Interparental Agreement on Internalizing, Externalizing, and Total Behavior Problems: A Meta-analysis

Amy M. Duhig, Kimberly Renk, Monica K. Epstein, and Vicky Phares
University of South Florida

Previous studies have addressed the degree of correspondence between interparental reports of children's behavior problems, but have not examined the discrepancies in these reports. A meta-analysis containing 60 studies and 126 independent effect sizes was conducted. Results suggest that maternal and paternal ratings exhibit moderate correspondence in ratings of internalizing behavior problems in children and large correspondence in ratings of externalizing and total behavior problems in children. In terms of discrepancy of reports, parents reported similar levels of all types of problems. Age, gender, and socioeconomic status were found to moderate correspondence between mothers' and fathers' ratings but did not moderate discrepancies in mothers' and fathers' ratings.

Key words: **children's emotional and behavioral problems, fathers, interparental agreement, meta-analytic review, mothers.** ***[Clin Psychol Sci Prac 7:435–453, 2000]***

A crucial issue in the study of psychopathology in children and adolescents concerns whether or not there is an optimal informant in rating the emotional and behavioral problems of children and adolescents. Different informants, such as mothers, fathers, children, teachers, and peers, have been found to differ in the information they provide for child and adolescent assessments (Achenbach, McConaughy, & Howell, 1987). However, some have suggested that parents are the most important source of information for clinicians in diagnosing children's emotional and behavioral problems (Achenbach et al., 1987; Kadzin, 1988). Still, there is no "gold standard" regarding who is the best informant of child functioning. The information received from different informants is thought of as unique and important in conceptualizing the difficulties of children and adolescents.

Low to moderate agreement among raters has been handled in several ways. It is sometimes assumed that the rater reporting a given problem is accurate and that the informant who does not endorse the same problem is inaccurate (e.g., Loeber, Green, Lahey, & Stouthamer-Loeber, 1989). Although understandable, this approach neglects one of the main constraints on the information received by informants, that of situational specificity. That is, those informants who interact with the child or adolescent in the same environment (e.g., mothers and fathers) tend to have higher agreement in their reports of functioning than those who interact with the child or adolescent in different environments (e.g., teachers and parents; Achenbach et al., 1987). Additionally, reports from different informants may have different antecedents and different predictive utility in predicting later developmental outcomes (e.g., Stanger, McConaughy, & Achenbach, 1992). Accordingly, the importance of multiaxial assessment of children and adolescents has been emphasized (Achenbach, 1985).

A comprehensive meta-analysis on the cross-informant ratings of emotional and behavioral problems in children and adolescents was conducted by Achenbach and colleagues (1987). Utilizing Pearson's correlation coefficients as effect sizes, they examined multiple informants' reports of these problems in children and adolescents ranging in age from 6 to 18 years. Overall, when examining interparental agreement (e.g., mothers' and fathers' agreement in

Address correspondence to Vicky Phares, Ph.D., University of South Florida, Department of Psychology, 4202 E. Fowler Avenue, BEH 339, Tampa, FL 33620. Electronic mail may be sent to phares@luna.cas.usf.edu.

rating problematic behavior), they found that mothers' and fathers' reports of behavior correlated moderately but significantly. Levels of agreement were found to be similar both for clinical and nonclinical samples and for boys and girls, but varied depending on the type of problem exhibited by the target child and the age of the target child. In particular, Achenbach and colleagues (1987) found that interparental agreement did not differ for externalizing problems and internalizing problems. Additionally, when comparing younger children (ages 6 to 11) and adolescents (ages 12 to 18), there was higher mother-father agreement for the younger age group.

Individual studies examining interparental agreement have found conflicting patterns of results. When examining interparental agreement at the specific behavior or item level, studies have found that interparental agreement is low (Christensen, Margolin, & Sullaway, 1992). Some studies reported that mothers tended to report more behavior problems in children than fathers (Christensen et al., 1992; Jensen, Taylor, Xenakis, & Davis, 1988; Thurber & Osborn, 1993). Others have found that mothers and fathers reported similar numbers of problems in their children (Achenbach, 1991; Stanger & Lewis, 1993). Some studies have reported a gender of parent by gender of child interaction, with mothers reporting more problems for sons and fathers reporting more problems for daughters (Friedlander, Weiss, & Traylor, 1986; Graham & Stevenson, 1985; Jensen et al., 1988). Other studies have failed to find such an interaction (Achenbach, Howell, Quay, & Connors, 1991; Stanger & Lewis, 1993). Regarding gender of the adolescent, both mothers and fathers tended to rate boys as having more internalizing and externalizing problems than girls (Thurber & Osborn, 1993). This pattern of rating boys as having more problems than girls, however, does not appear to manifest itself in different levels of interparental agreement (e.g., correlation coefficients) for boys and girls (Achenbach et al., 1987; Christensen et al., 1992; Thurber & Osborn, 1993).

Meta-analytic studies of ratings have yet to include children under 6 years of age. Individual studies, however, have examined this issue to some extent. For example, Achenbach (1992) found that when examining behavior problems of 2- to 3-year-old children, interparental agreement was found to be moderate for this age group. He also found that mothers tended to report more behavior problems than fathers for 2-year-old children. For 3-year-old children, however, parents did not differ significantly in their reports of behavior problems. At these younger ages, no effects of gender of the child were found for overall interparental agreement or for the number of behavior problems endorsed by mothers and fathers (Achenbach, 1992; Walker & Bracken, 1996). This trend appears to hold for preschool children as well (Walker & Bracken, 1996). For preschool children, correspondence between parents for internalizing problems, externalizing problems, and total problems have been found to be similar (Eisenstadt, McElreath, Eyberg, & McNeil, 1994; Walker & Bracken, 1996; Zahn-Waxler, Schmitz, Fulker, Robinson, & Emde, 1996). This pattern, however, may depend on the type of instrument used to examine the behaviors (Walker & Bracken, 1996). In another study of parental agreement in preschoolers, Baker and Heller (1996) found that parents had a higher rate of agreement for externalizing problems than for internalizing problems. Thus, the findings appear to be somewhat mixed.

Other potential moderators of interparental agreement are socioeconomic status and ethnicity, but very little work has been conducted in this area. Small discrepancies between mothers' and fathers' ratings of behavior problems in children and adolescents have been found when taking into account socioeconomic differences (Achenbach, 1991). To our knowledge, however, no studies have examined the effect of ethnicity on interparental agreement. Thus, regarding socioeconomic status, and more so ethnicity, it is unclear as to what effects these variables may have on interparental agreement.

Although studies have examined the *correspondence* (i.e., correlations) of maternal and paternal reports of emotional and behavioral problems in children and adolescents (i.e., Do maternal and paternal ratings correlate? Cf. Achenbach et al., 1987; Kolko & Kadzin, 1993; Stanger & Lewis, 1993), the *discrepancy* (i.e., mean differences) of these reports (i.e., Do maternal and paternal ratings differ in the severity of problems reported? Cf. Verhulst & van der Ende, 1992) has yet to be examined in meta-analytic studies. Interparental agreement can mean something different depending on the two types of questions asked. Further, findings from the aforementioned studies emphasize the importance of taking the gender of the parent, the gender of the child or adolescent, ethnicity, socioeconomic status, and treatment status into account when comparing the perspectives of mothers and fathers.

The current meta-analytic study adds considerable

information to that conducted by Achenbach and colleagues (1987), as well as to the individual studies conducted in this area. First, instead of focusing on multiple cross-informants, this meta-analysis focused solely on the contribution of mothers and fathers in rating emotional and behavioral problems in children and adolescents. Although most researchers and clinicians include mothers in their studies and clinical practices, fathers have been relatively neglected in these settings (Phares, 1992; Phares & Compas, 1992). Mothers and fathers, in addition to teachers and children themselves, provide valuable information about emotional and behavioral problems in children and adolescents. Even though a large number of children and adolescents do not live with both parents, a majority do have contact with both parents, and as such, both parents could serve as important sources of information for diagnostic purposes (Phares, 1996). Second, the current meta-analytic study incorporates measures for the correspondence as well as the discrepancy of maternal and paternal ratings of emotional and behavioral problems in children and adolescents. In addition, potential moderators of maternal and paternal ratings are incorporated into this meta-analytic study. Finally, the current meta-analytic study includes studies conducted after the meta-analysis of Achenbach and colleagues (1987) was conducted, providing a review of research conducted more recently than that of the previous meta-analysis.

THE PRESENT STUDY

In the present study, meta-analytic techniques were used to examine the correspondence (i.e., correlations) and discrepancy (i.e., mean differences) of interparental agreement on internalizing, externalizing, and total behavior problems in children and adolescents. As mentioned above, in their meta-analysis of multiple informants' ratings of emotional and behavioral problems in children and adolescents, Achenbach and colleagues (1987) examined correspondence between informants by looking at correlation coefficients as effect sizes. Findings about the level of correspondence (i.e., correlation coefficients) do not fully describe the discrepancy of maternal and paternal ratings. That is, interparental agreement on a dimension may be moderate, as found in a previous meta-analysis (i.e., Achenbach et al., 1987), but one does not know the pattern of these ratings (e.g., which parent is rating fewer or greater numbers of emotional and behavioral problems) from utilizing this information. Additionally, if there was high correspondence between maternal and paternal ratings, the information received from the two informants may not necessarily be similar. Along these lines, the second purpose of the present study was to examine discrepancy between maternal and paternal ratings of internalizing, externalizing, and total behavior problems in children and adolescents. In the present study, we combined studies to examine the level of correspondence and discrepancy between maternal and paternal ratings using meta-analytic procedures suggested by Hedges (1982), Hedges and Olkin (1985), and Rosenthal (1991).

