J Abnorm Child Psychol (2008) 36:567–589
DOI 10.1007/s10802-007-9201-9

A Meta-analytic Review of Components Associated with Parent Training Program Effectiveness

Jennifer Wyatt Kaminski · Linda Anne Valle · Jill H. Filene · Cynthia L. Boyle

Published online: 19 January 2008

Abstract This component analysis used meta-analytic techniques to synthesize the results of 77 published evaluations of parent training programs (i.e., programs that included the active acquisition of parenting skills) to enhance behavior and adjustment in children aged 0–7. Characteristics of program content and delivery method were used to predict effect sizes on measures of parenting behaviors and children's externalizing behavior. After controlling for differences attributable to research design, program components consistently associated with larger effects included increasing positive parent–child interactions and emotional communication skills, teaching parents to use time out and the importance of parenting consistency, and requiring parents to practice new skills with their children during parent training sessions. Program components consistently associated with smaller effects included teaching parents problem solving; teaching parents to promote children's cognitive, academic, or social skills; and providing other, additional services. The results have implications for selection and strengthening of existing parent training programs.

Keywords Parent training · Meta-analysis · Child behavior problems · Component analysis

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

J. Wyatt Kaminski (✉) · L. A. Valle
Division of Violence Prevention,
National Center for Injury Prevention and Control,
Centers for Disease Control and Prevention,
4770 Buford Hwy NE, MS-K60,
Atlanta, GA 30341, USA
e-mail: JKaminski@cdc.gov

L. A. Valle
e-mail: LValle@cdc.gov

J. H. Filene
James Bell Associates,
1001 19th Street, North, Suite 1500,
Arlington, VA 22209, USA
e-mail: Filene@jbassoc.com

C. L. Boyle
University of Kansas,
c/o Marcus Institute, 1920 Briarcliff Road,
Atlanta, GA 30329, USA
e-mail: lebezej777@aol.com

Early childhood behavior problems are generally characterized by oppositional, aggressive, impulsive, and inattentive behaviors. Although discrete instances of such behaviors are typical in very young children, pervasive and unremitting aggression and conduct problems in childhood reliably predict delinquent, aggressive, and risky behaviors in adolescence (e.g., Broidy et al. 2003; Fergusson et al. 1994; Tolan and Gorman-Smith 1998). As well, adolescents whose problem behaviors began in childhood commit more serious and violent acts and account for a disproportionate number of all youth offenses than adolescents without an early history of conduct problems (Farrington et al. 2003; Loeber et al. 1998; Tolan and Gorman-Smith 1998; Thornberry et al. 2003). Although most adolescent deviance discontinues at the end of the teenage years, individuals who exhibited conduct problems in childhood are more likely to engage in "life-course-persistent" antisocial behavior that continues through adolescence into adulthood (Moffitt and Caspi 2001; Moffitt et al. 2002). The life-course persistent pathway from childhood conduct problems to adult criminality and violent behaviors may best be interrupted early in life, when behavioral patterns are more easily modified (Tremblay 2006).

The late 1960s saw a shift in addressing children's problematic behaviors from child therapy, adolescent institutionalization, or juvenile adjudication focused exclusively on changing the child's undesirable behaviors to interventions focused on changing parents' behavior. This shift resulted from (a) the realization that parents, not just highly trained therapists, could act as agents of children's behavior change; and (b) the growing understanding of how parents contribute to their children's desirable and undesirable behaviors (e.g., Bandura 1969). Parenting programs have since proliferated, with different programs emphasizing different content (e.g., knowledge of typical child development, parenting self efficacy, communication skills, discipline or behavior management strategies), delivery settings (e.g., clinic-based therapy, community-based group sessions, individual home visits), delivery techniques used to engage parents and teach relevant content (e.g., group discussions, homework assignments, role playing), and types of families served (e.g., children with identified behavior problems, low-income adolescent parents, primiparous mothers) while still being considered parent training programs.

Over the years, the objectives of parent training programs have also extended well beyond the original intent of ameliorating identified child behavior problems. Other intended outcomes of parent training programs have included children's cognitive development (e.g., Cicchetti et al. 2000), anxiety (e.g., Barrett et al. 1996), and physical health (e.g., Reifsnider 1998). In addition, parenting programs are widely used by child welfare services for improving parenting practices among families at risk for child maltreatment, with over 800,000 families referred to parenting programs each year (Barth et al. 2005).

Not surprisingly, a number of meta-analyses have been published on parent training programs, including studies focusing exclusively on the effects of one or two specific programs (e.g., Cedar and Levant 1990; Thomas and Zimmer-Gembeck 2007), the effects associated with a specific delivery setting (e.g., Sweet and Appelbaum 2004), and the effects associated with programs using different theoretical orientation or approaches, such as behavioral parent training (e.g., Lundahl et al. 2006a, b; Maughan et al. 2005; Serketich and Dumas 1996). Three meta-analyses have also attempted to tease apart important elements of parent training programs by examining moderators of effect sizes, such as the characteristics of participating families (Lundahl et al. 2006a, b; Reyno and McGrath 2006). It is now generally accepted that parent training approaches can be effective. However, none of these meta-analyses examined specific program components, such as differences in content or delivery methods.

Since the inception of parent training, it has been assumed that the specific components of parent training approaches *are* the active ingredients that cause behavior change (Bernal et al. 1980; Westen et al. 2004). But despite the observed variation in how programs can be comprised, the components associated with more or less effective programs have rarely been examined. Westen et al. (2004) observed that the impossibility of determining if any manualized, packaged program is the most effective possible approach necessitates moving beyond examining complete treatment packages or theoretical orientation to examining individual strategies and processes of change. In such an analysis, individual program components (e.g., teaching empathic listening skills; engaging parents in role-playing of new skills) would be the variables of interest. Thus, one could examine the extent to which particular components are associated with larger or smaller program effects. Determining which components appear to be essential (or nonessential) across a variety of programs has important implications for practice. For example, when choosing among programs to implement, such information could be used to select programs containing components associated with greater program effectiveness. Existing programs could increase their effectiveness by integrating specific components reliably associated with greater effectiveness, rather than experiencing the higher costs, extensive time for staff retraining, and other barriers associated with adopting and implementing an entirely different packaged program. Similarly, it may be possible to eliminate components reliably associated with less effective programs, thereby minimizing the burden on practitioners and families.

The objective of the current study was to use meta-analytic techniques to determine which program components (i.e., content covered, delivery methods used) are reliably associated with more successful outcomes of parent training programs that targeted the prevention and/or remediation of early childhood behavior problems. Search parameters and inclusion criteria were selected to ensure a set of parent training program evaluations with a common set of outcomes, but with enough variability to investigate relationships between program characteristics and program effects. Data abstraction was structured to code for specific program components and strategies, rather than program type, to disassemble packaged programs into individual components. Analyses were designed to determine which components consistently appear to be associated with more effective programs, as indicated by the common metric of effect size. Multivariate analyses were utilized to separate the influence of program components from the influence of study method variance. Thus, these analyses decompose previously studied broad categories of theoretical approaches and packaged programs to provide a richer source of information about the key ingredients in effective parent training programs.

The decision to focus on parent training programs for managing children's behavior (rather than on programs intending only to change parental well-being or other outcomes) was based on the substantial body of research illustrating the direct effects of parenting practices on early child behavior problems. Although parental well-being is an important outcome, the effects of characteristics such as parental mental health and self esteem on child behaviors are frequently indirect effects, mediated by parenting practices (Capaldi et al. 2002; Conger et al. 1995; DeGarmo et al. 2004; Patterson et al. 2004). Parenting and child behavior problems were thus selected as the main outcomes of interest. Given the role of parents as agents of change for children's behavior problems, we predicted that effect sizes for parenting (immediate, proximal outcomes for a parent training program) would be larger than effect sizes for child behaviors (more distal outcomes).

As program evaluations often contain a variety of outcome measures related to changes in knowledge, attitudes, and behaviors, we hypothesized differences across types of outcome measures. Previous literature suggests that cognitive variables, such as self-reported knowledge and attitudes, change more readily than behavior (e.g., Albarracín et al. 2003; Webb and Sheeran 2006). In addition, changes on measures of proposed mediators of actual parenting behavior, such as parental distress or attitudes, following intervention may not ensure changes in parenting behavior and subsequent child outcomes (e.g., Chaffin and Valle 2003). Thus, we expected smaller effect sizes for parenting behaviors than for parenting knowledge, attitudes, or self-efficacy.

"Parent training program" was defined for this study as an intervention in which parents actively acquire parenting skills and may or may not have included other educational methods. Programs with the sole intent of providing parents with information (e.g., "parent education programs," parenting newsletters) that presume to change parents' behavior but did not utilize an active skills acquisition mechanism, were not included. Many, if not most, parent training programs also employ didactic or passive educational strategies, in addition to skills training. Thus, programs containing any active skills training component were included, whether or not they also contained an information–provision component. The decision to focus on programs containing an active acquisition component was based on decades of research showing that active learning approaches are superior to passive approaches (e.g., Arthur et al. 1998; Joyce and Showers 2002; Salas and Cannon-Bowers 2001; Swanson and Hoskyn 2001). We hypothesized that even within this set of parent skills training programs, components requiring more active parental involvement (e.g., practice during the session) would be associated with larger program effects, whereas components that required less active involvement (assigning homework that may or may not be completed and may or may not be completed correctly) would not.

Other hypotheses were designed to examine three commonly held, but rarely tested, assumptions and values in the field of parent training. First, it is often assumed that instruction in child development is not only necessary, but also sufficient, to ensure parental behavioral changes. This assumption may be an outgrowth of research suggesting maltreating parents have inappropriate behavioral expectations of their children (e.g., Tolliver et al. 1998). Although a focus on teaching child development knowledge is a logical response to such findings, intervention research from a wide variety of health outcomes (e.g., smoking, physical activity, nutrition, HIV) indicates that changes in knowledge do not guarantee changes in behavior. Although parent training programs ought to be teaching developmentally appropriate parenting skills, we predicted that program content specific to knowledge of child development and developmental expectations would not be strongly associated with program effects.

A second assumption is that manualized programs, which have historical connections with empirically supported therapies (Westen et al. 2004), produce better outcomes than programs without a curriculum. Given that the main goal of standardization is to minimize variability of program implementation and ensure delivery of the program's effective ingredients (Westen et al. 2004), we expected our data to support that assumption. Thus, programs utilizing a manual or curriculum would report larger effect sizes than programs without a standardized protocol.

A third strongly held assumption in the field is that "more is better"—that is, providing families with more services will lead to better outcomes (e.g., Lundahl and Harris 2006). Parent training programs often include other supplemental services, such as anger or stress management, substance abuse treatment, or job skills training. Although the provision of a larger array of services might lead to changes in other parent or family outcomes (e.g., abstinence from substance use, improved economic circumstances), we assert that these ancillary services are less likely to manifest in proximal, immediate changes in parenting or child behaviors than program components directly targeting parenting skills. Such additional services may present an overwhelming burden or impede parents' ability to focus on and master parenting skills. We therefore hypothesized that supplemental program services would be associated with smaller effect sizes.

The current component analysis was designed to allow for a meta-analytic comparison of approaches across a range of program content and delivery strategies. As in all meta-analyses, study selection and analytic decisions have

important ramifications for the generalization of results. A common inclusion requirement is the use of random assignment. Although the randomized design has been labeled the gold standard of intervention research, there are convincing theoretical reasons and empirical evidence to suggest that studies using random assignment should not be the only studies included in a meta-analysis such as this one.

From a theoretical perspective, including randomized and non-randomized designs broadens the scope of programs being summarized. In a meta-analysis of 382 juvenile delinquency interventions, Lipsey found randomized designs were more often used to evaluate "research/demonstration" programs, whereas non-randomized designs were more often employed by "practice" programs (Lipsey 2003; Wilson and Lipsey 2003). If non-randomized designs are excluded from a meta-analysis, the results are likely to estimate poorly the effects of interventions practiced outside the research laboratory, which are often the source of innovations that inform the development and empirical testing of laboratory-based interventions (Westen et al. 2004). In addition, the results would be limited to the small number of parent training programs that have been evaluated under tightly controlled conditions. To draw conclusions representative of the entire field, non-randomized designs must be allowed to contribute to the results.

