1993). Stress management training for adolescents with diabetes. *Journal of Pediatric Psychology, 18*, 29–45.

Boroffice, O. B. (1991). Self-help in sickle-cell disease in Nigeria. *Health Promotion International, 6*, 241–246.

Broome, M. E., Maikler, V., Kelber, S., Bailey, P., & Lea, G. (2001). An intervention to increase coping and reduce health care utilization for school-age children and adolescents with sickle cell disease. *Journal of the National Black Nurses Association, 12*, 6–14.

Brown, R. T., Kaslow, N. J., Doepke, K., Buchanan, I., Eckman, J., Baldwin, K., et al. (1993). Psychosocial and family functioning in children with sickle cell syndrome and their mothers. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*, 545–553.

Brown, S. J., Lieberman, D. A., Gemeny, B. A., Fan, Y. C., Wilson, D. M., & Pasta, D. J. (1997). Educational video game for juvenile diabetes: Results of a controlled trial. *Medical Informatics, 22*, 77–89.

Buckelew, S. P., & Parker, J. C. (1989). Coping with arthritis pain. A review of the literature. *Arthritis Care Research, 2*, 136–145.

Buckley, L. M., Vacek, P., & Cooper, S. M. (1990). Educational and psychosocial needs of patients with chronic disease. A survey of preferences of patients with rheumatoid arthritis. *Arthritis Care Research, 3*, 5–10.

Carr, A. (2002). Introduction. In A. Carr (Ed.), *Prevention: What works with children and adolescents* (pp. 1–16). Houl, East Sussex, UK: Brunner-Routledge.

Ceccoli, V. C. (1992). Developing group treatment for a pediatric diabetic population: The Bellevue experiment. *Journal of Child and Adolescent Group Therapy, 2*, 67–76.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*, 7–18.

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

Clark, N. M., & Valerio, M. A. (2003). The role of behavioural theories in educational interventions for paediatric asthma. *Paediatric Respiratory Reviews, 4*, 325–333.

Cohen, P. A. (1971). Psychosocial problems in treatment of cystic fibrosis. *Pediatrics, 47*, 472.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, T. D., & Campbell, D. T. (Eds.) (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

Creed, F. (1990). Psychological disorders in rheumatoid arthritis: A growing consensus? *Annals of Rheumatic Diseases, 49*, 808–812.

Dragone, M. A., Bush, P. J., Jones, J. K., Bearison, D. J., & Kamani, S. (2002). Development and evaluation of an interactive CD-ROM for children with leukemia and their families. *Patient Education and Counselling, 46*, 297–307.

Favara-Scacco, C., Smirne, G., Schiliro, G., & Di Cataldo, A. (2001). Art therapy as support for children with leukemia during painful procedures. *Medical Pediatric Oncology, 36*, 474–480.

Garrett, C. J. (1985). Effects of residential treatment on adjudicated delinquents: A meta-analysis. *Journal of Research in Crime and Delinquency, 45*, 287–308.

Gartstein, M. A., Short, A. D., Vannatta, K., & Noll, R. B. (1999). Psychosocial adjustment of children with chronic illness: An evaluation of three models. *Journal of Developmental and Behavioural Pediatrics, 20*, 157–163.

Gil, K. M., Anthony, K. K., Carson, J. W., Redding-Lallinger, R., Daeschner, C. W., & Ware, R. E. (2001). Daily coping practice predicts treatment effects in children with sickle cell disease. *Journal of Pediatric Psychology, 26*, 163–173.

Hazzard, A., Celano, M., Collins, M., & Markov, Y. (2002). Effects of STARBRIGHT World on knowledge, social support, and coping in hospitalized children with sickle cell disease and asthma. *Children's Health Care, 31*, 69–86.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Hinds, P. S., Quarqnenti, A., Bush, A. J., Pratt, C., Fairclough, D., Rissmiller, G., et al. (2000). An evaluation of the impact of a self-care coping intervention on psychological and clinical outcomes in adolescents with newly diagnosed cancer. *European Journal of Oncology Nursing, 4*, 6–17.

Horvath, P. (1988). Placebos and common factors in two decades of psychotherapy research. *Psychological Bulletin, 104*, 214–225.

Hunt, M. (1997). *How science takes stock*. New York: Russell Sage Found.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy. *Journal of Consulting and Clinical Psychology, 59*, 12–19.