Although independent studies and Achenbach and colleagues' (1987) meta-analytic study have examined these components of interparental agreement, no studies to date have examined specific moderators that may be related to variation in interparental agreement for internalizing, externalizing, and total behavior problems in children and adolescents. To investigate these issues, effect sizes were examined for different characteristics of children and adolescents (e.g., age, gender, ethnicity, socioeconomic status, and belonging to a special population) participating in research on emotional and behavioral problems. The ratings of internalizing, externalizing, and more global ratings of emotional and behavioral problems were examined separately.

METHOD

Retrieval of Studies

We conducted searches for studies published between 1990 and 1997 that contained quantitative data from both mothers and fathers on the emotional and behavioral problems of children and adolescents. Our computer-based information search was conducted using PsychLit and PsychINFO with the following key words and/or phrases: mothers and fathers (2,895 abstracts), maternal and paternal (579 abstracts), parent(s) (5,000 abstracts), parent ratings (823 abstracts), and cross-informant ratings (558 abstracts). In addition, relevant journals published between 1990 and 1997 (*Journal of Consulting and Clinical Psychology, Journal of Abnormal Psychology, Psychological Assessment, Child Development, Developmental Psychology, Journal of Abnormal Child Psychology, Journal of Clinical Child Psychology, Journal of the American Academy of Child and Adolescent Psychiatry, Journal of Child Psychology and Psychiatry, Development and Psychopathology,* and *Journal of Family Psychology*) were perused to locate studies not identified in

our computerized review of abstracts. One hundred sixty-six studies were found to have included both maternal and paternal reports of the emotional and/or behavioral functioning of children and adolescents during the data collection process.

Inclusion Criteria

Studies were included in the meta-analysis if they contained a measure of children's emotional and/or behavioral functioning completed by both their mother and father. Studies involving clinical and medically challenged groups were included as long as there was information on the emotional and behavioral functioning of the children. Results had to be reported in sufficient detail to permit calculation of an effect size, using either *r* or *g*. Studies reporting an effect size of *r* (correlation coefficient) or *g* (converted from data reported as means and standard deviations for both parents) were analyzed separately. Studies indicating only that results were not statistically significant or statistically significant were excluded because insufficient data were provided and *r* or *g* values could not be computed. Data collected from studies used in the meta-analysis relied exclusively on questionnaire measures where the psychometric properties were known and sound. The majority of the data were collected in these studies using the Child Behavior Checklist (CBCL; Achenbach, 1991). Other measures used in data collection included the Eyberg Child Behavior Inventory (ECBI; Robinson, Eyberg, & Ross, 1980), the Revised Behavior Problem Checklist (RBPC; Quay & Peterson, 1983), and the Children's Depression Inventory (CDI; Kovacs, 1994). Reliabilities for these measures ranged from .70 to .98. All studies were published in English-language journals.

Types of problems were divided into internalizing and externalizing problem groups. This is a distinction that is frequently found in the empirical derivation of syndromes of children's emotional and behavioral problems and is sometimes referred to as overcontrolled versus undercontrolled problems (see Achenbach & Edelbrock, 1978). Beside scales with internalizing or externalizing labels, the internalizing problems included scales that measured withdrawn, anxious, depressed, and fearful behavior, while externalizing problems included scales designated as hyperactive, delinquency, antisocial, and aggressive. Included in the total behavior problems group were scales with such a label, as well as scales that combined internalizing, externalizing, and other types of problems to compute a global indicator of emotional and behavioral problems.

Study Sample

Of the 166 studies identified in our initial retrieval search, 106 studies were rejected because sufficient data were not reported (e.g., the researchers merely reported that ratings were "statistically significant" or "not statistically significant" and did not report correlations or means and standard deviations) for both mothers and fathers to allow for calculation of an effect size. Of the 60 studies meeting criteria and providing specific data, 44 studies yielding 71 independent effect sizes (internalizing = 18; externalizing = 22; total behavior = 17; temperament = 8; social behavior = 6) were analyzed using *r* as the effect size estimator. Where *t* values were reported (two studies), conversion equations provided by Rosenthal (1991) were used to convert *t*s to *r*s. Sixteen studies contained 55 independent mean and standard deviation pairs for both mothers and fathers (internalizing = 15; externalizing = 13; total behavior = 27). Effect sizes for these studies were analyzed using Hedges *g* (1982) as the effect size estimator. Summaries of these studies are provided in Tables 1 and 2, respectively. Due to the small number of studies available for the temperament and social behavior categories, these studies were not included in the moderator analyses.[1] In studies where multiple samples were analyzed separately (boys vs. girls), effect sizes from each sample were treated independently.

Plan for Data Analyses

Both a mean effect size and a weighted mean effect size were calculated separately for studies containing a correlation (*r*) between mothers and fathers and for studies containing means for both mothers and fathers by using *g*. In the calculation of mean effect sizes with *r,* Fisher's formula for *r* to *z* transformation was used to adjust for population values as suggested by Rosenthal (1991). The degrees of freedom ($n - 3$) provided in each study were used in the calculation of the mean weighted effect size. In calculating mean effect sizes with *g,* procedures provided by Hedges (1982) were followed.

Categorical model testing was used to determine the homogeneity of our sample studies and to identify potential moderator variables that may explain inconsistencies in reporting the emotional and behavioral problems of

Table 1. Summary of studies reporting correlations of maternal and paternal ratings

Study	*r*	*d*	95% Confidence Interval Lower	Upper	*p*	Categorical Variables[a]
Internalizing behavior problems						
Asher & Wakefield (1990)	.20	.41	−.07	.88	.12	9-3-3-4
Baker & Heller (1996)	.12	.24	−.33	.81	.29	1-3-3-4
Banez & Compas (1990)	.66	1.76	1.29	2.23	.00	2-3-1-4
Eiser, Havermans, Pancer, & Eiser (1992)	.01	.02	−.17	.21	.44	2-3-9-9
Jaycox & Repetti (1993)	.51	1.19	.79	1.58	.00	2-3-3-4
Jouriles, Norwood, McDonald, Vincent, & Mahoney (1996)	.34	.72	.34	1.11	.01	2-3-1-9
O'Brien, Margolin, & John (1995)	.45	1.01	.68	1.33	.00	2-3-3-9
Rosenbaum & Ronen (1997)	.56	1.35	1.06	1.64	.00	9-3-4-4
Rosenberg, Brown, & Singer (1994)	.50	1.15	.56	1.75	.01	3-3-9-9
Rothbaum, Rosen, Pott, & Beatty (1995)	.50	1.15	.59	1.72	.00	9-3-1-4
Sawyer, Baghurst, & Clark (1992)	.66	1.76	1.58	1.94	.00	9-3-9-4
Stanger & Lewis (1993)	.33	.70	.41	.99	.00	3-3-1-4
Sternberg et al. (1993)	.24	.49	.09	.90	.05	2-3-5-1
Thurber & Osborn (1993)	.48	1.09	.70	1.49	.00	3-2-3-2
	.62	1.58	1.11	2.05	.00	3-1-3-2
Walker & Bracken (1996)	.74	2.20	1.58	2.82	.00	2-3-3-9
Westerman & Schonholtz (1993)	.07	.14	−.55	.83	.40	2-3-3-4
Zahn-Waxler, Schmitz, Fulker, Robinson, & Emde (1996)	.28	.58	.41	.75	.00	9-3-3-9
Externalizing behavior problems						
Armistead, Klein, Forehand, & Wierson (1997)	.50	1.15	.79	1.52	.00	2-9-3-9
Baker & Heller (1996)	.55	1.32	.69	1.94	.00	1-3-3-4
Campbell, Pierce, Moore, Marakovitz, & Newby (1996)	.55	1.32	.98	1.65	.00	2-1-1-2
Deater-Deckard & Scarr (1996)	.47	1.06	.94	1.19	.00	1-3-3-4
Henggeler, Watson, & Whelen (1990)	.59	1.46	1.09	1.83	.00	3-3-3-4
Jaycox & Repetti (1993)	.65	1.71	1.29	2.13	.00	2-3-3-4
Jouriles, Norwood, McDonald, Vincent, & Mahoney (1996)	.47	1.06	.67	1.46	.00	2-3-1-9
Kearney & Silverman (1993)	.59	1.46	.97	1.95	.00	2-3-3-2
Mann & MacKenzie (1996)	.70	1.96	1.48	2.44	.00	2-1-3-4
O'Brien, Margolin, & John (1995)	.49	1.12	.80	1.45	.00	2-3-3-9
Rosenberg, Brown, & Singer (1994)	.60	1.50	.87	2.13	.00	3-3-9-9
Rothbaum, Rosen, Pott, & Beatty (1995)	.68	1.85	1.23	2.48	.00	9-3-1-4
Sawyer, Baghurst, & Clark (1992)	.74	2.20	2.01	2.39	.00	9-3-9-4
Smith & Jenkins (1991)	.30	.63	.35	.91	.00	9-3-9-9
Stanger & Lewis (1993)	.50	1.15	.85	1.46	.00	3-3-1-4
Sternberg et al. (1993)	.30	.63	.22	1.03	.02	2-3-5-1
Stice & Barrera (1995)	.65	1.71	1.56	1.86	.00	3-3-3-9
Thurber & Osborn (1993)	.70	1.96	1.51	2.41	.00	3-2-3-2
	.70	1.96	1.46	2.46	.00	3-1-3-2
Walker & Bracken (1996)	.74	2.20	1.58	2.82	.00	2-3-3-9
Westerman & Schonholtz (1993)	.62	1.58	.79	2.37	.01	2-3-3-4
Zahn-Waxler, Schmitz, Fulker, Robinson, & Emde (1996)	.32	.68	.50	.85	.00	9-3-3-9
Total behavior problems						
Coyne (1996)	.72	2.08	1.52	2.63	.00	2-3-9-4
Eisenberg, Fabes, Murphy, Maszk, Smith, & Karbon (1995)	.50	1.15	.78	1.53	.00	2-3-3-2
Eisenstadt, McElreath, Eyberg, & McNeil (1994)	.61	1.54	1.06	2.02	.00	2-3-3-9
Jenkins & Smith (1991)	.36	.77	.51	1.04	.00	9-3-9-9
Kitzmann & Emery (1994)	.44	.98	.17	1.79	.07	2-3-3-9
	.20	.41	−.53	1.34	.30	2-3-3-9
Phares & Compas (1990)	.54	1.28	.84	1.72	.00	3-3-1-4
Radovanoic (1993)	.19	.39	−.03	.81	.11	2-3-9-4
Rosenberg, Brown, & Singer (1994)	.58	1.42	.80	2.04	.00	3-3-9-9
Rothbaum, Rosen, Pott, & Beatty (1995)	.57	1.39	.80	1.97	.00	9-3-1-4
Sawyer, Baghurst, & Clark (1992)	.73	2.14	1.95	2.33	.00	9-3-9-4
Sloper, Knussen, Turner, & Cunningham (1991)	.77	2.41	2.02	2.80	.00	2-3-9-4
Smith & Jenkins (1991)	.36	.77	.49	1.06	.00	9-3-9-9
Thurber & Osborn (1993)	.53	1.25	.85	1.65	.00	3-2-3-2
	.67	1.81	1.32	2.29	.00	3-1-3-2
Walker & Bracken (1996)	.75	2.27	1.64	2.90	.00	2-3-3-9
Watson, Henggeler, & Whelan (1990)	.33	.70	.20	1.20	.03	3-3-3-4