From an empirical perspective, by only including randomized designs, a meta-analysis is potentially biased on a number of additional dimensions. Lipsey (2003; Wilson and Lipsey 2003) demonstrated that random assignment was associated with a number of methodological and programmatic variables that were significant moderators of effect size (e.g., attrition, sample size, participant gender, and intervention dosage). Evidence further suggests that randomization, absent from other study factors associated with randomization, accounts for a limited amount of the variance in effect sizes (e.g., Wilson and Lipsey 2003). Thus, randomized and non-randomized studies are included here.

Another common meta-analytic strategy is to assign quality ratings or scores to individual studies, which are then used as effect size weights or statistical controls. In some instances, quality ratings merely reflect coders' judgments about study design choices, such as coders' perceptions of the appropriateness of statistical analyses or of the reliability of the outcome measure. This approach has been strongly criticized, as quality ratings are subjective, not objective, measures of methodological rigor, and may or may not be valid indicators of the extent and direction of bias related to study design (Greenland 1994).

In other meta-analyses, additive indices of methodological rigor (e.g., quality scores) have been computed based on an assumption of accumulated benefits (i.e., the more indicators of methodological rigor embodied by a particular study, the more accurate effect size estimates will be). However, this type of computation is only appropriate when the individual indicators predict effect size in the same direction—for example, if the use of random assignment, use of standardized measures, and assessment of treatment fidelity all predict smaller effect sizes—an assumption which is rarely tested. Further, an additive combination will mask any potential differences in the strength of the relationship between particular indicators and overall effect size. It is thus important to examine how individual indicators of methodological rigor are related to effect sizes and to include those indicators separately in analyses (Greenland 1994) to control for the possibility the indicators are differentially associated with effect size estimates.

Materials and Methods

Document Search

Because the objective was to conduct a broadly inclusive analysis of parent training programs, methodology was designed to err on the side of over-inclusion. The search was conducted in September 2002 and limited to publications from 1990 forward. A wide range of search strings was enacted in PsycInfo and Medline under four broad headings (Program Descriptors, Program Targets, Evaluation Descriptors, and Program Outcomes). Program Descriptors were Keyword search terms to elicit parent training programs: parent (or parenting) program, parent (or parenting) education, and parent training. Thirteen Program Targets were Subject Heading search terms to identify articles related to parenting behaviors (e.g., parenting skills, family relations, child management). Four Evaluation Descriptors returned publications describing an outcome study (e.g., treatment, intervention). Due to the large number of citations returned, this list was further restricted by requiring the documents to include the word "result" or "outcome" and the word "parent" or "family." The Program Outcomes were 28 terms describing the domains of intended program effects (e.g., conduct problems, social skills). These Program Descriptors, Program Targets, Evaluation Descriptors, and Program Outcome lists were crossed to create three matrices of search terms resulting in 175 pairs of search terms, which were used as search strings in each database. The number of unduplicated studies returned by the multi-tiered strategy was 8,277. Of these, 93 were literature reviews and meta-analyses, from which we identified additional relevant publications. A secondary search was conducted on program and author names that appeared in the original search results at least twice. These

two follow-up strategies and the initial search results provided 8,499 total abstracts for possible inclusion.

Document Review and Retrieval

Inclusion criteria were selected to define the scope of the meta-analysis to evaluations of parent training programs (defined as programs that included the active acquisition of parenting skills) for children's early behavior problems. "Parent" is used generically here to refer to primary caregivers, so programs for adoptive parents, foster parents, grandparents, and other relatives were considered relevant. Because our focus was on early problem behaviors, we selected the target ages of 0- to 7-year-old children. Programs targeted solely to literacy outcomes or intended only to increase parent–child shared reading times and sensorimotor stimulation therapies (e.g., "kangaroo care," infant massage) were excluded for their explicit focus on literacy or physical outcomes rather than early child behavior problems.

Criteria were also selected to ensure that conclusions made about specific program content and delivery components were generalizable to a broad population base. Thus, programs targeted to parents with developmental disabilities or parents of children with developmental disabilities were excluded, as were programs intended for children who had experienced bereavement, traumatic brain injury, and feeding or other developmental disorders of infancy.

Finally, we only included programs that were conducted in English and published as an article, book, or book chapter. Although limiting a meta-analysis to published documents introduces the potential for bias in favor of significant results, there are accepted methods for examining the extent to which the file-drawer problem may influence the results. Limitations in our institution's ability to procure unpublished documents resulted in a large percentage of unpublished documents that could not be obtained, and there were no systematic means of assessing differences between obtained and unobtainable documents. Because there is currently no accepted method to estimate the magnitude or direction of the bias this procurement problem would have introduced, the standard of computing and presenting the fail-safe N was followed rather than introduce an unknown source of bias.

Review of all 8,499 abstracts was conducted independently by three of the investigators (Kaminski, Filene and Boyle). Articles were excluded at this stage (n=7493) only if all three investigators agreed that it met at least one of the exclusion criteria. Following retrieval and cursory review of the remaining 1006 documents, an additional 205 were rejected. There were also 47 studies evaluating parent training programs that met the inclusion criteria, but used single-case evaluation methods. Such designs require different effect size computations and should not be combined with other designs in a meta-analysis (Faith et al. 1997). Next, because computation of the traditional effect size for meta-analysis involves the standardized difference between mean outcome scores of the "treatment" group and the "control" group (Hedges and Olkin 1985), fifty studies that lacked a control or comparison group were excluded. A large number of studies (n=576) were excluded because they did not contain enough or appropriate statistical information with which to calculate effect sizes (e.g., standard deviations were not reported). The resulting pool of 128 evaluations was appropriate for meta-analysis.

Data Abstraction

Before coding independently, data abstractors were trained to criteria of coding three consecutive articles with greater than 90% accuracy in each of the broad categories of variables of interest (described in more detail below). Forty studies were randomly selected at different time intervals and independently coded by a second data abstractor to assess reliability and control for drift over time. Total percent agreement (PA) and Cohen's kappas (for categorical variables) were computed. For the variables included in the current study, mean PA=96.50 (SD=3.75) and median PA=96.07, with a range of 82.50 to 100.00, whereas mean kappa=0.89 (SD=0.10) and median kappa=0.92, with a range of 0.64 to 1.00. Interrater agreement was considered adequate for the purposes of the present study, with 10% of kappas falling in the substantial range and 90% falling in the almost perfect reliability ranges described by Landis and Koch (1977). In addition, 100% of the outcome and sample size data used to compute effect sizes were verified for accuracy by the first or second author.

Based on recommendations for conducting meta-analysis of evaluation studies (e.g., Stock 1994; Wilson and Lipsey 2001), coding forms were designed to capture information about the document, intervention, participants, evaluation design, outcome measures, and results. Secondary sources cited as providing additional information about the program or study (e.g., "a more complete description of the program was reported in...") were retrieved, and information specifically referenced in the initial article was coded. Data elements characterizing the intervention consisted of items related to the parent training program's location, dosage, content, and delivery. Information about participants included the number, age, gender, and ethnicity of parents or children. Table 1 presents descriptive information, when available, indicating whether the program was focused on prevention or treatment of existing child behavior problems, the program intent, and participants' gender and child age. Tables 2 and 3 describe the variables coded for

Table 1 Program and participant descriptive information

Citation	Prevention (P) or treatment (T) of child behavior problems	Program intent	Parent participants[a]	Child gender[b] (% male)	Child age[c], in years (mean and [standard deviation] or range)
Adams and Kelley (1992)	T	Treatment of sibling aggression	Mothers	45[e]	5.72 (2.59)[e]
Anastopoulos et al. (1993)	T	Treatment of Attention-Deficit/ Hyperactivity Disorder (ADHD)	Mothers	NR	8.14 (1.06)[e]
Barber (1992)	P	Enhance parenting (of mothers on public assistance)	Mothers	NR	0–15[d]
Barhava-Monteith et al. (1999)	P	Promote positive school behavior and readiness	NR	NR	NR
Barkley et al. (1996)	T	Treatment of aggressive-hyperactive-impulsive behaviors	NR	66[e]	4–5[e]
Barrett et al. (1996)	T	Treatment of childhood anxiety	Both	58[e]	10.10 (1.9)[d]
Barth (1991)	P	Prevent child abuse	Mothers	NR	NR
Bradley et al. (1994)	P	Promote child development (in premature/low-birthweight children)	NR	50[d]	NR
Bratton and Landreth (1995)	P	Enhance parenting (of single parents)	Either	55[d]	4.45 (NR)[d]
Budd and Itzkowitz (1990)	P	Promote child social competence	Both	NR	8.08 (NR)[e]
Cicchetti et al. (2000)	P	Prevent child problem behaviors (with depressed mothers)	Mothers	51[d]	NR
Connell et al. (1997)	T	Treatment of oppositional behavior	Both	58[d]	NR
Connolly et al. (2001)	T	Treatment of behavior problems	NR	76[d]	NR
Costas and Landreth (1999)	P	Enhance parenting (of nonoffending parents of children who have been sexually abused)	Either (or grandmother)	23[d]	4–10[e]
Cunningham et al. (1995)	T	Treatment of disruptive behavior disorders	Both	57[d]	4.36 (0.38)[d]
Davis and Spurr (1998)	T	Treatment of multiple psychosocial problems	Both	62[d]	2.50 (1.49)[d]
Evans and Okifuji (1992)	T	Treatment of child behavior problems	NR	64[d]	NR
Evans et al. (1993)	T	Treatment of child behavior problems and family dysfunction	Both	65[d]	4–9[d]
Fennell and Fishel (1998)	P	Prevent child abuse	Either	30[d]	7.50 (1.50)[d]
Fewell and Wheeden (1998)	P	Promote positive parenting and child development (in adolescent mothers)	Mothers	46[d]	1.11 (0.67)[d]
Fitzpatrick et al. (1996)	P	Promote child development	Mothers	56[d]	1.47 (NR)[d]
Forgatch and DeGarmo (1999)	P	Enhance parenting (of divorcing mothers with sons)	Mothers	100[d]	NR
Forgatch and DeGarmo (2002)	P	Prevent child problem behaviors (with divorcing mothers with sons)	Mothers	100[d]	7.80 (0.93)[e]
Gaudin et al. (1990)	P	Enhance parenting (of neglectful parents)	NR	NR	NR
Gelfand et al. (1996)	P	Promote child development (with depressed mothers)	Mothers	54[d]	0.53 (0.24)[d]
Glover and Landreth (2000)	P	Promote positive parenting (with Native American parents)	Mothers (or grandmother or stepmother)	55[d]	5.50 (NR)[d]
Griff (1999)	T	Treatment of child behavior problems	Both	NR	2–6[e]
Gross et al. (1995)	T	Treatment of child behavioral difficulties	Both	NR	2–3[e]
Harris and Landreth (1997)	P	Promote parent–child relationships (with incarcerated mothers)	Mothers	NR	5 (NR)[e]

Table 1 (continued)