Kaplan, R., Chadwick, M., & Schimmel, L. (1985). Social learning intervention to promote metabolic control in type-I diabetes mellitus: Pilot experimental results. *Diabetes Care, 8*, 152–155.

Kaslow, N. J., Collins, M. H., Rashid, F. L., Baskin, M. L., Griffith, J. R., Hollins, L., et al. (2000). The efficacy of a pilot family psychoeducational intervention for pediatric sickle cell disease (SCD). *Families, Systems and Health, 18*, 381–404.

Kazdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology, 66*, 19–36.

Kibby, M. Y., Tyc, V. L., & Mulhern, R. K. (1998). Effectiveness of psychological intervention for children and adolescents with chronic medical illness: A meta-analysis. *Clinical Psychology Review, 18*, 103–117.

King, K., & Hanson, V. (1986). Psychosocial aspects of juvenile rheumatoid arthritis. *Pediatric Clinics of North America, 33*, 1221–1237.

Kolko, D., & Rickard-Figueroa, J. L. (1985). Effects of video games on the adverse colloraries of chemotherapy in pediatric oncology patients. *Journal of Consulting and Clinical Psychology, 53*, 223–228.

La Greca, A. M., & Varni, J. W. (1993). Editorial: Interventions in pediatric psychology: A look toward the future. *Journal of Pediatric Psychology, 18*, 667–679.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

McQuaid, E. L., & Nassau, J. H. (1999). Empirically supported treatments of disease-related symptoms in pediatric psychology: Asthma, diabetes, and cancer. *Journal of Pediatric Psychology, 24*, 305–328.

Moore, K. E., Geffken, G. R., & Royal, G. P. (1995). Behavioral intervention to reduce child distress during self-injection. *Clinical Pediatrics, 34*, 530–534.

Pendley, J. S., Kasmen, L. J., Miller, D. L., Donze, J., Swenson, C., & Reeves, G. (2002). Peer and family support in children and adolescents with type 1 diabetes. *Journal of Pediatric Psychology, 27*, 429–438.

Rose, N. I., Firestone, P., Heick, H. M. C., & Faught, A. K. (1983). The effects of anxiety management training on the control of juvenile diabetes mellitus. *Journal of Behavioral Medicine, 6*, 381–395.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59–82.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Shepherd, M., & Sartorius, N. (1989). *Non-specific aspects of treatment*. Toronto and Lewiston, NY: H. Huber.

Stark, L. J., Knapp, L. G., Bowen, A. M., Powers, S. W., Jelalian, E., Evans, S., et al. (1993). Increasing calorie consumption in children with cystic fibrosis: Replication with 2-year follow-up. *Journal of Applied Behavior Analysis, 26*, 435–450.

Stark, L. J., Mulvihill, M. M., Powers, S. W., Jelalian, E., Keating, K., Kreveling, S., et al. (1996). Behavioral intervention to improve caloric intake of children with cystic fibrosis: Treatment versus wait list control. *Journal of Pediatric Gastroenterology and Nutrition, 22*, 240–253.

Thomas, V. J., Dixon, A. L., & Milligan, P. (1999). Cognitive-behaviour therapy for the management of sickle cell disease pain: An evaluation of a community-based intervention. *British Journal of Health Psychology, 4*, 209–229.

Varni, J. W., Katz, E. R., Colegrove, R., & Dolgin, M. J. (1993). The impact of social skills training on the adjustment of children with newly diagnosed cancer. *Journal of Pediatric Psychology, 18*, 751–767.

White, D. M., Rusch, F. R., Kazdin, A. E., & Hartmann, D. P. (1989). Applications of meta-analysis in individual-subject research. *Behavioral Assessment, 11*, 281–296.

Wolf, F. M., Guevara, J. P., Grum, C. M., Clark, N. M., & Cates, C. J. (2003). Educational interventions for asthma in children. *Cochrane Database Systematic Reviews, 1*, CD000326.

Zeltzer, L. K. (1999). Commentary: Psychological treatment of physical symptoms: Who should get what? *Journal of Pediatric Psychology, 24*, 333–334.

Zeltzer, L., Dolgin, M. J., LeBaron, S., & LeBaron, C. (1991). A randomized, controlled study of behavioral intervention for chemotherapy distress in children with cancer. *Pediatrics, 88*, 34–42.