[a] Refer to Table 3 for key to categorical variables.

Table 2. Summary of studies reporting maternal and paternal means

Study	95% Confidence Interval				Categorical Variables[a]
	g	Lower	Upper	*p*	
Internalizing behavior problems					
Brody, Stoneman, Flor, McCrary, Hastings, & Conyers (1994)	.26	−.04	.55	.04	9-3-2-1-2
Cohen, Coyne, & Duvall (1993)	.24	−.34	.82	.21	3-3-9-9-1
	.14	−.49	.77	.33	2-3-9-9-1
	.12	−.45	.69	.34	2-3-9-9-2
	−.03	−.65	.59	.46	2-3-9-9-2
Crockenberg & Lourie (1996)	−.02	−.75	.70	.47	2-1-3-9-2
	−.54	−1.19	.12	.05	2-2-3-9-2
Fagot (1995)	.24	.01	.48	.02	1-3-1-4-2
Kazak, Christakis, Alderfer, & Coiro (1994)	.17	−.18	.52	.17	3-3-3-9-1
Reeve, Bernstein, & Christenson (1992)	.89	−.32	2.09	.08	2-3-3-4-1
Sawyer, Baghurst, & Mathias (1992)	−.08	−.51	.35	.36	3-1-9-4-2
	.44	.00	.87	.03	3-2-9-4-2
	.18	−.16	.51	.15	3-1-9-4-1
	−.10	−.59	.39	.35	3-2-9-4-1
Steele, Forehand, & Armistead (1997)	.14	−.19	.48	.20	3-3-3-9-2
Externalizing behavior problems					
Cohen, Coyne, & Duvall (1993)	.51	−.08	1.09	.05	3-3-9-9-1
	.04	−.58	.67	.45	2-3-9-9-1
	−.05	−.62	.51	.43	2-3-9-9-2
	−.12	−.74	.50	.35	2-3-9-9-2
Crockenberg & Lourie (1996)	.03	−.69	.75	.47	2-1-3-9-2
	−.61	−1.27	.05	.04	2-2-3-9-2
Fagot (1995)	.04	−.20	.27	.38	1-3-1-4-2
Kazak, Christakis, Alderef, & Coiro (1994)	.18	−.17	.53	.16	3-3-3-4-1
Reeve, Bernstein, & Christenson (1992)	.33	−.83	1.50	.29	2-3-3-4-1
Sawyer, Baghurst, & Mathias (1992)	.28	−.15	.72	.10	3-1-9-4-2
	.27	−.16	.70	.11	3-2-9-4-2
	.11	−.23	.45	.26	3-1-9-4-1
	−.17	−.66	.32	.25	3-2-9-4-1
Total behavior problems					
Abidin, Jenkins, & McGaughey (1992)	.06	−.38	.51	.39	9-1-1-2-2
	.04	−.37	.45	.43	9-2-1-2-2
Cohen, Coyne, & Duval (1993)	.45	−.13	1.04	.07	3-3-9-9-1
	.15	−.48	.78	.32	2-3-9-9-1
	.02	−.55	.58	.48	2-3-9-9-2
	−.14	−.76	.48	.33	2-3-9-9-2
Cuskelly & Dadds (1992)	−.03	−.96	.89	.47	9-2-9-9-1
	.41	−.40	1.22	.16	9-1-9-9-1
	.14	−.66	.94	.37	9-2-9-9-2
	.47	−.46	1.41	.16	9-1-9-9-2
Jensen, Bloedau, DeGroot, Ussery, & Davis (1990)	.24	.00	.48	.03	2-3-9-9-2
	.20	−.04	.44	.05	2-3-9-9-1
Jensen, Watanabe, Richters, Cortes, Roper, & Lui (1995)	.13	−.05	.30	.07	9-3-3-9-2
Kazak, Christakis, Alderfer, & Coiro (1994)	.25	−.10	.60	.08	3-3-3-9-1
Sawyer, Baghurst, & Mathias (1992)	.32	.12	.75	.08	3-1-9-4-2
	.38	−.05	.82	.04	3-2-9-4-2
	.19	−.14	.53	.13	3-1-9-4-1
	−.16	−.65	.33	.26	3-2-9-4-1
Sharp et al. (1995)	−.02	−.65	.60	.47	1-1-9-9-2
	.15	−.39	.70	.29	1-1-9-9-2
Webster-Stratton (1992)	.21	−.18	.59	.15	2-3-9-4-1
	.02	−.46	.49	.47	2-3-9-4-1
Webster-Stratton & Hammond (1990)	.08	−.26	.42	.32	1-3-9-4-1
Webster-Stratton & Hammond (1997)	.70	.06	1.34	.02	2-3-3-9-1
	.34	−.24	.92	.13	2-3-3-9-1
	.36	−.25	.98	.13	2-3-3-9-1
	−.13	−.79	.53	.35	2-3-3-9-1

[a]Refer to Table 3 for key to categorical variables.

children and adolescents. This procedure was also used to reduce the dependency among single studies when multiple effect sizes were utilized. Dependency upon studies where multiple effect sizes were examined but analyzed separately was reduced because effect size indicators were partitioned to the specific moderator (Erel & Burman, 1995). The following moderator variables were included in our analyses: (a) age of child, (b) gender of child, (c) ethnicity of child, and (d) socioeconomic status. A special population moderator (e.g., clinical vs. nonclinical child and/or parent; HIV-positive child and/or parent) was examined as well. Because of the small number of studies available for a thorough analysis and due to the heterogeneity of study populations, however, meaningful interpretations were unable to be made and these studies were not included in the moderator analyses.[2] The moderator variables and their categories are shown in Table 3. The effect sizes from the collected studies were used in order to analyze the influence of these different moderator variables in the case of *r* and *g*. Both statistical outcomes for homogeneity (Q_W) and between category differences (Q_B) are reported. To control for the inflation of familywise error rate, the alpha level for between-class testing was fixed at .001. Significant differences in between-class testing were further explored using post hoc contrasts.

Table 3. Potential moderator variables and their categories

Moderator	Categories
Age of children	1. 3–5 years 2. 6–12 years 3. 13–19 years
Gender of children	1. Boys only 2. Girls only 3. Boys and Girls
Ethnicity of children	1. Caucasian (all or primarily) 2. African American (all or primarily) 3. Diverse ethnicities (Caucasian and other) 4. Diverse ethnicities without Caucasian 5. Jewish
Socioeconomic status	1. Low 2. Middle 3. High 4. Diverse status

Note. All variables were coded based on the information presented in the studies used in the meta-analyses. Studies were assigned to age groups in the Age of Children moderator based on the mean age provided for the majority of children in the study. Gender, Ethnicity, and Socioeconomic Status of Children moderators were decided based on the information explicitly reported by the studies used in the meta-analysis. For each moderator, a value of 9 indicates that the information is missing from the published study. Categorical variables from Tables 1 and 2 are listed in the order indicated in this table.

Table 4. Maternal and paternal correspondence for behavior problems (*r*)

Behavior Problem	No. of Effect Sizes	Mean *r*	Mean Weighted *r*
Internalizing	18	.46	.45
Externalizing	22	.66	.63
Total	17	.61	.70

Note. All values are significant at the $p < .000$ level.

RESULTS

How Well Do Mothers and Fathers Correspond in Their Ratings of Children's Behavior Problems?

Mean Composite Effect Sizes (r). The results of the computations of mean *r* effect sizes and mean weighted *r* effect sizes for maternal and paternal ratings of internalizing, externalizing, and total behavior problems are presented in Table 4, which shows that the mean weighted effect sizes range from .45 to .70. Although the correspondence for maternal and paternal ratings is moderate for children's internalizing behavior problems, maternal and paternal ratings show a large correspondence for children's externalizing and total behavior problems. Thus, mothers and fathers exhibited moderate correspondence in their ratings of children's internalizing behavior problems, whereas they tend to show high levels of correspondence in their ratings of children's externalizing and total behavior problems.

The number of independent effect sizes describing maternal and paternal correspondence in the reporting of children's emotional and behavioral problems varies with the type of problem examined. Many studies reported maternal and paternal correspondence for externalizing behavior problems, and slightly fewer studies reported maternal and paternal correspondence for internalizing and total behavior problems.