Citation	Prevention (P) or treatment (T) of child behavior problems	Program intent	Parent participants[a]	Child gender[b] (% male)	Child age[c], in years (mean and [standard deviation] or range)
Horn et al. (1990)	T	Treatment of ADHD	Both	81[e]	8.66 (1.55)[d]
Hutcheson et al. (1997)	P	Promote positive parenting and child development (in children with failure-to-thrive)	Mothers	46[d]	1.10 (0.48)[e]
Infant Health and Development Program (1990)	P	Promote child development (in low-birthweight children)	NR	NR	NR
Jason et al. (1994)	P	Prevent child substance use	Either	NR	NR
Johnson et al. (1993)	P	Promote positive parenting and child development (with first-time, low-income mothers)	Mothers	49[d]	NR
Kale and Landreth (1999)	P	Prevent child behavior problems (among children with learning difficulties)	Either (or grandmother)	NR	NR
Kissman (1992)	P	Promote positive parenting and child development (with adolescent mothers)	Mothers	NR	NR
Koniak-Griffin and Verzemnieks (1991)	P	Promote positive parenting (with adolescent mothers)	Mothers	NR	NR
Koniak-Griffin et al. (1992)	P	Promote positive parenting (with adolescent mothers)	Mothers	NR	NR
Landreth and Lobaugh (1998)	P	Enhance parenting (of incarcerated fathers)	Fathers	6[d]	5.94 (NR)[d]
Lee and Holland (1991)	P	Promote effective foster parenting	NR	NR	NR
Letourneau et al. (2001)	P	Promote child mental health (among low-income and/or adolescent parents)	Mothers	NR	NR
Lyons-Ruth et al. (1990)	P	Promote positive parenting (with low-income, depressed mothers)	Mothers	NR	0.39 (NR)[e]
MacKenzie and Hilgedick (1999)	P	Prevent child behavior problems	Either	49[e]	4.26 (0.73)[e]
Magen and Rose (1994)	T	Treatment of aggressive or noncompliant behavior	Either	68[e]	7.89 (NR)[d]
Maguin et al. (1994)	P	Prevent child conduct problems (with children of alcoholic fathers)	Mothers	NR	4.40 (NR)[e]
McBride (1991a)	P	Promote positive parenting (with fathers)	Fathers	53[d]	3.42 (0.85)[d]
McBride (1991b)	P	Promote positive parenting (with fathers)	Fathers	NR	3.68 (NR)[e]
McNeil et al. (1999)	T	Treatment of disruptive behavior problems	NR	NR	2–8[e]
McNeil et al. (1991)	T	Treatment of child behavior problems	NR	100[d]	4.5 (NR)[d]
Metropolitan Area Child Study Research Group (2002)	P	Prevent child aggression	Both	61[e]	7–13[e]
Mullin and Quigley (1994)	P	Promote positive child behavior	Mothers	NR	0.25–14[e]
Myers et al. (1992)	P	Prevent child problem behaviors	NR	NR	NR
Nelson and Levant (1991)	P	Prevent child problem behaviors (in stepfamilies)	Either	50[d]	14.00 (4.49)[d]
Nicholson et al. (2002)	P	Promote positive parenting (with low-income parents using harsh discipline)	Both	NR	1–5[e]
Nixon et al. (2003)	T	Treatment of Oppositional Defiant Disorder (ODD)	Both	82[d]	3.90 (0.55)[e]
Odom (1996)	T	Treatment of ADHD	Mothers	NR	5–11[e]
Pehrson and Robinson (1990)	P	Promote positive parenting	Either	NR	NR

Table 1 (continued)

Citation	Prevention (P) or treatment (T) of child behavior problems	Program intent	Parent participants[a]	Child gender[b] (% male)	Child age[c], in years (mean and [standard deviation] or range)
Pisterman et al. (1992a)	T	Treatment of hyperactivity	Either	87[d]	4.06 (NR)[d]
Pisterman et al. (1992b)	T	Treatment of ADHD	Either	96[d]	3.89 (0.62)[d]
Reid et al. (1999)	T	Treatment of child's bedtime behavior problems	Both	NR	NR
Reifsnider (1998)	P	Promote positive parenting and child development (in children with growth deficiency)	Mothers	48[d]	1.5 (NR)[d]
Sanders et al. (2000a)	T	Treatment of conduct problems	Both	66[d]	3.36 (0.29)[d]
Sanders et al. (2000b)	P	Prevent disruptive behaviors	Mothers	57[d]	4.8 (1.35)[d]
Schuhmann et al. (1998)	T	Treatment of ODD	Both	NR	4.93 (1.03)[e]
Schuler et al. (2002)	P	Promote positive parenting and child development (with drug-using mothers)	Mothers	43[d]	NR
Sonuga-Barke et al. (2001)	T	Treatment of ADHD	Mothers	62[e]	NR
Taylor et al. (1998)	T	Treatment of child behavior problems	Both	67[d]	5.6[e]
Thompson et al. (1996)	T	Treatment of child behavior problems	Either	62[d]	10.18 (3.82)[d]
Vines and Williams-Burgess (1994)	P	Enhance parenting (of first-time, at-risk mothers) and prevent child abuse or neglect	Mothers	NR	0–0.17[d]
Wahler and Meginnis (1997)	P	Increase child compliance	Mothers	42[d]	7.45 (NR)[e]
Walker et al. (1998)	T	Treatment of child behavior problems	NR	NR	NR
Webster-Stratton (1990)	T	Treatment of conduct problems	Both	79[e]	5.1 (NR)[e]
Webster-Stratton (1992)	T	Treatment of conduct problems	Both	NR	5.02 (NR)[e]
Webster-Stratton (1998)	P	Prevent child conduct problems	Both	54[d]	4.70 (0.36)[d]
Webster-Stratton and Hammond (1997)	T	Treatment of conduct problems	Both	81[d]	5.33 (1.27)[d]
Webster-Stratton et al. (2001)	P	Prevent child conduct problems	Both	50[d]	4.59 (0.36)[e]
Wilczak and Markstrom (1999)	P	Promote positive parenting (with incarcerated fathers)	Fathers	NR	NR

NR Not reported

[a] Whether the intervention included: Mothers only, Fathers only, Either parent (mother or father), or Both parents.

[b] The reported percent of treatment group or whole sample that were boys. Because some evaluations included three or more groups, "whole sample" demographics are not necessarily representative of the children who received the focal intervention included here. When available, child gender is reported for treatment group ([d]) of interest only. If not available, child gender is reported for the whole sample ([e]).

[c] When available, child age is reported for treatment group ([d]) of interest only. If not available, child gender is reported for the whole sample ([e]).

[d] Treatment group

[e] Whole sample

program content and delivery components, respectively. We defined five categories of parent outcome measures and four categories of child outcome measures (see Table 4). Although the primary outcomes of interest were parenting behaviors and child behavior problems, we included parenting outcomes often considered to be direct mediators of parenting behaviors, such as attitudes about parenting and parenting self-efficacy. We also included child outcomes in addition to children's externalizing behaviors when they were presented. For each reported outcome measure within these nine categories, the method of data collection (e.g., survey, observation) and the reporter for that measure (e.g., child care provider) were also coded.

Guided by Wilson and Lipsey's (2001) analysis of 319 meta-analyses, we expected that effect size would be related to particular study indicators of methodological rigor. Thus,

Table 2 Program content variables and description (coded as present or absent)

Coded variable	Description
Child development knowledge and care	Providing developmentally appropriate physical care and environment (e.g., feeding, diapering, home safety); learning about typical child development and behavior; fostering children's positive emotional development (e.g., self-esteem, providing stimulating environment)
Positive interactions with child	Learning the importance of positive, non-disciplinary interactions with children; using skills that promote positive parent–child interactions (e.g., demonstrating enthusiasm, following child's interests, offering appropriate recreational options); providing positive attention
Responsiveness, sensitivity, and nurturing	Responding sensitively to child's emotional and psychological needs (e.g., soothing); providing developmentally appropriate physical contact and affection
Emotional communication	Using relationship-building communication skills (e.g., active listening); helping children identify and appropriately express emotions
Disciplinary communication	Giving clear and developmentally appropriate directions; setting limits and rules; stating behavioral expectations and consequences
Discipline and behavior management	Coded into specific variables: A. Attitudes about discipline strategies B. Attributions about child misbehaviors C. Monitoring and supervision practices D. Specific reinforcement and punishment techniques: e.g., planned ignoring, positive reinforcement, time out E. Problem solving about child behaviors F. Consistent responding or generalization
Promoting children's social skills or prosocial behavior	Educating parents to teach children to share and cooperate, use good manners, and get along with peers, siblings, or adults
Promoting children's cognitive or academic skills	Using incidental teaching; fostering children's language or literacy development; enhancing child's school readiness

four items about potential threats to internal validity were created to gauge the degree to which the reported effect sizes reflected true treatment effects (rather than variability related to methodological artifacts):

1. Participant assignment to condition was coded as *randomly assigned individuals to conditions* or *other assignment procedure* (e.g., random assignment of groups, non-random assignment).
2. The extent to which authors examined the initial equivalence of groups on outcome measures at baseline was coded as *assessed* or *not assessed*.
3. The type of comparison group in each study was coded as *comparison/control group received no treatment* (including using a wait-list comparison/control group), or *comparison/control group received alternate treatment*. This distinction would allow for an examination of absolute versus relative efficacy of the programs (Wampold 2001).
4. Specificity of the treatment condition was coded as *parent training program alone* (if the parent training program was being evaluated as a stand-alone program) or *parent training program plus additional interventions* (if the parent training program was being

Table 3 Program delivery variables and description (coded as present or absent)

Coded variable	Description
Curriculum or manual	Use of an established course of parent training, as evidenced by authors' report of a curriculum or manual
Modeling	Presenting live or recorded demonstrations of parenting behaviors
Homework	Written, verbal, or behavioral assignments to complete between sessions
Rehearsal, role-playing, or practice	In-session opportunities for parents to practice skills through rehearsal and role-playing situations, coded more specifically when possible: A. Role playing with the parent trainer or a peer B. Parent practicing skills with the parent's own child
Separate child instruction	Child participated in behavioral skills training separately from the parent
Ancillary services	Program included supplementary content beyond that specific to parenting skills (e.g., mental health or substance abuse services, case management or referrals, social support, stress/anger management, educational assistance)

Table 4 Means and standard errors for effect sizes of specific outcomes

Outcome category	Number of studies	Mean (SE) effect size
All parenting measures	54	0.43 (0.03)
Knowledge or information acquisition	6	0.88 (0.14)
Self-efficacy	19	0.49 (0.06)
Attitudes and values	14	0.47 (0.09)
Behaviors and skills	40	0.39 (0.03)
All child measures	58	0.30 (0.03)
Externalizing behaviors (e.g., noncompliance, aggressive or hyperactive behavior)	48	0.25 (0.03)
Internalizing behaviors (e.g., anxiety, depression)	11	0.40 (0.08)
Educational and cognitive outcomes (e.g., school grades, literacy, cognitive development)	9	0.26 (0.04)
Social skills and social competence	15	0.13 (0.06)
Parent–child interaction (e.g., observational measures of dyadic interaction)	7	0.22 (0.10)

evaluated as part of a wider array of interventions). This treatment condition specificity variable was distinct from the ancillary services component presented in Table 3. Both variables assessed whether or not the treatment group received assistance beyond parent training specifically for child behavior. However, the treatment condition specificity variable was a research design factor, indicating that families received additional interventions of unknown dosage, type, and quality, the effects of which could not be distinguished from the effects of the parent training program in the evaluation. The ancillary services variable (a program component) indicated that participants received additional services that were a routine part of the program and were delivered by program staff.

Effect Size Calculation

The conventional effect size used for treatment outcome studies is the standardized mean difference statistic (Hedges and Olkin 1985), in which the post-test mean of the comparison group is subtracted from the post-test mean of the treatment group, the result of which is divided by the pooled standard deviation. Effect sizes were calculated from the unadjusted means, standard deviations, and sample sizes, when available, for the treatment and comparison groups at post-test for each measure of interest. The ES software program (Shadish et al. 1999) was used to calculate the standardized mean difference from other statistics such as the independent samples t test or the F statistic from an analysis of variance. Information from adjusted means or analyses of covariance was not included. Because different authors and statisticians rarely use the same set of covariates, combining and comparing differentially adjusted effect sizes would limit our ability to determine if differences in program effects were due to program characteristics or to statistical choices of study authors. Hedges' (1983) small sample correction was applied to all effect sizes to obtain unbiased estimates prior to analysis.

Evaluations frequently reported results on multiple outcome measures, multiple time points, multiple subsamples, or multiple treatment conditions. To avoid overrepresenting any study with multiple effect sizes and violating the statistical independence assumption, only one effect size per study can be included in any analysis. The two recommended strategies for handling multiple effect sizes per study are selection of effect sizes or aggregation of effect sizes within a study (Wilson and Lipsey 2001).

Multiple Published Reports When two or more publications included results from the same sample, at the same time point, on the same measures, the most comprehensive publication (e.g., the one with the larger sample or with more reported outcomes) was selected. If other publications from the same trial reported other effect sizes, they were coded only for the additional effect sizes.

Multiple Outcome Measures Because evaluations frequently included more than one measure of each defined outcome of interest (see Table 4), effect sizes were computed separately and then selected or aggregated as appropriate to the level of analysis. For example, one set of analyses concerns "overall parenting outcomes." In these analyses, effect sizes based on reports of parent measures were selected, and multiple effect sizes of parenting outcomes within a particular study were averaged to produce one parenting outcome effect size per study.

Multiple Time Points A number of studies reported outcome measurements at mid-treatment, post-treatment, and/or at follow-up periods after the program had ended. The intent of the current analyses is to report on immediate post-test effects, so only those studies or effect sizes were selected for inclusion.

Table 5 Mean effect sizes associated with indicators of methodological rigor

Rigor indicator	Percent of studies	Mean (SE) effect size
Participants randomly assigned?		
Yes	58.4	0.39 (0.03)[a]
No	41.6	0.26 (0.04)[a]
Initial equivalence assessed?		
Yes	32.5	0.23 (0.04)[a]
No	67.5	0.39 (0.03)[a]
True "no-treatment" control group?		
Yes	83.1	0.36 (0.03)
No	16.9	0.31 (0.04)
Parent training program evaluated as stand-alone program?		
Yes	75.3	0.46 (0.03)[a]
No	24.7	0.22 (0.03)[a]

[a] Means for "yes" versus "no" were significantly different in an inverse-variance-weighted oneway ANOVA ($p<0.01$).

Multiple Samples For some studies, outcomes were presented separately for different sub-samples, such as mothers and fathers, depressed and non-depressed mothers, or primiparous and multiparous mothers. In these cases, an effect size was computed separately for each subgroup and averaged to produce a single effect size for that sample.

Multiple Treatment Conditions Many evaluations did not fall into a two-condition "treatment" versus "no treatment" design. A common design included a "no treatment" control/comparison group, a "treatment" group (i.e., a group that received only parent training), and an "enhanced treatment" group (e.g., a group that received parent training plus a separate preschool-based children's program). The effect size of primary interest was the one representing differences between a group that received a parent training program and a group who did not receive the parent training program. Thus, the "treatment" vs. "no treatment" effect size was preferred. When the study design precluded that effect size, either a "treatment" vs. "alternate treatment" or an "enhanced treatment" vs. "no treatment" effect size was chosen. Differences in effect sizes due to these different designs were accounted for by inclusion of two of the methodological rigor variables in the final analyses (type of comparison group, and evaluating the parent training program as a stand-alone intervention).

Through the described selection and aggregation processes, the pool of 128 published studies was reduced to 77 studies for which appropriate, relevant comparisons were reported. We followed the conventional method of accounting for differences in studies' ability to accurately reflect population differences, by weighting each effect size by the inverse of the variance (Hedges and Olkin 1985). All effect sizes reported are weighted effect sizes.

Results

The overall weighted effect size of the final set of 77 studies across all coded outcomes was 0.34 (95% CI=0.29–0.39), reflecting a significant mean difference between treatment and comparison groups at immediate post-test of just larger than a third of a standard deviation. A calculation of Orwin's (1983) fail-safe N revealed that 250 unpublished studies with non-significant results would have to exist in order to reduce the overall effect size to 0.08 (below the conventional significance level of $p<0.05$). Effect sizes ranged from −0.61 to 3.69.[1]

The Q test of homogeneity of effect sizes was significant (Q [76]=330.91, $p<0.001$), indicating marked variability in reported effect sizes across studies, thereby warranting examination of potential moderators of effect size. Moderator analyses were initially undertaken based on an assumption of fixed effects, as our hypotheses involved investigating sources of variability related to effect size (e.g., different content and delivery methods) that are not randomly assigned or distributed across studies (Hedges 1994). However, the final analyses were subsequently repeated with the inclusion of a random-effects variance component, to test the robustness of the results to potential

[1] Inspection of the full set of effect sizes suggested a small number of potential outliers on the positive and negative tails of the distribution. Adjusting those values to less extreme ones (i.e., Winsorizing) reduced the overall effect size slightly, but not significantly. Because the important question was whether or not extreme values influenced the results of the more complex hypothesis-driven analyses, all subsequent analyses were conducted with and without adjusting potential outliers. None of the patterns of results changed as a function of Winsorizing. For simplicity in presentation, all reported results are based on original computed effect size values (i.e., to avoid describing multiple Winsorizing processes across different sets of effect sizes).

program and method variations not represented in this sample of studies (Raudenbusch 1994).

First, an overall "parent" effect size and an overall "child" effect size were computed for each study by aggregating across all coded outcome measures with the parent or child as referent, respectively. The overall parent effect represents an aggregate of measures of parenting knowledge, attitudes, values, self-efficacy, and behavior; and the overall child effect represents an aggregate of measures of children's externalizing, internalizing, academic, or socially competent behavior. Each category of measure (e.g., parent knowledge and information acquisition, parenting behaviors and skills) was also examined separately. Visual inspection revealed that the obtained effect sizes adhered to hypothesized patterns, as shown in Table 4. The overall parent effect size was 0.43, compared to the overall child effect size of 0.30. Within the parent measures, there were also marked differences. The average effect size for Knowledge and Information Acquisition was large (0.88) and more than twice the magnitude of the average effect size for Behaviors and Skills (0.39). Among the child behavior measures, internalizing behavior outcomes appeared to have the largest effect sizes, and social skills outcomes had the smallest effect sizes. Unfortunately, because individual studies frequently assessed multiple outcomes, assumptions of independence would be violated by testing for significant differences across referents or categories.

The next set of moderators examined four indicators of methodological rigor in the evaluation design to ascertain the extent to which effect sizes specifically reflected the impact of the parent training program rather than methodological influences or biases. Using available macros for SPSS (Wilson 2002), analogues to oneway ANOVAs were conducted on overall study effect sizes with each of the four independent variable indicators of rigor. As shown in Table 5, three of the four rigor variables predicted significant differences in effect sizes. Studies that randomly assigned participants to conditions reported significantly larger effect sizes than studies that did not randomly assign participants. Studies in which initial equivalence of groups was assessed reported significantly smaller effect sizes than those that did not assess initial equivalence. Studies with a true no-treatment control/comparison group reported similar average effect sizes as studies in which the control/comparison group received alternate treatment or services. Finally, when the parent training program was being tested as a stand-alone intervention, effect sizes were larger than when the program was embedded in a package of interventions.

The analyses examining program content and delivery components are presented in Tables 6 and 7. First, we examined bivariate associations between the presence versus absence of each component and effect sizes. Given the observed variation in effect sizes across the different measure categories shown in Table 4 (e.g., relatively larger

Table 6 Effect sizes and predictors of parenting behaviors/skills outcomes

Program component or strategy	Effect size (SE) for programs with this component	Effect size (SE) for programs without this component	Standardized regression weight, controlling for indicators of methodological rigor and parent self-report
Child development knowledge and care	0.37 (0.05)	0.41 (0.05)	0.021
Positive interactions with child	0.39 (0.04)	0.38 (0.06)	0.198[b]
Responsiveness, sensitivity, and nurturing	0.39 (0.09)	0.39 (0.03)	0.017
Emotional communication	1.47 (0.17)[a]	0.35 (0.03)[a]	0.437[b]
Disciplinary communication	0.45 (0.05)	0.35 (0.04)	0.090
Positive reinforcement	0.38 (0.05)	0.40 (0.04)	−0.019
Time out	0.52 (0.08)	0.36 (0.04)	0.013
Problem solving	0.13 (0.11)[a]	0.41 (0.03)[a]	−0.247[b]
Consistent responding	0.59 (0.10)[a]	0.36 (0.03)[a]	0.083
Promoting children's social skills	0.23 (0.07)[a]	0.42 (0.04)[a]	−0.144
Promoting children's cognitive/academic skills	0.29 (0.05)[a]	0.48 (0.05)[a]	−0.243[b]
Curriculum or manual	0.38 (0.04)	0.40 (0.07)	0.073
Modeling	0.36 (0.05)	0.41 (0.04)	0.048
Homework	0.39 (0.05)	0.38 (0.04)	−0.003
Role play	0.45 (0.05)	0.35 (0.04)	0.169
Practice with own child	0.91 (0.11)[a]	0.33 (0.03)[a]	0.375[b]
Separate child instruction	0.37 (0.05)	0.41 (0.05)	0.041
Ancillary services	0.30 (0.04)[a]	0.61 (0.06)[a]	−0.205[b]

[a] Means for "with" versus "without" were significantly different in an inverse-variance-weighted oneway ANOVA ($p<0.05$).
[b] Component's regression weight from inverse-variance-weighted multiple linear regression was significant ($p<0.05$).

Table 7 Effect sizes and predictors of child externalizing behavior outcomes

Program component or strategy	Effect size (SE) for programs with this component	Effect size (SE) for programs without this component	Standardized regression weight, controlling for indicators of methodological rigor
Child development knowledge and care	0.25 (0.05)	0.24 (0.04)	0.054
Positive interactions with child	0.36 (0.04)[a]	0.13 (0.04)[a]	0.284[b]
Responsiveness, sensitivity, and nurturing	0.58 (0.11)[a]	0.22 (0.03)[a]	0.087
Emotional communication	0.03 (0.06)[a]	0.33 (0.04)[a]	−0.159
Disciplinary communication	0.24 (0.04)	0.25 (0.05)	0.034
Positive reinforcement	0.23 (0.04)	0.27 (0.05)	0.114
Time out	0.54 (0.06)[a]	0.15 (0.03)[a]	0.170[b]
Problem solving	0.49 (0.10)[a]	0.22 (0.03)[a]	0.096
Consistent responding	0.22 (0.05)	0.26 (0.04)	0.333[b]
Promoting children's social skills	0.02 (0.04)[a]	0.38 (0.04)[a]	−0.198[b]
Promoting children's cognitive/academic skills	0.27 (0.05)	0.23 (0.04)	0.062
Curriculum or manual	0.21 (0.03)[a]	0.45 (0.08)[a]	−0.118
Modeling	0.39 (0.05)[a]	0.16 (0.04)[a]	0.076
Homework	0.20 (0.04)	0.31 (0.05)	−0.112
Role play	0.25 (0.04)	0.25 (0.05)	0.138
Practice with own child	0.69 (0.08)[a]	0.18 (0.03)[a]	0.234[b]
Separate child instruction	0.21 (0.04)	0.28 (0.04)	0.116
Ancillary services	0.16 (0.03)[a]	0.51 (0.06)[a]	−0.130

[a] Means for "with" versus "without" were significantly different in an inverse-variance-weighted oneway ANOVA ($p<0.05$).
[b] Component's regression weight from inverse-variance-weighted multiple linear regression was significant ($p<0.05$).

effect size for Parent Knowledge or Information Acquisition and relatively smaller effect size for Parenting Behaviors and Skills), multivariate analyses would ideally have been employed to control statistically for those differences (i.e., estimating the association between a component and effect size, controlling for category of outcome measure). However, because studies usually included more than one category of outcome measure, inclusion of a control variable representing outcome category would have violated the independence assumption on which analysis of variance is based. Instead, we selected the two outcome measure categories that were of greatest interest for this meta-analysis, Parenting Behaviors/Skills and Child Externalizing Behaviors, and restricted analyses of variance to effect sizes on these outcomes. Each component was separately entered as the independent variable in a oneway ANOVA with effect size as the dependent variable. Thus, the average effect size of programs reporting the presence of a particular component (e.g., teaching parents to use time out) was compared to the average effect size of programs that did not report including that component.