Correspondence for Internalizing Behavior Problems. An overall model of maternal and paternal correspondence for internalizing behavior problems in children was examined to determine the homogeneity of effect sizes in the sample. It was determined that these effect sizes were not homogeneous, Q_T (df = 17) = 257.33, $p < .001$. Therefore, categorical model testing was used to examine whether potential moderator categories accounted for variation in the magnitude of effect sizes. The results of

Table 5. Tests of categorical models for *r* effect sizes: internalizing behavior problems

Variable and Category	Between-Classes Effect (Q_B)	No. of Effect Sizes	*N*	*r*	Mean Weighted Effect Size (*d*)	Confidence Interval for *d*		Within Class (Q_w)
						Lower	Upper	
Age of children	14.62***							
Early childhood		1	24	.12	.24	−.33	.81	0
Middle childhood		8	566	.28	.59	.47	.72	103.44***
Adolescence		4	226	.45	1.00	.81	1.20	10.56*
Gender of children	9.09							
Boys only		1	46	.48	1.09	.70	1.49	0
Girls only		1	57	.62	1.58	1.11	2.05	0
Boys and girls		16	1468	.41	.89	.80	.97	248.24***
Ethnicity of children	14.27							
Caucasian		4	229	.43	.94	.74	1.14	16.01**
Diverse ethnicity		8	596	.39	.84	.72	.97	50.68***
Ethnic groups excluding caucasian		1	112	.56	1.35	1.06	1.64	0
Jewish		1	49	.24	.49	.09	.90	0
Socioeconomic status	15.71***							
Low		1	49	.24	.49	.09	.90	0
Middle		2	103	.54	1.30	.99	1.60	2.43
Diverse status		8	721	.56	1.34	1.23	1.45	69.93***

Note. For Q_w, significance indicates rejection of homogeneity.
*$p < .05$, **$p < .01$, ***$p < .001$.

the categorical model testing for internalizing behavior problems are shown in Table 5.

The examination of between-class effects demonstrated that there was a significant difference in maternal and paternal correspondence among age groups of children for internalizing behavior problems (see first column of numbers in Table 5). Post hoc contrasts revealed that the average effect size for children in early childhood was not significantly different from that for children in middle childhood ($z^2 = 1.38$, NS). In contrast, the average effect sizes for children in early childhood and in middle childhood were significantly lower than that for adolescents ($z^2 = 6.15$, $p < .05$, and $z^2 = 12.19$, $p < .01$, respectively). This finding suggested that parental correspondence is higher for adolescents than for younger children when examining internalizing behavior problems.

Between-class effects testing revealed that there were no significant differences in average effect sizes among studies examining boys only, girls only, or both boys and girls. In addition, between-class effects testing demonstrated that there were no significant differences in average effect sizes across different ethnic categories.

The examination of between-class effects demonstrated that there was a significant difference in maternal and paternal correspondence among different categories of socioeconomic status. Post hoc contrasts revealed that the average effect size for children in the low socioeconomic category was significantly lower than the average effect sizes for children in the middle socioeconomic and mixed socioeconomic categories ($z^2 = 9.74$, $p < .01$, and $z^2 = 15.93$, $p < .001$, respectively). The average effect size for children in the middle socioeconomic category, however, was not significantly different from that for children in the mixed socioeconomic category ($z^2 = .07$, NS). This finding suggested that maternal and paternal correspondence for internalizing problems was higher for children in middle and mixed socioeconomic status groups than for children in a lower socioeconomic group.

Correspondence for Externalizing Behavior Problems. An overall model of maternal and paternal correspondence for externalizing behavior problems in children and adolescents was examined to determine the homogeneity of effect sizes in the sample. It was determined that these effect sizes were not homogeneous, Q_T (df = 21) = 253.17, $p < .001$. Therefore, categorical model testing was used to examine whether potential moderator categories accounted for variation in the magnitude of effect sizes. The results of the categorical model testing for externalizing behavior problems can be found in Table 6.

The examination of between-class effects demonstrated that there was a significant difference in maternal and paternal correspondence among age groups of children for externalizing behavior problems (see first column

Table 6. Tests of categorical models for *r* effect sizes: externalizing behavior problems

Variable and Category	Between-Classes Effect (Q_B)	No. of Effect Sizes	*N*	*r*	Mean Weighted Effect Size (*d*)	Confidence Interval for *d*		Within Class (Q_w)
						Lower	Upper	
Age of children	41.50***							
Early childhood		2	613	.47	1.07	.95	1.20	.60
Middle childhood		10	534	.55	1.31	1.17	1.45	33.53***
Adolescence		6	737	.63	1.63	1.50	1.75	15.28**
Gender of children	15.55***							
Boys only		3	178	.63	1.63	1.39	1.87	6.84*
Girls only		1	57	.70	1.96	1.51	2.41	0
Boys and girls		17	2327	.54	1.28	1.22	1.35	230.04***
Ethnicity of children	9.17							
Caucasian		4	263	.53	1.25	1.06	1.43	4.94
Diverse ethnicity		14	1854	.54	1.27	1.20	1.33	129.63***
Jewish		1	49	.30	.63	.22	1.04	0
Socioeconomic status	17.62***							
Low		1	49	.30	.63	.22	1.04	0
Middle		4	226	.63	1.61	1.39	1.82	7.52
Diverse status		9	1270	.58	1.42	1.33	1.51	108.13***

Note. For Q_w, significance indicates rejection of homogeneity.
*$p < .05$, **$p < .01$, ***$p < .001$.

of numbers in Table 6). Post hoc contrasts revealed that the average effect size for children in early childhood was significantly lower than those for children in middle childhood and in adolescence ($z^2 = 6.17$, $p < .05$, and $z^2 = 38.33$, $p < .001$, respectively). In addition, the average effect size for children in middle childhood was significantly lower than that for adolescents ($z^2 = 10.99$, $p < .01$). Thus, when examining externalizing behavior problems, parental correspondence varied according to the age of the child or adolescent, with correspondence being greater for adolescents than for children in early and middle childhood.

Between-class effects testing also revealed that there was a significant difference in maternal and paternal correspondence among children of different genders. Post hoc contrasts demonstrated that the average effect size for boys was not significantly different from that for girls ($z^2 = 1.59$, NS). In contrast, the average effect sizes for boys only and for girls only were both significantly higher than that for studies that examined boys and girls collectively ($z^2 = 7.67$, $p < .05$, and $z^2 = 8.64$, $p < .05$). This finding suggested that the degree of maternal and paternal correspondence varied according to how gender is grouped in studies examining externalizing behavior problems in children and adolescents, with studies that examined boys and girls separately having higher effect sizes, on average, than those examining boys and girls collectively.

Although the examination of between-class effects demonstrated that there were no significant differences in average effect sizes among different categories of ethnicity, between-class effects testing demonstrated that there was a significant difference in maternal and paternal correspondence among different categories of socioeconomic status. Post hoc contrasts revealed that the average effect size for children in the low socioeconomic category was significantly lower than those for children in the middle socioeconomic and in the mixed socioeconomic categories ($z^2 = 17.39$, $p < .001$, and $z^2 = 13.89$, $p < .001$, respectively). In contrast, the average effect size for children in the middle socioeconomic category was not significantly different from that for children in the mixed socioeconomic category ($z^2 = 2.59$, NS). This finding suggested that parental correspondence in ratings of externalizing behavior problems varied as a function of level of socioeconomic status of the family, with correspondence being greater in middle and mixed socioeconomic groups than in lower socioeconomic groups.

Correspondence for Total Behavior Problems. An overall model of maternal and paternal correspondence for total behavior problems in children was examined to determine the homogeneity of effect sizes in the sample. It was determined that these effect sizes were not homogeneous, Q_T (df = 16) = 177.72, $p < .001$. Therefore, categorical

Table 7. Tests of categorical models for *r* effect sizes: total behavior problems

Variable and Category	Between-Classes Effect (Q_B)	No. of Effect Sizes	*N*	*r*	Mean Weighted Effect Size (*d*)	Confidence Interval for *d*		Within Class (Q_w)
						Lower	Upper	
Age of children	1.90							
Middle childhood		8	331	.59	1.48	1.30	1.65	67.70***
Adolescence		5	209	.54	1.28	1.07	1.50	9.98
Gender of children	3.07							
Boys only		1	46	.67	1.81	1.32	2.29	0
Girls only		1	57	.53	1.25	.85	1.65	0
Boys and girls		15	1022	.58	1.43	1.32	1.53	174.65***
Ethnicity of children	.00							
Caucasian		2	76	.55	1.32	.97	1.67	.08
Diverse ethnicity		8	297	.55	1.31	1.14	1.49	24.59***
Socioeconomic status	8.90							
Middle		3	166	.56	1.35	1.11	1.59	4.65
Diverse status		7	615	.66	1.76	1.63	1.90	91.45***

Note. For Q_w, significance indicates rejection of homogeneity.
***$p < .001$.

model testing was used to examine whether potential moderator categories accounted for variation in the magnitude of effect sizes. The examination of between-class effects demonstrated that there were no significant differences across age, gender, ethnicity, and socioeconomic categories. The results of the categorical model testing for total behavior problems can be found in Table 7 (see first column of numbers for between-class effects).

Are There Discrepancies Between Mothers' and Fathers' Ratings of Children's Behavior Problems?

Mean Composite Effect Sizes (g). The results of the computation of the mean *g* effect sizes and mean weighted *g* effect sizes for maternal and paternal ratings of internalizing, externalizing, and total behavior problems are shown in Table 8: The mean weighted effect sizes for maternal and paternal discrepancies in the rating of emotional and behavioral problems in children and adolescents ranged from .08 to .17. Although each of these effect sizes can be categorized as small in nature, the mean weighted effect size for differential ratings of externalizing behavior problems (.08) was somewhat smaller than those for differential ratings of internalizing (.16) and total behavior problems (.17). This finding indicates that mothers and fathers exhibit somewhat fewer differences in their ratings of externalizing behavior problems than in their ratings of internalizing or total behavior problems, but the discrepancies in their ratings were not statistically, and perhaps not clinically, significant. Further, there was a tendency for mothers to report more behavior problems than fathers, as indicated by the positive effect sizes (which were calculated by subtracting fathers' mean ratings from mothers' mean ratings of behavior problems).

Table 8. Maternal and paternal discrepancies in rating behavior problems (*g*)

Behavior Problem	No. of Effect Sizes	Mean *g*	Mean Weighted *g*
Internalizing	15	.14	.16
Externalizing	13	.06	.08
Total	27	.18	.17

Note. The *p* values for the listed mean *g* values are .44, .48, and .43, respectively. The *p* values for the mean weighted *g* values are .44, .47, and .43, respectively. These values indicated that maternal and paternal ratings of internalizing, externalizing, and total behavior problems were not significantly different from one another.

From the number of effect sizes described in Table 8, it is apparent that the number of independent effect sizes pertaining to each type of behavior problem in children and adolescents varies greatly. Many studies reported maternal and paternal mean ratings on some measure of total behavior problems, whereas fewer studies reported maternal and paternal mean ratings on measures assessing internalizing or externalizing behavior problems.

Discrepancies in Ratings of Internalizing Behavior Problems. An overall model of maternal and paternal mean ratings of internalizing behavior problems in children and adolescents was examined to determine the homogeneity of effect sizes in the sample. It was determined that these effect sizes were homogeneous, Q_T (df = 14) = 11.15, NS. Therefore, none of the potential moderator catego-

ries used for categorical model testing accounted for variation in the magnitude of effect sizes. These results are reflected in the lack of significant between-class effects across age, gender, ethnicity, and socioeconomic status categories. These results for internalizing behavior problems are presented in Table 9 (see first column of numbers for between-class effects).

Discrepancies in Ratings of Externalizing Behavior Problems. An overall model of maternal and paternal mean ratings of externalizing behavior problems in children was examined to determine the homogeneity of effect sizes in the sample. It was determined that these effect sizes were homogenous, Q_T (df = 12) = 10.17, NS. Therefore, none of the potential moderator categories used for categorical model testing accounted for variation in the magnitude of effect sizes. These results are reflected in the lack of significant between-class effects across age, gender, ethnicity, and socioeconomic status categories. These results for externalizing behavior problems are presented in Table 10.

Discrepancies in Ratings of Total Behavior Problems. An overall model of maternal and paternal mean ratings of

Table 9. Tests of categorical models for *g* effect sizes: internalizing behavior problems

Variable and Category	Between-Classes Effect (Q_B)	No. of Effect Sizes	*N*		Mean Weighted Effect Size (*g*)	Confidence Interval for *g*		Within Class (Q_w)
			Mom	Dad		Lower	Upper	
Age of children	1.77							
Early childhood		1	151	126	.24	.00	.48	0
Middle childhood		6	114	96	−.00	−.28	.27	5.04
Adolescence		7	345	333	.15	−.01	.30	3.85
Gender of children	1.46							
Boys only		3	128	121	.07	−.18	.32	.90
Girls only		3	95	90	.06	−.24	.35	6.74*
Boys and girls		9	477	434	.20	.08	.33	2.32
Ethnicity of children	1.23							
Caucasian		1	151	126	.24	.00	.48	0
African American		1	90	90	.26	−.03	.55	0
Diverse ethnicity		5	191	159	.09	−.13	.31	5.60
Socioeconomic status	.16							
Low		1	90	90	.26	−.03	.55	0
Diverse status		6	343	313	.19	.04	.34	5.51

Note. For Q_w, significance indicates rejection of homogeneity.
*$p < .05$.

Table 10. Tests of categorical models for *g* effect sizes: externalizing behavior problems

Variable and Category	Between-Classes Effect (Q_B)	No. of Effect Sizes	*N*		Mean Weighted Effect Size (*g*)	Confidence Interval for *g*		Within Class (Q_w)
			Mom	Dad		Lower	Upper	
Age of children	3.25							
Early childhood		1	151	126	.04	−.20	.28	0
Middle childhood		6	276	264	−.11	−.39	.17	3.20
Adolescence		6	114	96	.18	.01	.34	3.71
Gender of children	1.24							
Boys only		3	128	121	.16	−.09	.41	.50
Girls only		3	95	90	−.06	−.35	.23	5.18
Boys and girls		7	318	275	.09	−.07	.26	3.24
Ethnicity of children	.00							
Caucasian		1	151	126	.04	−.20	.28	0
Diverse ethnicity		4	122	90	.03	−.25	.30	4.61
Socioeconomic status	0							
Diverse status		6	343	313	.10	−.06	.25	2.91

Note. For Q_w, significance indicates rejection of homogeneity.

Table 11. Tests of categorical models for *g* effect sizes: total behavior problems

Variable and Category	Between-Classes Effect (Q_B)	No. of Effect Sizes	*N*		Mean Weighted Effect Size (*g*)	Confidence Interval for *g*		Within Class (Q_w)
			Mom	Dad		Lower	Upper	
Age of children	.95							
Early childhood		3	148	91	.08	−.18	.34	.17
Middle childhood		11	526	480	.19	.07	.32	6.06
Adolescence		6	276	264	.24	.06	.41	3.68
Gender of children	.43							
Boys only		7	227	206	.19	.01	.38	1.74
Girls only		5	140	141	.10	−.13	.33	2.90
Boys and girls		15	984	870	.18	.09	.26	7.81
Ethnicity of children	.64							
Caucasian		2	85	85	.05	−.25	.35	.01
Diverse ethnicity		6	440	384	.18	.05	.32	4.56
Socioeconomic status	.40							
Middle		2	85	85	.05	−.25	.35	.01
Diverse status		7	372	314	.16	.01	.31	3.83

Note. For Q_w, significance indicates rejection of homogeneity.

total behavior problems in children was examined to determine the homogeneity of effect sizes in the sample. It was determined that these effect sizes were homogeneous, Q_T (df = 26) = 12.88, NS. Therefore, none of the potential moderator categories used for categorical model testing accounted for variation in the magnitude of effect sizes. These results are reflected in the lack of significant between-class effects across age, gender, ethnicity, socioeconomic, and special population categories. These results for total behavior problems are presented in Table 11.

Overall, mothers and fathers exhibited very small and nonsignificant discrepancies in their ratings of emotional and behavioral problems in children and adolescents. The effect sizes examined were homogeneous. As a result, none of the potential moderators examined in this study proved to be important predictors of discrepancies in maternal and paternal ratings of emotional and behavioral problems in children and adolescents.

DISCUSSION

What Is the Correspondence Between Mothers and Fathers in Their Ratings of Emotional and Behavioral Problems in Children and Adolescents?

Maternal and paternal correspondence in ratings for internalizing problems tended to be of moderate size, whereas mothers and fathers tended to have large correspondence for externalizing and total behavior problems. These findings are similar to those reported by the meta-analytic study conducted by Achenbach and colleagues (1987).

Regarding moderators of maternal and paternal ratings of internalizing and externalizing behavior problems, the results differed depending on the type of problem being rated. For both internalizing and externalizing problems, there was a significant difference in maternal and paternal correspondence among age groups of children, with average effect sizes for children in early childhood and middle childhood being significantly lower than ratings for adolescents. This finding is inconsistent with that of Achenbach and colleagues (1987), who found higher agreement for younger children than for adolescents across all combinations of informants. However, data presented separately for parents in the CBCL manual show negligible differences between correlations for parents' ratings of children versus adolescents (Achenbach, 1991). Parents in this study also tended to demonstrate higher correspondence in their ratings of internalizing and externalizing behaviors when they were rating children and adolescents in the middle or mixed socioeconomic status category than when rating children and adolescents in the low socioeconomic category. This finding is consistent with the findings from the national normative sample of the CBCL (Achenbach, 1991), which showed differences in maternal and paternal ratings with regard to socioeconomic status. However, this conclusion should be viewed with caution due to the small number of studies in this analysis. When gender was examined separately rather than collectively, maternal and paternal correspondence was higher for externalizing problems, and it is simply not clear why

combining gender categories in studies would create this result. Our findings are also inconsistent with past research which found more mother-father agreement on observable behaviors that violate social role expectations (e.g., conduct problems in girls; Jensen et al., 1988). Additionally, Tarullo, Richardson, Radke-Yarrow, and Martinez (1995) found that interparental agreement varied as a function of gender and type of problem, depending on the age of the child. There were no significant differences in the correspondence between maternal and paternal ratings when ethnic categories were examined for internalizing and externalizing behavior problems.

For total behavior problems, parental correspondence did not differ across age, gender, ethnicity, or socioeconomic status categories. It may be that certain categorical moderator variables that were significant in predicting internalizing and externalizing problems were not related to correspondence for total behavior problems because total behavior problems also include other types of behavior problems that do not fall neatly into the internalizing and externalizing categories. For example, on the CBCL (Achenbach, 1991), included in the total behavior problems scale are subscales examining thought problems, social problems, and attention problems, which when rated by parents, may not have similar moderators to those found for internalizing and externalizing categories because of the nature of those types of problems.

Do Mothers and Fathers Actually Differ in Their Ratings of Emotional and Behavioral Problems in Children and Adolescents?

Regarding maternal and paternal discrepancies in the rating of emotional and behavioral problems, findings of the present meta-analytic study suggested that mothers and fathers tended to exhibit fewer differences in their ratings of externalizing behavior problems than in their ratings of internalizing or total behavior problems, but the difference was small and nonsignificant. For internalizing, externalizing, and total behavior problems, none of the potential moderator categories accounted for variation in the magnitude of mean weighted effect sizes regarding the discrepancy between maternal and paternal reports. It is hypothesized that because the effect sizes were so small (ranging from .08 to .17), there was not sufficient variability to detect moderators as there was with the examination of correspondence.