The next step in identifying program components associated with better outcomes was to conduct a set of multiple linear regressions to determine the robustness of the ANOVA results while controlling for indicators of methodological rigor. Using available SPSS macros (Wilson 2002), effect sizes for Parenting Behaviors/Skills and Child Externalizing Behaviors were regressed on each program component (in separate regressions) with the four rigor variables (random assignment, assessment of initial equivalence, use of a true "no-treatment" control/comparison group, and evaluating the program as a stand-alone intervention) as additional predictors. It is important to note that the set of studies for the Parenting Behaviors/Skills analyses and the set of studies for the Child Externalizing Behavior analyses were not identical. There were 26 studies that reported both Parenting Behaviors/Skills and Child Externalizing Behavior outcomes, an additional 14 that only reported Parenting Behaviors/Skills outcomes, and an additional 22 that only reported Child Externalizing Behavior Outcomes.

Parenting Behaviors and Skills Of the eighteen program components tested, seven were significantly associated with effect sizes on Parenting Behaviors/Skills outcomes in oneway ANOVAs (see Table 6). Three components (Emotional Communication, Consistent Responding, and Practicing with Their Own Child) were each predictive of larger program effects, whereas four components (Problem Solving, Promoting Children's Social Skills, Promoting Children's Cognitive/Academic Skills, and Ancillary Services) were each predictive of smaller program effects. The eighteen components were next regressed separately onto Parenting Behaviors/Skills effect sizes, controlling for the four threats to internal validity. Because demand characteristics or self-reporting bias were likely to have influenced

effect sizes for programs that relied solely on parent's self reports of their parenting behavior, a fifth control variable was included in the regressions for Parenting Behaviors/Skills. This dichotomous variable distinguished between studies that only utilized parents as reporters of their own behaviors and studies that also used other reporters of parents' behaviors (e.g., research assistants).

Six of the 18 program components significantly predicted parenting behavior when methodological rigor and parent self-report were controlled (see Table 6). Three components produced significant positive coefficients, indicating that their presence was reliably associated with more successful programs: Positive Interactions with Child, Emotional Communication, and Practicing with Their Own Child. Three components resulted in significant negative coefficients, indicating that their presence was reliably associated with less successful programs: Problem Solving, Promoting Children's Cognitive/Academic Skills, and Ancillary Services. Although not central to our hypotheses, it is interesting to note that across all regressions on Parenting Behaviors/Skills, only two methodological rigor variables consistently resulted in significant coefficients: the use of a true no-treatment control/comparison condition and using only parent report of outcomes (both associated with larger effect sizes).

Child Externalizing Behavior In oneway ANOVAs, 10 of 18 program components significantly predicted effect sizes on measures of Child Externalizing Behaviors (see Table 7). Six components were predictive of larger program effects: Positive Interactions with Child; Responsiveness, Sensitivity, and Nurturing; Time Out; Problem Solving; Modeling; and Practicing with Own Child. Four were predictive of smaller program effects: Emotional Communication; Promoting Social Skills; Having a Curriculum or Manual; and Ancillary Services. Each of the program components was then used to predict Child Externalizing Behavior outcomes in regression analyses controlling for the four indicators of methodological rigor. Few studies included child reports of their own behaviors, so controlling for self-reporting bias in these analyses was not necessary.

Four components produced significant positive regression coefficients (Positive Interactions with Child, Time Out, Consistent Responding, and Practicing with Own Child) indicating their presence was associated with larger effect sizes (see Table 7). One produced a significant negative coefficient (Promoting Children's Social Skills), suggesting smaller effect sizes for programs with that component. Across all regressions on Child Externalizing Behavior, only two methodological rigor variables consistently resulted in significant coefficients: assessment of initial equivalence and evaluation of the program as part of a package of interventions (both associated with smaller effect sizes).

Test of Robustness and Generalizability

The regression analyses, given the inclusion of at least four control variables (five in the case of parent behavior outcomes), provide a relatively conservative test of the relations between program components and program outcomes. As a further test, we used the maximum likelihood version of Wilson's (2002) SPSS macros to conduct regression analyses based on a mixed-effects model, which incorporates a random effects variance component into the analysis along with the hypothesized fixed effects of the predictor variables (Wilson and Lipsey 2001). This introduction of the random effects variance component into the regression decreases the amount of variance that the program characteristic variables can predict, so variables that continue to be significant predictors in mixed-effects models are more robust than those that are only significant in fixed-effects models (Rosenthal 1995). Controlling for the same indicators of methodological rigor, four components emerged as significant predictors of effect size. For Parenting Behaviors/Skills outcomes, two of the fixed-effects predictors remained significant in the mixed-effects models: Emotional Communication and Practicing with Own Child, both of which were associated with larger effect sizes. Similarly, two of the significant fixed-effects predictors of Child Externalizing Behavior were also significant mixed-effects predictors: Positive Interactions with Child and Consistent Responding, also both associated with larger effect sizes.

Discussion

Consistent with our hypothesis and the results of previous meta-analyses of behavioral parent training (e.g., Lundahl et al. 2006b; Maughan et al. 2005; Reyno and McGrath 2006; Serketich and Dumas 1996), there was a significant, positive overall effect size, supporting the use of such programs in changing parenting behavior and in preventing or ameliorating early child behavior problems. Hypotheses about different outcome constructs were also upheld. The mean effect size for parenting outcomes appeared larger than the mean effect size for child outcomes; the effect size for parenting behaviors and skills were smaller than the effect sizes for parenting knowledge, attitudes, or self-efficacy. Similar patterns have been observed in other meta-analyses (e.g., Lundahl et al. 2006a; Skowron and Reinemann 2005).

For child outcomes, larger effect sizes were observed for internalizing behaviors than for externalizing behaviors and cognitive or educational skills, for which larger effect sizes were observed than for children's social skills or prosocial behavior. The results pertaining to differences in child outcomes were unexpected, and there are few precedents in the literature to help explain these differences. It is possible that internalizing problems (which are potentially more distressful to children than to others in their environment and less likely to be a traditional focus of parenting interventions) are more amenable to change through parent training approaches. Parenting approaches are increasingly being recommended for addressing childhood trauma and other internalizing problems (e.g., Chadwick Center on Children and Families 2004). In contrast, parenting interventions targeting social skills only through a family-based approach may be less effective, as the child might have limited opportunities to practice the skills with peers and others outside the immediate family.

As expected, the results of the regression analyses demonstrated that, independent of study design characteristics, there were identifiable program components and strategies that were associated with larger effect sizes. For both parenting behavior outcomes and child externalizing behavior outcomes, programs that included parent training in creating positive interactions with their child and programs that required parents to practice new skills with their own child during sessions reported significantly larger effect sizes (i.e., the treatment groups exhibited higher levels of parenting behaviors and skills and/or lower levels of child externalizing problems) than programs without those components, regardless of other program content or delivery. Lesser levels of active parental involvement in their own training (modeling, homework, and other types of role playing) were not predictive of program outcomes in the regression analyses. Slightly different patterns were observed for three components, suggesting that different components may be specifically associated with certain outcomes. For parenting behaviors and skills only, programs that included parent training in emotional communication reported significantly larger positive differences between treatment and control/comparison groups at immediate post-test. For child externalizing behaviors only, programs that included parent training in using time-out as a disciplinary technique and responding consistently to children's behaviors reported significantly larger effects.

With respect to three common assumptions in the field, two of the hypotheses were supported. As predicted, teaching parents about child development was not related to program effect sizes. However, knowledge about child development may have been incorporated in other components that were related to effect sizes. For example, in order to teach parents to interact positively with their child or to respond effectively to a 2-year-old's temper tantrum, a program may explain the cognitive, emotional, and behavioral capabilities of typical toddlers. It is possible that components focused solely on learning about child development are less critical to program outcomes than components that translate such knowledge into concrete, developmentally appropriate parenting behaviors and skills.

Also as predicted, providing parents with other ancillary services as part of the parent training program was associated with smaller program effects on parent behaviors and skills outcomes, a result that has been found in other meta-analyses (e.g., Crosby and Perkins 2004; Lundahl et al. 2006b; Serketich and Dumas 1996). The focus on other objectives may divert providers' and parents' attention from the actual acquisition of new parenting behaviors and skills. Although there is strong support in the field for addressing the needs and problems of at-risk families, further research is needed to examine the circumstances (e.g., types of families, services, and problems addressed) and timing (i.e., before, during, or after the parent training program) in which ancillary services might be beneficial. For example, there is evidence to suggest that parental self-efficacy and mental health problems, such as depression, improve with parent training alone (e.g., DeGarmo et al. 2004; Hutchings et al. 2002; McCart et al. 2006; Patterson 1980; Serketich and Dumas 1996; Tonge et al. 2006) albeit maturation effects (i.e., change naturally occurring over time in the absence of an intervention effect) can often not be ruled out.

Contrary to predictions and to commonly held beliefs, reporting the use of a standardized curriculum or manual was not related to effect size in any regression model. The expected relationship between using a manual and effect sizes may be confounded with the quality of the components selected for inclusion in the manual. Because manuals were designed to describe and standardize implementation, Westen et al. (2004) suggested that whether or not a program is manualized may be less important when the goal is to determine which strategies are associated with positive outcomes. In other words, merely having a program manual does not guarantee that the program is effective.

Three components produced significantly negative regression weights, indicating smaller mean differences between treatment and control/comparison groups at post-test. Programs that included parent training in problem solving and promoting children's academic/cognitive skills reported smaller effects on parenting behavior and skills outcomes than programs without those components. Similarly, programs providing parent training in promoting children's social skills reported significantly smaller effects on child externalizing behavior outcomes. These components may be less relevant either to the specific outcomes

targeted in this meta-analysis or to the families of very young children.

The remaining components were not significant in regression models, suggesting that none of those components alone differentiate between more and less effective programs. However, some components might be more effective, or only effective, in the presence of other components. Unfortunately, current statistical tools do not allow for tests of such interaction effects among components in an analogue to linear regression with effect sizes. Future research could be directed to address questions about combinations of components in other ways.

If one assumes the even more conservative test of program characteristics by including a random-effects variance component in a mixed-effects model, four of the identified components and strategies continued to be significant predictors of effect sizes (i.e., predictors of treatment groups with better parenting skills and behaviors and/or lower levels of child externalizing problems): requiring in vivo practice with the parent's own child, teaching skills related to emotional communication, teaching parents to interact positively with their children, and disciplinary consistency. Programs that required in vivo practice with the parent's own child (e.g., the parent directed her/his child through an activity while program facilitators observed) reported larger effects on parenting outcomes. Teaching parents emotional communication skills, which was also predictive of larger parenting effects, included: training in active listening skills, such as reflecting back what the child says; teaching parents to help children identify and deal with emotions; and teaching parents to reduce negative communication (e.g., criticism, sarcasm). Larger effects on children's externalizing behavior were reported by programs that included teaching parents to interact positively with their children in non-disciplinary situations. This component included teaching parents how to demonstrate enthusiasm and positive attention for appropriate behavior, how to interact on the child's level during play and to let the child take the lead during a play activity, and how to provide appropriate recreational activity choices for the child. Finally, inclusion of teaching disciplinary consistency (i.e., that by responding to a particular misbehavior every time it occurs and with the same consequence, the misbehavior will extinguish more quickly) was also predictive of larger effects on child externalizing behaviors. Although other components were robust while accounting for the methodological indicators in fixed-effects models, these four were especially robust in the even more conservative mixed-effects test. Hence, of the program components and strategies included in our analyses, these four represent the greatest stability in predicting significant, positive outcomes for parenting behaviors and skills and child externalizing problems.