The slight disagreement that did exist between mothers and fathers appeared to be systematic. Overall, mothers tended to report slightly more behavior problems than fathers. This finding has been demonstrated by other researchers as well (Christensen et al., 1992; Eisenstadt et al., 1994; Robinson et al., 1980). To interpret this tendency, several explanations have been proposed. One possibility is that mothers spend more time than fathers with their children (Biller, 1993) and, as a result, have more exposure to the behavior of their children and adolescents. Another possibility is that parental discrepancies may reflect actual differences in the observed behavior of their children. For example, in preschool children, it has been found that children obey their fathers more frequently than their mothers, and that children are more likely to obey their mothers in the presence of their fathers (Campbell, 1991; Patterson, 1980). Comparable findings were reported for a sample of school-aged hyperactive children (Tallmadge & Barkley, 1983). Thus, it may be that fathers are less likely to see disruptive behavior in children and adolescents than mothers, and parents' ratings may be an accurate reflection of what they experience. This tendency may hold for internalizing problems, older children, and adolescents as well.

Limitations of the Present Study

Results of this study should be examined in terms of the limitations of meta-analyses in general, and in terms of the limitations of the present study. First, like most meta-analyses, data analyzed in this study were taken from previously published works. Commonly called the "file drawer problem," a difficulty lies in the possibility of a retrievability bias, in which unpublished studies do not contribute to the database for the present study (Rosenthal, 1991). Second, meaningful information can be lost when research is summarized by a single parameter as is the practice in meta-analytic reviews (Rosenthal, 1991). Finally, common to meta-analyses, variables that are heterogeneous are frequently combined (Rosenthal, 1991). In response to this problem, potential moderators were examined in the current study. It is hoped that the problem of heterogeneity has been lessened and that the loss of important information has been decreased by using this procedure.

There are also limitations specific to the current meta-analysis. First, many of the studies failed to provide information regarding family factors (e.g., distress level of the family, biological versus adoptive children, parents' levels

of psychological symptoms, marital status) that could have served as moderators of parents' ratings of child and adolescent behavior. For example, Christensen and colleagues (1992) found that parental discrepancies tended to increase with the distress level of the family. It has also been found that parents' own levels of psychological symptoms and the types of symptoms that parents experienced were related to interparental agreement (Griest, Wells, & Forehand, 1979; Jensen et al., 1988). Thus, many of these variables would have provided valuable information relevant to interparental ratings of emotional and behavioral problems in children and adolescents, but could not be adequately studied in the present work. Second, many studies failed to report ethnicity or the socioeconomic status of the participants, resulting in small sample sizes in some of the analyses. Phares and Lum (1996) argued that, although these data are often absent in published studies, demographic data need to be reported in a more thorough manner in order to understand the sample characteristics. Third, many studies examining mothers' and fathers' ratings of problematic behavior were excluded from our analyses because specific statistics (e.g., means, correlations) were not reported in the original articles for us to utilize in our study. It is unknown whether or not these studies, if included, would have influenced our results. Along these lines, the validity of our current conclusions would be compromised if there was, in fact, a selection bias. Certainly more information would have been obtained to facilitate the examination of moderators if these studies did provide statistical information. Finally, different patterns of results were found when examining parents' correspondence in their ratings of emotional and behavioral problems than when examining parents' discrepancy in their ratings of these problems, even though the same significant moderators should have emerged. This issue may relate to the different questions that were asked (i.e., do maternal and paternal ratings correlate? vs. do maternal and paternal ratings differ?). As a result of the different emphases in these research questions, it is possible to find different patterns of results based on correlations and on differences in means (as was found in this study).

Implications of Findings and Future Directions for Research

Similar to the work of Achenbach and colleagues (1987), maternal and paternal ratings tended to have a high degree of correspondence when rating problematic behaviors in children and adolescents. The results of this study suggested that one parent would typically provide a reasonable estimate of what would be provided by another parent who sees a child or adolescent under similar conditions. Due to practical constraints and unavailable resources, it is not always possible to gather information from both parents. This study suggests that, if necessary, one parent's ratings may suffice in reporting information about emotional and behavioral problems in children and adolescents. Further, the finding that mothers and fathers in general report about the same level of problems has direct implications for individual cases. For example, if major discrepancies are found between ratings by a particular mother and father, the discrepancies should not be written off as typical of mothers' and fathers' reports in general, and may be clinically meaningful. Instead, the clinician may want to examine possible factors related to their disagreement for that particular family. For clinicians who use the CBCL and related measures (Achenbach, 1991), cross-informant patterns of agreement can actually be calculated for an individual case and then compared with the normative data for those measures.

Clinicians and researchers should be aware that different moderators are related to the correspondence of maternal and paternal reports when considering individual cases. Additionally, it is important to keep in mind that other factors (e.g., parental psychological functioning, distress level of the family, marital status), which were not examined directly in the present study, may alter the correspondence and discrepancies in maternal and paternal ratings. In future studies utilizing mothers and fathers as informants, it is suggested that researchers include these variables and descriptive characteristics, as well as other important demographic data to facilitate study in this area. It is also necessary to remember the tendency for parents (Phares, 1997) and professionals (Loeber, Green, & Lahey, 1990) to perceive mothers as the most useful and accurate informants of emotional and behavioral problems in children and adolescents. There has also been a tendency to place less emphasis on fathers in research and clinical settings (Phares, 1992). The important role of fathers in the development of their children's psychopathology is becoming increasingly apparent (Phares, 1996), however, and this area of study is becoming more influential.

It is also important to keep in mind that it would be ideal to gather data from multiple informants in the assessment process (Achenbach et al., 1987) due to the variabil-

ity in children's and adolescents' behavior across situations. Whereas clinicians can feel somewhat confident that behavior ratings from one parent would be comparable to the ratings from the other parent, a thorough multiaxial assessment (Achenbach, 1985) with multiple informants is likely to lead to a better understanding of the child's or adolescent's emotional and behavioral functioning. Along these lines, there is evidence that ratings by different informants are related to one another in different ways (Achenbach et al., 1987), that mothers' versus fathers' ratings may be related to other informants' ratings in different fashions, and that there is greater utility when collecting multiple reports. For example, when examining parents' perceptions of child behaviors, and how they are related to others' ratings, findings have been equivocal. Several studies have found that mothers' ratings of externalizing behaviors related to teachers' ratings, but that fathers' ratings did not (Baker & Heller, 1996; Schaughency & Lahey, 1985). Yet another study found just the opposite: Fathers' ratings correlated with teachers' ratings whereas mothers' ratings did not (Webster-Stratton, 1988). Additionally, there appears to be increased incremental utility when combining teacher with parent reports for particular problems (Power et al., 1998). It would be interesting in future work to determine if there is a difference in utility when examining mothers' versus fathers' reports when combined with teachers' reports. Further, in future studies, it may be beneficial to determine if combining mothers' and fathers' ratings actually predict others' ratings such as teachers', clinicians', or children's ratings better than one parent's ratings alone.

Pertaining to assessment in general, information from multiple informants appears to be important for a variety of age groups, not just children and adolescents. Regarding young adults, upward extensions of the CBCL and Youth Self Report forms have been developed to facilitate multiple informant data to address specific problems faced by this age group (Achenbach, 1995). Multiple informants' information (often referred to as collateral reports) in the assessment of substance use, which is applicable to a very broad age group, has also been studied. Whereas there are certain drawbacks in using multiple informant data for this population, information from others (e.g., spouse, coworker, friend, family member) is recommended in certain situations (Carroll, 1995). Pertaining to the elderly, when studying cognitive decline in a community sample, it was found that agreement between the identified subject and close relatives or friends was poor, thus suggesting the importance of multiple ratings (Jorm et al., 1994). The use of multiple informants has also been suggested for providing clinical information in depressive symptoms associated with Alzheimer's disease (Gilley et al., 1995). Thus, although the use of multiple informant data has been examined in relation to childhood psychopathology, it has also received some study in the area of adult psychopathology and is becoming increasingly important. Future study in this area is warranted.

NOTES

1. Although manuscripts examining maternal and paternal ratings of social behavior and temperament were initially collected for inclusion in this study, these manuscripts did not provide a substantial number of effect sizes. Six studies provided *r* effect sizes for social behavior (Asher & Wakefield, 1990; D'Ilio & Karnes, 1992; Eiser, Havermans, Pancer, & Eiser, 1992; Henggeler, Watson, & Whelan, 1990; Sloper, Knussen, Turner, & Cunningham, 1991; Walker & Bracken, 1996). The average effect size for these studies was .41, whereas the average weighted effect size for these studies was .30. Seven studies provided eight *r* effect sizes for temperament (Boer & Westenberg, 1994; Jewsuwan, Luster, & Kostelnik, 1993; Saudino, McGuire, Hetherington, Reiss, & Plomin, 1995; Schaughency & Fagot, 1993; Walker & Bracken, 1996; Whiffen, 1990; Wolk, Zeanah, Coll, & Carr, 1992). The average effect size for these studies was .57, whereas the average weighted effect size for these studies was .53. These studies were not included in further analyses because they did not provide an adequate number of effect sizes to conduct categorical model testing.

2. For the special population moderator, studies were assigned to the category if a portion or all of the children in the sample were labeled as experiencing a medical or psychological disorder. Because of the heterogeneity of the sample, and difficulty with interpretation of the results, the moderator was not included in the main report of the findings. The studies, however, were used in the calculation of effect sizes for other moderator variables and for internalizing, externalizing, and total behavior problems effect size calculations.