Not surprisingly, some of the recurring predictors of larger program effects can be traced to the earliest parent training programs and have continued through divergent iterations of those programs. The two-stage model presented in 1969 by Constance Hanf consisted of the Child's Game (in which mothers were taught to reinforce desirable behaviors with attention) and the Mother's Game (in which mothers were taught to extinguish undesirable behaviors and increase compliance). Interventions that followed the Hanf model (e.g., Barkley 1997, Eyberg 2003; McMahon and Forehand 2003; Webster-Stratton 2000) are characterized by engaging parents in role-play (following instruction and therapist modeling) of a common core of specific child behavior management skills: attending, positive reinforcement, planned ignoring, providing clear instructions, and using time-out from positive reinforcement (McMahon and Forehand 2003). The regression analyses provide further validation of the importance of two of those components in changing parents' and children's behavior. Parent training in attending (positive parent–child interactions) was predictive of both parents' and children's behavioral outcomes, and teaching parents to use time-out was predictive of children's behavioral outcomes. Teaching parents to provide positive reinforcement and clear instructions (i.e., disciplinary communication) were not predictive of program effects. As with teaching child development knowledge, these might be necessary, but not sufficient, for changing the outcomes examined here. Although planned ignoring was a component included in the coding scheme, it unfortunately was reported by too few programs to be included in these analyses.

Interestingly, of the four most robust predictors of outcomes, only one was specifically related to discipline. Two of the four (teaching positive interactions and teaching emotional communication skills) were focused on providing parents with the skills to enhance the overall quality of the parent-child relationship—such as increasing parental activity and positivity with the child, decreasing parental negativity, and teaching the parent to communicate in ways that promotes the child's emotional health and well-being. These findings are consistent with decades of basic developmental theory and research suggesting that the quality of the parent–child relationship has a great deal of influence on a child's behavior or misbehavior (e.g., Brooks 2005; Coie and Dodge 1998). Having parents practice with their own child during parent training sessions (another component of programs that follow Hanf's model) was consistently associated with better parent and child outcomes. This makes intuitive sense. First, it allows parents and children to practice new skills and behaviors with each other. This is a more naturalistic approach than other methods that require parents to generalize skills learned in other settings to their interactions with the children, during

which any number of child behaviors may occur to disrupt the process. Second, this approach allows the parent training facilitator to provide immediate reinforcement and corrective feedback to ensure parents' mastery of the skills. These results are also consistent with the educational literature, which demonstrates that learning in context is more effective (e.g., Hattie et al. 1996).

Several study design features were related to effect sizes, providing additional support to the growing understanding of how methodological choices influence study outcomes. Similar to Wilson and Lipsey (2001, 2003), our results suggest that random assignment is unlikely to be the most important study design variable contributing to different effect sizes. Although a study's use of random assignment was predictive of overall effect size, random assignment was the only methodological variable included in our analyses that never significantly contributed in regression models. The significant difference in the overall effect size might be explained by the greater use of randomized designs by research/demonstration programs (Lipsey 2003; Wilson and Lipsey 2003)—which are more likely to have been standardized, tested, and refined than practice programs. That random assignment was not a uniquely significant predictor in the regression analyses simply indicates that a program's components were stronger or more reliable predictors of program outcomes than whether or not it had been evaluated with a randomized design.

Design and method variables related to type of comparison group, assessment of initial equivalence, assessment of the program as a stand-alone intervention, and reliance on parent self-report reliably accounted for between-study variance in effect size, with some indicators being positively associated and others being negatively associated with larger effect sizes. This is consistent with Wilson and Lipsey's (2003) conclusion that methodological rigor may function more as error than as a consistent bias and suggests that the common practice of using cumulative or subjective indices to assess or control for study quality has limitations. Although some researchers have successfully constructed a priori cumulative indices of study quality that are predictably related to effect sizes (e.g., Lundahl et al. 2006a; McCart et al. 2006), other researchers have failed to find such relationships (e.g., Lundahl et al. 2006b; Skowron and Reinemann 2005). Our results suggest that study quality variables should be examined independently and objectively in meta-analyses.

The present study included evaluations published between 1990 and 2002. Clearly, additional parent training evaluations have been published in the last few years that were not included in the current study. Although parent training components appear to be consistently represented across the time frame of the current study (and in the earlier parenting approaches not included here), more recent innovations in parent training are not represented. In addition, as with any analysis based on published literature, our results thus cannot be considered representative of programs for which evaluation results have not been published. However, the programs included varied widely in their content and delivery components, which allowed examination of the contribution of these components to program effect sizes and testing of the hypotheses. In addition, the intentional inclusion of nonrandomized studies increases the applicability of the results beyond the small set of "research/demonstration" programs that have been subjected to controlled trials to include "practice" programs. Although the generalizability of the results should be interpreted with caution (which is true of any study), we attempted to minimize the likelihood that programs not meeting the publication inclusion criteria would be substantially different from included programs.

As is true in all meta-analyses, the results are correlational. Meta-analysts have no experimental control over the studies they include and must take the field of study "as is." Thus, it would be inappropriate to claim that particular components or strategies caused program success or that the inclusion of other components led to less optimal outcomes. The results speak only to the extent to which certain components were consistently associated with greater differences between treatment and control/comparison groups on parenting behaviors/skills and on child externalizing behavior at immediate post-test across a range of program content, delivery, and evaluation methodologies. The remaining components (i.e., those that did not significantly predict effect size) did not distinguish more successful programs from less successful ones. Although these components may or may not contribute to program outcomes as precursors to or through interactions with other components, they appear less likely to be independently sufficient to ensure program success.

Another limitation pertains to the completeness of reporting within individual studies. For some variables, the extent of missing data was known (e.g., total number of contact hours could not be confidently calculated in 58% of included studies). These missing data limited the ability to conduct moderator analyses of great interest such as the impact of intervention dosage, participant demographic characteristics (such as age or gender of the children or parents), intervention location, and facilitator/provider qualifications on effect sizes. The absence of such information in publications is well established and has been commented on by others (e.g., Borrelli et al. 2005; McCart et al. 2006; Skowron and Reinemann 2005; Westen et al. 2004). For other variables, especially those related to program components and strategies, a lack of mention within the article was coded as a lack of use (e.g., if an article did not report assigning homework to parents, the

program was coded as not including homework). Thus, the results are reliant upon program descriptions that are sufficiently descriptive to be captured in all of the coded variables applicable to that program. Although the use of secondary sources referenced by the study authors attempted to ameliorate this problem, the extent to which program characteristics were not reported is unknown, as are the effects of such under-reporting on the results.

The limitation imposed by incomplete reporting about populations, treatments, and methods warrants repeating in the form of recommendations for the field. Quality of reporting can have significant impact on what can be learned from any particular study. When evaluation reports omit critical information about the population and treatment being studied and the methods used to study them, the audience's ability to draw conclusions about the intervention is limited. Further, without clear descriptions, a reader cannot confidently make inferences about how effective an intervention is likely to be in a different population or using a different evaluation approach. Providing a set of comprehensive guidelines for reporting on a program evaluation is beyond the scope of this paper. Based on the missing information encountered in this meta-analysis, we offer general recommendations for including the following information when reporting evaluations of parent training programs (whether published or unpublished):

- Basic demographic information for participants (i.e., age, gender, and ethnicity, as well as any other notable characteristics, such as diagnosis or selection criteria specifically relevant to a particular treatment)
- Details of recruitment and assignment to treatment or control/comparison conditions
- Details of the intervention, including content covered, delivery method(s) used, intervention setting(s), and the amount and duration of contact time
- Details of program/intervention facilitators' professional and program-specific training
- Details (and results) of treatment fidelity assessments
- Outcome measure information, including the name and source of each measure, the reporter for each measure, and the timing and circumstances of data collection; and
- Attrition information, including the number of participants who dropped out of each condition and how their data were handled in the analyses

More complete recommendations are available elsewhere for reporting the results of randomized clinical trials (Altman et al. 2001; Moher et al. 2001) and other program evaluation designs (Moskowitz 1989).

Statistical reporting quality also impacts how well-represented a study will be within the context of a meta-analysis such as this one. Fortunately, the statistical requirements for inclusion in a meta-analysis are simple: unadjusted means, standard deviations, and sample sizes for the treatment and control/comparison groups. Neither advanced statistical techniques nor software packages are needed—most spreadsheet software programs will compute these statistics. We were thus disheartened that so many studies could not be included because they were missing one or more basic pieces of information. It is also important to note that some of the excluded studies reported relatively high-level outcome analyses, but failed to report the simple group statistics. Thus, even when advanced inferential techniques are employed, reports should also contain basic descriptive statistics.

A final consideration with respect to the obtained results is that the set of programs included those intended to be preventive as well as those intended to treat identified child behavior problems. We had no a priori reason or evidence to suggest that any program components would be more (or less) effective in prevention versus treatment programs. Although children with identified behavior problems (and parents of those children) may require more intensive interventions, the categories of skills as defined in this meta-analysis that parents need to either prevent or remediate behavior problems are likely to be the same. Thus, there was no reason to hypothesize that different components would be differentially effective in prevention or treatment programs. Where differences may arise, however, is that prevention programs may be less likely than treatment programs to show changes in children's externalizing behavior at immediate post-test. This might be because treatment programs would be more likely to offer additional child-directed behavioral training in addition to parent training, or because the lower initial levels of child behavior problems in a true prevention program leave less room for change during the intervention period. Regardless of the cause, this issue underscores the importance of including longer-term follow-up assessments in evaluations of prevention programs, as immediate post-test results will only answer part of the effectiveness question.

This meta-analysis marks a distinct departure from the commonly conducted, "best practice" approach to recommending effective programs. Although best practice recommendations provide useful information for practitioners considering adoption of a packaged program, any particular program may not include the best possible combination of components to produce maximum results. Given the current climate of decreasing resources and increasing accountability for impact, careful attention needs to be paid to optimizing returns on expenditures. Instead of considering each treatment as a black box, the coding scheme allowed for the decomposition of programs into specific content components and the methods used to teach them. Results from the component analyses conducted here can not only help in developing or selecting a parent training program

for use, but can also assist in critically appraising and improving programs already labeled effective or efficacious. Although careful evaluation of modifications or adaptations to existing programs would be critical, by adding, changing, or omitting components associated with larger or smaller effects, effective and efficacious programs are likely to become more potent. Such approaches may also minimize costs associated with adopting and implementing new approaches and increase uptake of more effective parent training components in the field (Barth et al. 2005). Our results do not necessarily imply that programs should be discarded if they contain components that did not contribute positively to effect sizes. However, these results suggest that if the intended outcomes are parenting behaviors and skills and externalizing behaviors in children ages 0–7, resources might best be redirected from strategies consistently associated with smaller effects (problem solving; teaching parents to promote children's cognitive, academic, or social skills; and providing an array of other services) to strategies consistently associated with larger effects, such as increasing positive parent–child interactions and emotional communication, teaching time out and the importance of parenting consistency, and requiring parents to practice new skills with their children during parent training sessions.

References

Adams, C. D., & Kelley, M. L. (1992). Managing sibling aggression: Overcorrection as an alternative to time-out. *Behavior Therapy*, *23*, 707–717.

Albarracín, D., McNatt, P. S., Klein, C. T. F., Ho, R. M., Mitchell, A. L., & Kumkale, G. T. (2003). Persuasive communications to change actions: An analysis of behavioral and cognitive impact in HIV prevention. *Health Psychology*, *22*, 166–177.

Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., et al. (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine*, *134*, 663–694.

Anastopoulos, A. D., Shelton, T. L., DuPaul, G. J., & Guevremont, D. C. (1993). Parent training for Attention-Deficit Hyperactivity Disorder: Its impact on parent functioning. *Journal of Abnormal Child Psychology*, *21*, 581–596.

Arthur Jr., W., Bennett Jr., W., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance*, *11*, 57–101.

Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart, and Winston.

Barber, J. G. (1992). Evaluating parent education groups: Effects on sense of competence and social isolation. *Research on Social Work Practice*, *2*, 28–38.

Barhava-Monteith, G., Harre, N., & Field, J. (1999). A promising start: An evaluation of the HIPPY program in New Zealand. *Early Child Development and Care*, *159*, 145–157.

Barkley, R. S. (1997). *Defiant children* (2nd Ed.) New York: The Guilford.

Barkley, R. A., Shelton, T. L., Crosswait, C., Moorehouse, M., Fletcher, K., Barrett, S., et al. (1996). Preliminary findings of an early intervention program with aggressive hyperactive children. *Annals of the New York Academy of Sciences*, *794*, 277–289.