Pertaining to correspondence for internalizing problems, between-class effects testing ($Q_B = 106.68$, $p < .001$) revealed that the average correspondence of maternal and paternal ratings was lower for children in the special population category than that for children who were not in the special population category. Five studies provided *r* effect sizes for these analyses (Eiser et al., 1992; Rosenberg, Brown, & Singer, 1994; Sternberg et al., 1993; Thurber & Osborn, 1993; Zahn-Waxler et al., 1996). Similar to findings for internalizing problems, the examination

of between-class effects ($Q_B = 25.97$, $p < .001$) for externalizing problems revealed that the average correspondence of maternal and paternal ratings was lower for children in the special population category than that for children who were not in the special population category. Six studies provided *r* effect sizes for these analyses (Henggeler et al., 1990; Kearney & Silverman, 1993; Rosenberg et al., 1994; Sternberg et al., 1993; Thurber & Osborn, 1993, Zahn-Waxler et al., 1996). The finding that correspondence was lower for both types of problems in the special population category is at odds with what Achenbach and colleagues (1987) found for clinical and nonclinical samples. Their study, however, did not include children with special needs or with various types of medical problems. For total behavior problems, the examination of between-class effects demonstrated that there was no significant difference between the special population categories. Six studies contributed *r* effect sizes for these analyses (Coyne, 1996; Eisenstadt et al., 1994; Rosenberg et al., 1994; Sloper et al., 1991; Thurber & Osborn, 1993; Watson, Henggeler, & Whelan, 1990).

Upon examining discrepancies in maternal and paternal reports of internalizing, externalizing, and total behavior problems, overall models for each type of problem were homogeneous. Therefore, the special population category was not an important predictor of discrepancies in parent reports. Four studies provided *g* effect sizes for internalizing problems (Cohen, Coyne, & Duvall, 1993; Kazak, Christakis, Alderfer, & Coiro, 1994; Reeve, Bernstein, & Christenson, 1992; Sawyer, Baghurst, & Mathias, 1992), four studies provided *g* effect sizes for externalizing problems (Cohen et al., 1993; Kazak et al., 1994; Reeve et al., 1992; Sawyer et al., 1992), and eight studies provided *g* effect sizes for total behavior problems (Cohen et al., 1993; Cuskelly & Dadds, 1992; Jensen, Bloedau, DeGroot, Ussery, & Davis, 1990; Kazak et al., 1994; Sawyer et al., 1992; Webster-Stratton, 1992; Webster-Stratton & Hammond, 1990, 1997).

ACKNOWLEDGMENT

This research was supported by National Institute of Mental Health Grant R29 49601-03.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

*Abidin, R. R., Jenkins, C. L., & McGaughey, M. C. (1992). The relationship of early family variables to children's subsequent behavioral adjustment. *Journal of Clinical Child Psychology, 21,* 60–69.

Achenbach, T. M. (1985). *Assessment and taxonomy of child and adolescent psychopathology.* Newbury Park, CA: Sage.

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist and 1991 Child Behavior Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1992). *Manual for the Child Behavior Checklist/2–3 and 1992 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1995). Empirically based assessment and taxonomy: Applications to clinical research. *Psychological Assessment, 7,* 261–274.

Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin, 85,* 1275–1301.

Achenbach, T. M., Howell, C. T., Quay, H. C., & Conners, C. K. (1991). National Survey of problems and competencies among four- to sixteen-year-olds: Parent's reports for normative and clinical samples. *Monographs of the Society for Research in Child Development, 56* (3, Serial No. 225).

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

*Armistead, L., Klein, K., Forehand, R., & Wierson, M. (1997). Disclosure of parental HIV infection to children in the families of men with hemophilia: Description, outcomes, and the role of family processes. *Journal of Family Psychology, 11,* 49–61.

*Asher, R. A., & Wakefield, J. A., Jr. (1990). The relationship between parents' personality and their ratings of their preschool children's behaviors. *Personality and Individual Difference, 11,* 1131–1136.

*Baker, B. L., & Heller, T. L. (1996). Preschool children with externalizing behaviors: Experience of fathers and mothers. *Journal of Abnormal Child Psychology, 24,* 513–532.

*Banez, G. A., & Compas, B. E. (1990). Children's and parents' daily stressful events and psychological symptoms. *Journal of Abnormal Child Psychology, 18,* 591–605.

Biller, H. B. (1993). *Fathers and families: Paternal factors in child development*. Westport, CT: Auburn House.

*Boer, F., & Westenberg, P. M. (1994). The factor structure of the Buss and Plomin EAS Temperament Survey (parental ratings) in a Dutch sample of elementary school children. *Journal of Personality Assessment, 62,* 537–551.

*Brody, G. H., Stoneman, Z., Flor, D., McCrary, C., Hastings, L., & Conyers, O. (1994). Financial resources, parent psychological functioning, parent co-caregiving, and early adolescent competence in rural two-parent African-American families. *Child Development, 65,* 590–605.

Campbell, S. B. (1991). *Behavior problems in preschool children: Clinical and developmental issues*. New York: Guilford Press.

*Campbell, S. B., Pierce, E. W., Moore, G., Marakovitz, S., & Newby, K. (1996). Boys' externalizing problems at elementary school age: Pathways from early behavior problems,

maternal control, and family stress. *Development and Psychopathology, 8,* 701–719.

Carroll, K. M. (1995). Methodological issues and problems in the assessment of substance use. *Psychological Assessment, 7,* 349–358.

Christensen, A., Margolin, G., & Sullaway, M. (1992). Interparental agreement on child behavior problems. *Psychological Assessment, 4,* 419–425.

*Cohen, N. J., Coyne, J., & Duvall, J. (1993). Adopted and biological children in the clinic: Family, parental and child characteristics. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 34,* 545–562.

*Coyne, J. C. (1996). Expressed emotion in families of clinically referred and nonreferred children: Toward a further understanding of the expressed emotion index. *Journal of Family Psychology, 10,* 97–106.

*Crockenberg, S., & Lourie, A. (1996). Parents' conflict strategies with children and children's conflict strategies with peers. *Merrill-Palmer Quarterly, 42,* 495–518.

*Cuskelly, M., & Dadds, M. (1992). Behavioural problems in children with Down's Syndrome and their siblings. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 33,* 749–761.

*Deater-Deckard, K., & Scarr, S. (1996). Parenting stress among dual-earner mothers and fathers: Are there gender differences? *Journal of Family Psychology, 10,* 45–59.

*D'Ilio, V. R., & Karnes, F. A. (1992). Student leaders' and their parents' perceptions of social skills. *Perceptual and Motor Skills, 74,* 863–866.

*Eisenberg, N., Fabes, R. A., Murphy, B., Maszk, P., Smith, M., & Karbon, M. (1995). The role of emotionality and regulation to children's social functioning: A longitudinal study. *Child Development, 66,* 1360–1384.

*Eisenstadt, T. H., McElreath, L. H., Eyberg, S. M., & McNeil (1994). Interparent agreement on the Eyberg Child Behavior Inventory. *Child and Family Behavior Therapy, 16,* 21–27.

*Eiser, C., Havermans, T., Pancer, M., & Eiser, J. R. (1992). Adjustment to chronic disease in relation to age and gender: Mothers' and fathers' reports of their children's behavior. *Journal of Pediatric Psychology, 17,* 261–275.

Erel, O., & Burman, B. (1995). Interrelatedness of marital relations and parent-child relations: A meta-analytic review. *Psychological Bulletin, 118,* 108–132.

*Fagot, B. I. (1995). Classification of problem behaviors in young children: A comparison of four systems. *Journal of Applied Developmental Psychology, 16,* 95–106.

Friedlander, S., Weiss, D. S., & Traylor, J. (1986). Assessing the influence of maternal depression on the validity of the Child Behavior Checklist. *Journal of Abnormal Child Psychology, 14,* 123–133.

Gilley, D. W., Wilson, R. S., Fleischman, D. A., Harrison, D. W., Goetz, C. G., & Tanner, C. M. (1995). Impact of Alzheimer's-type dementia and information source on the assessment of depression. *Psychological Assessment, 7,* 42–48.

Graham, P., & Stevenson, J. (1985). A twin study of genetic influences in behavioral deviance. *Journal of the American Academy of Child and Adolescent Psychiatry, 24,* 33–41.

Griest, D., Wells, K. C., & Forehand, R. (1979). An examination of predictors of maternal perceptions of maladjustment in clinic-referred children. *Journal of Abnormal Psychology, 88,* 277–281.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92,* 490–499.

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

*Henggeler, S. W., Watson, S. M., & Whelan, J. P. (1990). Peer relations of hearing-impaired adolescents. *Journal of Pediatric Psychology, 15,* 721–731.

*Jaycox, L. H., & Repetti, R. L. (1993). Conflict in families and the psychological adjustment of preadolescent children. *Journal of Family Psychology, 7,* 344–355.

*Jenkins, J. M., & Smith, M. A. (1991). Marital disharmony and children's behaviour problems: Aspects of a poor marriage that affect children adversely. *Journal of Child Psychology and Psychiatry, 32,* 793–810.

*Jensen, P. S., Bloedau, L., DeGroot, J., Ussery, T., & Davis, H. (1990). Children at risk I: Risk factors and child symptomatology. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 51–59.

Jensen, P. S., Traylor, J., Xenakis, S. N., & Davis, H. (1988). Child psychopathology rating scales and interrater agreement: I. Parents' gender and psychiatric symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 442–450.

*Jensen, P. S., Watanabe, H. K., Richters, J. E., Cortes, R., Roper, M., & Lui, S. (1995). Prevalence of mental disorder in military children and adolescents: Findings from a two-stage community survey. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1514–1524.

*Jewsuwan, R., Luster, T., & Kostelnik, M. (1993). The relation between parents' perceptions of temperament and children's adjustment to preschool. *Early Childhood Research Quarterly, 8,* 33–51.

Jorm, A. F., Christensen, H., Henderson, A. S., Korten, A. E., Mackinnon, A. J., & Scott, R. (1994). Complaints of cognitive decline in the elderly: A comparison of reports by subjects and informants in a community survey. *Psychological Medicine, 24,* 365–374.

*Jouriles, E. M., Norwood, W. D., McDonald, R., Vincent, J. P., & Mahoney, A. (1996). Physical violence and other forms of marital aggression: Links with children's behavior problems. *Journal of Family Psychology, 10,* 223–234.

Kadzin, A. (1988). The diagnosis of childhood disorders: Assessment issues and strategies. *Behavioral Assessment, 10,* 67–94.

*Kazak, A. E., Christakis, D., Alderfer, M., & Coiro, M. J. (1994). Young adolescent cancer survivors and their parents: Adjustment, learning problems, and gender. *Journal of Family Psychology, 8,* 74–84.

*Kearney, C. A., & Silverman, W. K. (1993). Measuring the function of school refusal behavior: The School Refusal Assessment scale. *Journal of Clinical Child Psychology, 22,* 85–96.

*Kitzmann, K. M., & Emery, R. E. (1994). Child and family coping one year after mediated and litigated child custody disputes. *Journal of Family Psychology, 8,* 150–159.

Kolko, D. J., & Kadzin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent, and teacher reports. *Journal of Child Psychology and Psychiatry, 34,* 991–1006.

Kovacs, M. (1994). *Children's Depression Inventory manual.* Toronto: Mental Health Systems.

Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. *Journal of Clinical Child Psychology, 19,* 136–143.

Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1989). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology, 1,* 317–337.

*Mann, B. J., & MacKenzie, E. P. (1996). Pathways among marital functioning, parental behaviors, and child behavior problems in school-age boys. *Journal of Clinical Child Psychology, 25,* 183–191.

*O'Brien, J., Margolin, G., & John, R. S. (1995). Relation among marital conflict, child coping, and child adjustment. *Journal of Clinical Child Psychology, 24,* 346–361.

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

Phares, V. (1992). Where's Poppa? The relative lack of attention to the role of fathers in child and adolescent psychopathology. *American Psychologist, 47,* 656–664.

Phares, V. (1996). *Fathers and developmental psychopathology.* New York: Wiley.

Phares, V. (1997). Accuracy of informants: Do parents think that mother knows best? *Journal of Abnormal Child Psychology, 25,* 165–171.

*Phares, V., & Compas, B. E. (1990). Adolescents' subjective distress over their emotional/behavioral problems. *Journal of Consulting and Clinical Psychology, 58,* 596–603.

Phares, V., & Compas, B. E. (1992). The role of fathers in child and adolescent psychopathology: Make room for daddy. *Psychological Bulletin, 111,* 387–412.

Phares, V., & Lum, J. J. (1996). Family demographics of clinically referred children: What we know and what we need to know. *Journal of Abnormal Child Psychology, 24,* 787–801.

Power, T. J., Andrews, T. J., Eiraldi, R. B., Doherty, M. J., Ikeda, M. J., DuPaul, G. J., & Landau, S. (1998). Evaluating attention deficit hyperactivity disorder using multiple informants: The incremental utility of combining teacher with parent reports. *Psychological Assessment, 10,* 250–260.

Quay, H. C., & Peterson, D. R. (1983). *Interim manual for the Revised Behavior Problem Checklist.* Unpublished manuscript, University of Miami.

*Radovanoic, H. (1993). Parental conflict and children's coping styles in litigating separated families: Relationships with children's adjustment. *Journal of Abnormal Child Psychology, 21,* 697–713.

*Reeve, E. A., Bernstein, G. A., & Christenson, G. A. (1992). Clinical characteristics and psychiatric comorbidity in children with trichotillomania. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 132–138.

Robinson, E. A., Eyberg, S. M., & Ross, A. W. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child Psychology, 9,* 22–28.

*Rosenbaum, M., & Ronen, T. (1997). Parents' and children's appraisals of each other's anxiety while facing a common threat. *Journal of Clinical Child Psychology, 26,* 43–52.

*Rosenberg, L. A., Brown, J., & Singer, H. S. (1994). Self-reporting of behavior problems in patients with tic disorders. *Psychological Reports, 74,* 653–654.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

*Rothbaum, F., Rosen, K. S., Pott, M., & Beatty, M. (1995). Early parent-child relationships and later problem behavior: A longitudinal study. *Merrill-Palmer Quarterly, 41,* 133–151.

*Saudino, K. J., McGuire, S., Hetherington, E. M., Reiss, D., & Plomin, R. (1995). Parent ratings of EAS temperaments in twins, full siblings, half siblings, and step siblings. *Journal of Personality and Social Psychology, 68,* 723–733.

*Sawyer, M. G., Baghurst, P., & Clark, J. (1992). Differences between reports from children, parents and teachers: Implications for epidemiological studies. *Australian and New Zealand Journal of Psychiatry, 26,* 652–660.

*Sawyer, M. G., Baghurst, P., & Mathias, J. (1992). Differences between informants' reports describing emotional and behavioural problems in community and clinic-referred children: A research note. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 33,* 441–449.

*Schaughency, E. A., & Fagot, B. I. (1993). The prediction of adjustment at age 7 from activity level at age 5. *Journal of Abnormal Child Psychology, 21,* 29–50.

Schaughency, E. A., & Lahey, B. B. (1985). Mothers' and fathers' perceptions of child deviance: Roles of child behavior, parental depression, and marital satisfaction. *Journal of Consulting and Clinical Psychology, 53,* 718–723.

*Sharp, D., Hay, D. F., Pawlby, S., Schmucker, G., Allen, H., & Kumar, R. (1995). The impact of postnatal depression on

boys' intellectual development. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 36,* 1315–1336.

*Sloper, P., Knussen, C., Turner, S., & Cunningham, C. (1991). Factors related to stress and satisfaction with life in families of children with Down's syndrome. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 32,* 655–676.

*Smith, M. A., & Jenkins, J. M. (1991). The effects of marital disharmony on prepubertal children. *Journal of Abnormal Child Psychology, 10,* 625–644.

*Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behavior problems. *Journal of Clinical Child Psychology, 22,* 107–115.

Stanger, C., McConaughy, S. H., & Achenbach, T. M. (1992). Three-year course of behavioral/emotional problems in a national sample of 4- to 16-year-olds: II. Predictors of syndromes. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 941–950.

*Steele, R. G., Forehand, R., & Armistead, L. (1997). The role of family processes and coping strategies in the relationship between parental chronic illness and childhood internalizing problems. *Journal of Abnormal Child Psychology, 25,* 83–94.

*Sternberg, K. J., Lamb M. E., Greenbaum, C., Cicchetti, D., Dawud, S., Cortes, R. M., Krispin, O., & Lorey, F. (1993). Effects of domestic violence on children's behavior problems and depression. *Developmental Psychology, 29,* 44–52.

*Stice, E., & Barrera, M. , Jr. (1995). A longitudinal examination of the reciprocal relations between perceived parenting and adolescent's substance use and externalizing behaviors. *Developmental Psychology, 31,* 322–334.

Tallmadge, J., & Barkley, R. A. (1983). The interactions of hyperactive and normal boys with their fathers and mothers. *Journal of Abnormal Child Psychology, 11,* 565–579.

Tarullo, L. B., Richardson, D. T., Radke-Yarrow, M., & Martinez, P. E. (1995). Multiple sources in child diagnosis: Parent-child concordance in affectively ill and well families. *Journal of Clinical Child Psychology, 24,* 173–183.

*Thurber, S., & Osborn, R. A. (1993). Comparisons of parent and adolescent perspectives on deviance. *The Journal of Genetic Psychology, 154,* 25–32.

Verhulst, F. C., & van der Ende, J. (1992). Agreement between parents' reports and adolescents' self-reports of problem behavior. *Journal of Child Psychology and Psychiatry, 33,* 1011–1023.

*Walker, K. C., & Bracken, B. A. (1996). Inter-parent agreement on four preschool behavior rating scales: Effects of parent and child gender. *Psychology in the Schools, 33,* 273–282.

*Watson, S. M., Henggeler, S. W., & Whelan, J. P. (1990). Family functioning and the social adaptation of hearing-impaired youths. *Journal of Abnormal Child Psychology, 19,* 143–163.

Webster-Stratton, C. (1988). Mothers' and fathers' perceptions of child deviance: Roles of parent and child behaviors and parent adjustment. *Journal of Consulting and Clinical Psychology, 56,* 909–915.

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

*Webster-Stratton, C., & Hammond, M. (1990). Predictors of treatment outcome in parent training for families with conduct problem children. *Behavior Therapy, 21,* 319–337.

*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.

*Westerman, M. A., & Schonholtz, J. (1993). Marital adjustment, joint parental support in a triadic problem-solving task, and child behavior problems. *Journal of Clinical Child Psychology, 22,* 97–106.

*Whiffen, V. E. (1990). Maternal depressed mood and perceptions of child temperament. *The Journal of Genetic Psychology, 151,* 329–339.

*Wolk, S., Zeanah, C. H., Coll, C. T., & Carr, S. (1992). Factors affecting parents' perceptions of temperament in early infancy. *American Journal of Orthopsychiatry, 62,* 71–82.

*Zahn-Waxler, C., Schmitz, S., Fulker, D., Robinson, J., & Emde, R. (1996). Behavior problems in five-year-old monozygotic and dizygotic twins: Genetic and environmental influences, patterns of regulation, and internalization of control. *Development and Psychopathology, 8,* 103–122.

Received May 12, 1999; revised February 3, 2000, and April 7, 2000; accepted May 2, 2000.