Barrett, P. M., Dadds, M. R., & Rapee, R. M. (1996). Family treatment of childhood anxiety: A controlled trial. *Journal of Consulting and Clinical Psychology*, *64*, 333–342.

Barth, R. P. (1991). An experimental evaluation of in-home child abuse prevention services. *Child Abuse and Neglect*, *15*, 363–375.

Barth, R. P., Landsverk, J., Chamberlain, P., Reid, J. B., Rolls, J. A., Hurlburt, M. S., et al. (2005). Parent-training programs in child welfare services: Planning for a more evidence-based approach to serving biological parents. *Research on Social Work Practice*, *15*, 353–371.

Bernal, M. E., Klinnert, M. D., & Schultz, L. Q. (1980). Outcome evaluation of behavioral parent training and client-centered parent counseling for children with conduct problems. *Journal of Applied Behavior Analysis*, *13*, 677–691.

Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A. J., Czajkowski, S., Breger, R., et al. (2005). A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of Consulting and Clinical Psychology*, *73*, 852–860.

Bradley, R. H., Whiteside, L., Mundfrom, D. J., Casey, P. H., Caldwell, B. M., & Barrett, K. (1994). Impact of the Infant Health and Development Program (IHDP) on the home environments of infants born prematurely and with low birthweight. *Journal of Educational Psychology*, *86*, 531–541.

Bratton, S., & Landreth, G. (1995). Filial therapy with single parents: Effects on parental acceptance, empathy, and stress. *International Journal of Play Therapy*, *4*, 61–80.

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., et al. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology*, *39*, 222–245.

Brooks, R. B. (2005). The power of parenting. In S. Goldstein, & R. B. Brooks (Eds.) *Handbook of resilience in children* (pp. 297–314). New York: Plenum.

Budd, K. S., & Itzkowitz, J. S. (1990). Parents as social skill trainers and evaluators of children. *Child and Family Behavior Therapy*, *12*, 13–30.

Capaldi, D., DeGarmo, D., Patterson, G. R., & Forgatch, M. (2002). Contextual risk across early life span and association with antisocial behavior. In J. Reid, J. R. Patterson, & J. Snyder (Eds.) *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention* (pp. 123–145). Washington, DC: American Psychological Association.

Cedar, B., & Levant, R. F. (1990). A meta-analysis of the effects of Parent Effectiveness Training. *The American Journal of Family Therapy*, *18*, 373–384.

Chadwick Center on Children and Families (2004). *Closing the quality chasm in child abuse treatment: Identifying and disseminating best practices*. San Diego, CA: Chadwick Center on Children and Families.

Chaffin, M., & Valle, L. A. (2003). Dynamic risk prediction characteristics of the Child Abuse Potential Inventory. *Child Abuse and Neglect*, *27*, 463–481.

Cicchetti, D., Rogosch, F. A., & Toth, S. L. (2000). The efficacy of Toddler–Parent Psychotherapy for fostering cognitive development of offspring in depressed mothers. *Journal of Abnormal Child Psychology*, *28*, 135–148.

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon, & N. Eisenberg (Eds.) *Handbook of child psychology: Vol 3: Social, emotional, and personality development* (5th Ed., pp. 779–862). New York: Wiley.

Conger, R. D., Patterson, G. R., & Ge, X. (1995). It takes two to replicate: A mediational model for the impact of parents' stress on adolescent adjustment. *Child Development*, *66*, 80–97.

Connell, S., Sanders, M. R., & Markie-Dadds, C. (1997). Self-directed behavioral family intervention for parents of oppositional children in rural and remote areas. *Behavior Modification*, *21*, 379–408.

Connolly, L., Sharry, J., & Fitzpatrick, C. (2001). Evaluation of a group treatment programme for parents of children with behavioural disorders. *Child Psychology and Psychiatry Review*, *6*, 159–165.

Costas, M., & Landreth, G. (1999). Filial therapy with nonoffending parents of children who have been sexually abused. *International Journal of Play Therapy*, *8*, 43–66.

Crosby, J., & Perkins, S. (2004). *The effectiveness of behavioral parent training in the treatment of conduct problems in children and adolescents.* Poster session presented at the 38th Annual Convention of the Association for Advancement of Behavior Therapy, November, New Orleans, LA.

Cunningham, C. E., Bremner, R., & Boyle, M. (1995). Large group community-based parenting programs for families of preschoolers at risk for disruptive behavior disorders: Utilization, cost effectiveness, and outcome. *Journal of Child Psychology and Psychiatry*, *36*, 1141–1159.

Davis, H., & Spurr, P. (1998). Parent counseling: An evaluation of a community child mental health service. *Journal of Child Psychology and Psychiatry*, *39*, 365–376.

DeGarmo, D. S., Patterson, G. R., & Forgatch, M. S. (2004). How do outcomes in a specified parent training intervention maintain or wane over time? *Prevention Science*, *5*, 72–89.

Evans, I. M., & Okifuji, A. (1992). Home-school partnerships: A behavioural-community approach to childhood behaviour disorders. *New Zealand Journal of Psychology*, *21*, 14–24.

Evans, I. M., Okifuji, A., Engler, L., Bromley, K., & Tishelman, A. (1993). Home-school communication in the treatment of childhood behavior problems. *Child and Family Behavior Therapy*, *15*, 37–60.

Eyberg, S. M. (2003). Parent–Child Interaction Therapy. In T. H. Ollendick, & C. S. Schroeder (Eds.) *Encyclopedia of clinical child and pediatric psychology* (pp. 446–447). New York: Plenum.

Faith, M. S., Allison, D. B., & Gorman, B. S. (1997). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.) *Design and analysis of single-case research* (pp. 245–277). Mahwah, NJ: Erlbaum.

Farrington, D. P., Joliffe, D., Hawkins, J. D., Catalano, R. F., Hill, K. G., & Kosterman, R. (2003). Comparing delinquency careers in court records and self-reports. *Criminology*, *41*, 933–958.

Fennell, D. C., & Fishel, A. H. (1998). Parent education: An evaluation of STEP on abusive parents' perceptions and abuse potential. *Journal of Child and Adolescent Psychiatric Nursing*, *11*, 107–120.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. (1994). The childhoods of multiple problem adolescents: A 15-year longitudinal study. *Journal of Child Psychology and Psychiatry*, *35*, 1123–1140.

Fewell, R. R., & Wheeden, C. A. (1998). A pilot study of intervention with adolescent mothers and their children: A preliminary examination of child outcomes. *Topics in Early Childhood Special Education*, *18*, 18–25.

Fitzpatrick, P., Molloy, B., & Johnson, Z. (1996). Community mothers' programme: Extension to the travelling community in Ireland. *Journal of Epidemiology and Community Health*, *51*, 299–303.

Forgatch, M. S., & DeGarmo, D. S. (1999). Parenting through change: An effective prevention program for single mothers. *Journal of Consulting and Clinical Psychology*, *67*, 711–724.

Forgatch, M. S., & DeGarmo, D. S. (2002). Extending and testing the social interaction learning model with divorce samples. In J. B. Reid, G. R. Patterson, & J. Snyder (Eds.) *Antisocial behavior in children and adolescents* (pp. 235–257). Washington, DC: American Psychological Association.

Gaudin, J. M., Wodarski, J. S., Arkinson, M. K., & Avery, L. S. (1990). Remedying child neglect: Effectiveness of social network interventions. *The Journal of Applied Social Sciences*, *15*, 97–123.

Gelfand, D. M., Teti, D. M., Seiner, S. A., & Jameson, P. B. (1996). Helping mothers fight depression: Evaluation of a home-based intervention program for depressed mothers and their infants. *Journal of Clinical Child Psychology*, *25*, 406–422.

Glover, G. J., & Landreth, G. L. (2000). Filial therapy with Native Americans on the Flathead Reservation. *International Journal of Play Therapy*, *9*, 57–80.

Greenland, S. (1994). A critical look at some popular meta-analytic methods. *American Journal of Epidemiology*, *140*, 290–296.

Griff, M. D. (1999). Intergenerational play therapy: The influence of grandparents in family systems. *Child and Youth Services*, *20*, 63–76.

Gross, D., Fogg, L., & Tucker, S. (1995). The efficacy of parent training for promoting positive parent–toddler relationships. *Research in Nursing and Health*, *18*, 489–499.

Harris, Z. L., & Landreth, G. (1997). Filial therapy with incarcerated mothers: A five-week model. *International Journal of Play Therapy*, *6*, 53–73.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66*, 99–136.

Hedges, L. V. (1983). A random effects model for effect sizes. *Psychological Bulletin*, *93*, 388–395.

Hedges, L. V. (1994). Fixed effects models. In H. Cooper, & L. V. Hedges (Eds.) *The handbook of research synthesis* (pp. 285–299). New York: Russell Sage Foundation.

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

Horn, W. F., Ialongo, N., Greenberg, G., Packard, T., & Smith-Winberry, C. (1990). Additive effects of behavioral parent training and self-control therapy with Attention-Deficit Hyperactivity disordered children. *Journal of Clinical Child Psychology*, *19*, 98–110.

Hutcheson, J. J., Black, M. M., Talley, M., Dubowitz, H., Howard, J. B., Starr Jr., R. H., et al. (1997). Risk status and home intervention among children with Failure-to-Thrive: Follow-up at age 4. *Journal of Pediatric Psychology*, *22*, 651–668.

Hutchings, J., Appleton, P., Smith, M., Lane, E., & Nash, S. (2002). Evaluation of two treatments for children with severe behaviour problems: Child behaviour and maternal mental health outcomes. *Behavioural and Cognitive Psychotherapy*, *30*, 279–295.

Infant Health and Development Program, (1990). Enhancing the outcomes of low-birth-weight, premature infants: A multisite, randomized trial. *Journal of the American Medical Association*, *263*, 3035–3042.

Jason, L. A., Pokorny, S. B., Kohner, K., & Benetto, L. (1994). An evaluation of the short-term impact of a media-based substance abuse prevention programme. *Journal of Community and Applied Social Psychology*, *4*, 63–69.

Johnson, Z., Howell, F., & Molloy, B. (1993). Community mothers' programme: Randomised controlled trial of non-professional intervention in parenting. *British Medical Journal*, *306*, 1449–1452.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd Ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Kale, A. L., & Landreth, G. L. (1999). Filial therapy with parents of children experiencing learning difficulties. *International Journal of Play Therapy*, *8*, 35–56.

Kissman, K. (1992). Parent skills training: Expanding school-based services for adolescent mothers. *Research on Social Work Practice*, *2*, 161–171.

Koniak-Griffin, D., & Verzemnieks, I. (1991). Effects of nursing intervention on adolescents' maternal role attainment. *Issues in Comprehensive Pediatric Nursing*, *14*, 121–138.

Koniak-Griffin, D., Verzemnieks, I., & Cahill, D. (1992). Using videotape instruction and feedback to improve adolescents' mothering behaviors. *Journal of Adolescent Health*, *13*, 570–575.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Landreth, G. L., & Lobaugh, A. F. (1998). Filial therapy with incarcerated fathers: Effects on parental acceptance of child, parental stress, and child adjustment. *Journal of Counseling and Development*, *76*, 157–165.

Lee, J. H., & Holland, T. P. (1991). Evaluating the effectiveness of foster parent training. *Research on Social Work Practice*, *1*, 162–174.

Letourneau, N., Drummond, J., Fleming, D., Kysela, G., McDonald, L., & Stewart, M. (2001). Supporting parents: Can intervention improve parent–child relationships? *Journal of Family Nursing*, *7*, 159–187.

Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *Annals of the American Academy of Political and Social Science*, *587*, 69–81.

Loeber, R., Farrington, D. P., & Waschbusch, D. A. (1998). Serious and violent juvenile offenders. In R. Loeber, & D. P. Farrington (Eds.) *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 13–29). Thousand Oaks, CA: Sage.

Lundahl, B., & Harris, N. (2006). Delivering parent training to families at risk to abuse: Lessons from three meta-analyses. *The APSAC Advisor*, *18*(3), 7–11.

Lundahl, B. W., Nimer, J., & Parsons, B. (2006a). Preventing child abuse: A meta-analysis of parent training programs. *Research on Social Work Practice*, *16*, 251–262.

Lundahl, B. W., Risser, H. J., & Lovejoy, M. C. (2006b). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, *26*, 86–104.

Lyons-Ruth, K., Connell, D. B., Grunebaum, H. U., & Botein, S. (1990). Infants at social risk: Maternal depression and family support services as mediators of infant development and security of attachment. *Child Development*, *61*, 85–98.

MacKenzie, E. P., & Hilgedick, J. M. (1999). The Computer-Assisted Parenting Program (CAPP): The use of a computerized behavioral parent training program as an educational tool. *Child and Family Behavior Therapy*, *21*, 23–43.

Magen, R. H., & Rose, S. D. (1994). Parents in groups: Problem solving versus behavioral skills training. *Research on Social Work Practice*, *4*, 172–191.

Maguin, E., Zucker, R. A., & Fitzgerald, H. E. (1994). The path to alcohol problems through conduct problems: A family-based approach to very early intervention with risk. *Journal of Research on Adolescence*, *4*, 249–269.

Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review*, *34*, 267–286.

McBride, B. A. (1991a). Parent education and support programs for fathers: Outcome effects on paternal involvement. *Early Child Development and Care*, *67*, 73–85.

McBride, B. A. (1991b). Parental support programs and paternal stress: An exploratory study. *Early Childhood Research Quarterly*, *6*, 137–149.

McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-braining and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, *34*, 527–543.

McMahon, R. J., & Forehand, R. L. (2003). *Helping the noncompliant child: Family-based treatment for oppositional behavior* (2nd Ed.), New York: The Guilford.

McNeil, C. B., Capage, L. C., Bahl, A., & Blanc, H. (1999). Importance of early intervention for disruptive behavior problems: Comparison of treatment and waitlist-control groups. *Early Education and Development*, *10*, 445–454.

McNeil, C. B., Eyberg, S., Eisenstadt, T. H., Newcomb, K., & Funderburk, B. (1991). Parent–Child Interaction Therapy with behavior problem children: Generalization of treatment effects to the school setting. *Journal of Clinical Child Psychology*, *20*, 140–151.

Metropolitan Area Child Study Research Group, (2002). A cognitive–ecological approach to preventing aggression in urban settings: Initial outcomes for high-risk children. *Journal of Consulting and Clinical Psychology*, *70*, 179–294.

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology*, *13*, 355–375.

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology*, *14*, 179–207.

Moher, D., Schulz, K. F., & Altman, D. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association*, *285*, 1987–1991.

Moskowitz, J. M. (1989). Preliminary guidelines for reporting outcome evaluation studies of health promotion and disease prevention programs. *New Directions in Program Evaluation*, *43*, 101–111.

Mullin, E., & Quigley, K. (1994). A controlled evaluation of the impact of a parent training programme on child behaviour and mothers' general well-being. *Counselling Psychology Quarterly*, *7*, 167–179.

Myers, H. F., Alvy, K. T., Arrington, A., Richardson, M. A., Marigna, M., Huff, R., et al. (1992). The impact of a parent training program on inner-city African-American families. *Journal of Community Psychology*, *20*, 132–147.

Nelson, W. P., & Levant, R. F. (1991). An evaluation of a skills training program for parents in stepfamilies. *Family Relations*, *40*, 291–296.

Nicholson, B., Anderson, M., Fox, R., & Brenner, V. (2002). One family at a time: A prevention program for at-risk parents. *Journal of Counseling and Development*, *80*, 362–371.

Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2003). Parent–Child Interaction Therapy: A comparison of standard and abbreviated treatments for oppositional defiant children. *Journal of Consulting and Clinical Psychology*, *71*, 251–260.

Odom, S. E. (1996). Effects of an educational intervention on mothers of male children with Attention Deficit Hyperactivity Disorder. *Journal of Community Health Nursing*, *13*, 207–220.

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, *8*, 157–159.

Patterson, G. R. (1980). Mothers: The unacknowledged victims. *Monographs of the Society for Research in Child Development*, *45*(5), 1–54 (Serial No. 186).

Patterson, G. R., DeGarmo, D., & Forgatch, M. S. (2004). Systematic changes in families following prevention trials. *Journal of Abnormal Child Psychology*, *32*, 621–633.

Pehrson, K. L., & Robinson, C. C. (1990). Parent education: Does it make a difference? *Child Study Journal*, *20*, 221–236.

Pisterman, S., Firestone, P., McGrath, P., Goodman, J. T., Webster, I., Mallory, R., et al. (1992a). The effects of parent training on

parenting stress and sense of competence. *Canadian Journal of Behavioural Science*, *24*, 41–58.

Pisterman, S., Firestone, P., McGrath, P., Goodman, J. T., Webster, I., Mallory, R., et al. (1992b). The role of parent training in treatment of preschoolers with ADDH. *American Journal of Orthopsychiatry*, *62*, 397–408.

Raudenbusch, S. W. (1994). Random effects models. In H. Cooper, & L. V. Hedges (Eds.) *The handbook of research synthesis* (pp. 301–336). New York: Russell Sage Foundation.

Reid, M. J., Walter, A. L., & O'Leary, S. G. (1999). Treatment of young children's bedtime refusal and nighttime wakings: A comparison of "standard" and graduated ignoring procedures. *Journal of Abnormal Child Psychology*, *27*, 5–16.

Reifsnider, E. (1998). Reversing growth deficiency in children: The effect of a community-based intervention. *Journal of Pediatric Health Care*, *12*, 305–312.

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems—A meta-analytic review. *Journal of Child Psychology and Psychiatry*, *47*, 99–111.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*, 183–192.

Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, *52*, 471–499.

Sanders, M. R., Markie-Dadds, C., Tully, L. A., & Bor, W. (2000a). The Triple-P Positive Parenting Program: A comparison of enhanced, standard, and self-directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology*, *68*, 624–640.

Sanders, M. R., Montgomery, D. T., & Brechman-Toussaint, M. L. (2000b). The mass media and the prevention of child behavior problems: The evaluation of a television series to promote positive outcomes for parents and their children. *Journal of Child Psychology and Psychiatry*, *41*, 939–948.

Schuhmann, E. M., Foote, R. C., Eyberg, S. M., Boggs, S. R., & Algina, J. (1998). Efficacy of Parent–Child Interaction Therapy: Interim report of a randomized trial with short-term maintenance. *Journal of Clinical Child Psychology*, *27*, 34–45.

Schuhler, M. E., Nair, P., & Black, M. M. (2002). Ongoing maternal drug use, parenting attitudes, and a home intervention: Effects on mother–child interaction at 18 months. *Developmental and Behavioral Pediatrics*, *23*, 87–94.

Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy*, *27*, 171–186.

Shadish, W. R., Robinson, L., & Lu, C. (1999). *ES: A computer program for effect size calculation*. St. Paul, MN: Assessment Systems.

Skowron, E., & Reinemann, D. H. S. (2005). Effectiveness of psychological interventions for child maltreatment: A meta-analysis. *Psychotherapy: Theory, Research, Practice, Training*, *42*, 52–71.

Sonuga-Barke, E. J. S., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: A randomized controlled trial with a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 402–408.

Stock, W. A. (1994). Systematic coding for research synthesis. In H. Cooper, & L. V. Hedges (Eds.) *The handbook of research synthesis* (pp. 125–138). New York: Russell Sage Foundation.

Swanson, H. L., & Hoskyn, M. (2001). A meta-analysis of intervention research for adolescent students with learning disabilities. *Learning Disabilities Research & Practice*, *16*, 109–119.

Sweet, M. A., & Appelbaum, M. I. (2004). Is home visiting an effective strategy? A meta-analytic review of home visiting programs for families with young children. *Child Development*, *75*, 1435–1456.

Taylor, T. K., Schmidt, F., Pepler, D., & Hodgins, C. (1998). A comparison of eclectic treatment with Webster-Stratton's Parents and Children Series in a children's mental health center: A randomized controlled trial. *Behavior Therapy*, *29*, 221–240.

Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral outcomes of Parent–Child Interaction Therapy and Triple P-Positive Parenting Program: A review and meta-analysis. *Journal of Abnormal Child Psychology*, *35*, 475–495.

Thompson, R. W., Ruma, P. R., Schuchmann, L. F., & Burke, R. V. (1996). A cost-effectiveness evaluation of parent training. *Journal of Child and Family Studies*, *5*, 415–429.

Tolan, P. H., & Gorman-Smith, D. (1998). Development of serious and violent offending careers. In R. Loeber, & D. P. Farrington (Eds.) *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 68–85). Thousand Oaks, CA: Sage.

Tolliver, R. M., Valle, L. A., Dopke, C. A., Serra, L. D., & Milner, J. S. (1998). Child physical abuse. In N. A. Jackson, & G. C. Oates (Eds.) *Violence in intimate relationships* (pp. 1–23). Boston: Butterworth-Heinemann.

Tonge, B., Brereton, A., Kiomall, M., Mackinnon, A., King, N., & Rinehart, N. (2006). Effects on parental mental health of an education and skills training program for parents of young children with autism: A randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 561–569.

Thornberry, T. P., Lizotte, A. J., Krohn, M. D., Smith, C. A., & Porter, P. K. (2003). Causes and consequences of delinquency: Findings from the Rochester Youth Development Study. In T. P. Thornberry, & M. D. Krohn (Eds.) *Taking stock of delinquency: An overview of findings from contemporary longitudinal studies* (pp. 11–46). New York: Plenum.

Tremblay, R. E. (2006). Prevention of youth violence: Why not start at the beginning? *Journal of Abnormal Child Psychology*, *34*, 481–487.

Vines, S. W., & Williams-Burgess, C. (1994). Effects of a community health nursing parent–baby (ad)venture program on depression and other selected maternal–child health outcomes. *Public Health Nursing*, *11*, 188–195.

Wahler, R. G., & Meginnis, K. L. (1997). Strengthening child compliance through positive parenting practices: What works? *Journal of Clinical Child Psychology*, *26*, 433–440.

Walker, H. M., Kavanagh, K., Stiller, B., Golly, A., Severson, H. H., & Feil, E. G. (1998). First Step to Success: An early intervention approach for preventing school antisocial behavior. *Journal of Emotional and Behavioral Disorders*, *6*, 66–80.

Wampold, B. E. (2001). *The great psychotherapy debate*. Mahwah, NJ: Erlbaum.

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *132*, 249–268.

Webster-Stratton, C. (1990). Enhancing the effectiveness of self-administered videotape parent training for families with conduct-problem children. *Journal of Abnormal Child Psychology*, *18*, 479–492.

Webster-Stratton, C. (1992). Individually administered videotape parent training: "Who benefits?". *Cognitive Therapy and Research*, *16*, 31–35.

Webster-Stratton, C. (1998). Preventing conduct problems in head start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, *66*, 715–730.

Webster-Stratton, C. (2000). *The Incredible Years training series bulletin*. Washington, DC: US Department of Justice, Office of Juvenile Justice and Delinquency Prevention.

Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent

training interventions. *Journal of Consulting and Clinical Psychology*, *65*, 93–109.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in Head Start. *Journal of Clinical Child Psychology*, *30*, 283–302.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, *130*, 631–663.

Wilczak, G. L., & Markstrom, C. A. (1999). The effects of parent education on parental locus of control and satisfaction of incarcerated fathers. *International Journal of Offender Therapy and Comparative Criminology*, *43*, 90–102.

Wilson, D. B. (2002). *SPSS macros*. Retrieved July 8, 2005, from http://mason.gmu.edu/~dwilsonb/ma.html.

Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Bulletin*, *6*, 413–429.

Wilson, D. B., & Lipsey, M. W. (2003). The role of method in treatment effectiveness research: Evidence from meta-analysis. In A. E. Kazdin (Ed.)*Methodological issues and strategies in clinical research* (3rd Ed., pp. 589–615). Washington, DC: American Psychological Association.