The relations among depression in fathers, children’s psychopathology, and father–child conflict: A meta-analysis

Peter Kane*, Judy Garber*

Vanderbilt University, Peabody 512, 230 Appleton Place, Nashville, TN 37203-5701, USA

Received 26 August 2002; received in revised form 13 October 2003; accepted 22 March 2004

Abstract

Research on parental depression is beginning to recognize the importance of studying fathers in relation to maladaptive outcomes in their offspring. Paternal depression is hypothesized to correlate with internalizing and externalizing psychopathology in children and adolescents and to compromise adaptive parent–child relationships (e.g., increased conflict). In the present paper, meta-analytic procedures were applied to this literature to address the magnitude and direction of covariation between paternal depression and children’s functioning. In addition, we tested whether variation in findings could be accounted for by study characteristics. Results indicated that paternal depression was significantly related to offspring internalizing and externalizing psychopathology and father–child conflict. Larger effects for internalizing symptoms were associated with the use of community samples and symptom rating scales of internalizing problems.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Fathers; Depression; Child psychopathology; Conflict

1. Introduction

The emergence of developmental psychopathology has led to an expansion of research on the developmental sequelae of offspring of depressed mothers. Studies consistently have shown that maternal depression is associated with internalizing and externalizing psychopathology in youth and irritable, hostile, and critical interactions between mothers and children (Conger, Patterson, & Ge, 1995; Cummings & Davies, 1999; Ge, Conger, Lorenz, & Simons, 1994; Harnish, Dodge, Valente, & Conduct

* Corresponding authors.

E-mail addresses: Peter.P.Kane@Vanderbilt.edu (P. Kane), Judy.Garber@Vanderbilt.edu (J. Garber).
Problems Prevention Research Group, 1995; Patterson & Dishion, 1988; Webster-Stratton & Hammond, 1988). Qualitative and quantitative reviews of the literature (Beck, 1999; Downey & Coyne, 1990; Goodman & Gotlib, 1999; Lovejoy, Graczyk, O’Hare, & Neuman, 2000) have concluded that consistent relations exist among maternal depression, child psychopathology, and negative mother–child interactions, and have suggested that conflict may mediate the association between maternal depression and child maladjustment (Conger et al., 1995; Ge et al., 1994; Harnish et al., 1995).

Ecological models of development, however, assert that relations between parent attributes and child outcomes may not be limited to mothers (Bronfenbrenner, 1986; Cicchetti & Toth, 1997). Psychopathology in fathers also may be related to maladaptive development of children and adolescents (Compas, Howell, Phares, Williams, & Ledoux, 1989a,b; Connell & Goodman, 2002). Empirical investigations of the relation between paternal depression and child outcomes have begun in recent years, thereby making a quantitative synthesis of these findings now possible.

Although previous reviews (Phares, 1996, 1997; Phares & Compas, 1992) have demonstrated links between paternal depression and child psychopathology, several limitations have reduced the accuracy of the conclusions and clarity of constructs reviewed in these papers. First, some studies used the same sample rather than independent samples (Radke-Yarrow, Cummings, Kuczynski, & Chapman, 1985; Radke-Yarrow, Nottelman, Martinez, Fox, & Belmont, 1992; Zahn-Waxler, Cummings, McKnew, & Radke-Yarrow, 1984). Thus, the inferences drawn from these studies may be limited by statistical dependencies among the samples.

Second, these papers have drawn conclusions based on samples of fathers with different types of mood disorders, including unipolar and bipolar depression. For example, Phares (1996) included in her review a study of cyclothymia in offspring of parents with bipolar disorder (Klein, Depue, & Slater, 1985). Although unipolar and bipolar depression are both mood disorders, they likely are distinct forms of psychopathology with different correlates and consequences (Keller, 1987).

Third, studies have varied with regard to whom was the informant about depressive symptoms in fathers. For instance, one study (Beardslee, Schultz, & Selman, 1987) used mothers’ report, whereas most of the other studies reviewed used fathers’ self-reports of symptoms. Spouse’s reports of depression may not be as valid as fathers’ self-report, and therefore may less accurately characterize relations between paternal and child psychopathology.

Fourth, these reviews have drawn conclusions about child functioning in general rather than focusing on different dimensions of child psychopathology. To examine the issue of specificity, the relation between paternal depression and children’s internalizing and externalizing symptoms should be explored separately. Moreover, some of the studies included in prior reviews did not focus on psychopathology, but rather examined other outcomes, such as attachment security. Although security of attachment may be linked with the development of psychopathology in children (Cummings & Cicchetti, 1990; Greenberg, Speltz, & DeKlyen, 1993; Shaw, Owens, Vondra, Keenan, & Winslow, 1996), it is not an index of psychopathology per se.

Finally, a qualitative synthesis cannot assess the degree of covariation between depression in fathers and child psychopathology. Thus, the magnitude and direction of effect sizes are still unknown. Are effect sizes of modest, moderate, or large magnitude? Are these relations positive or negative?

The present paper used meta-analysis to address these questions regarding effect sizes and to build upon the limitations of previous reviews by examining associations between depression in fathers and three outcome domains—child internalizing, externalizing, and father–child conflict. Thus, multiple reports from one sample as well as studies examining nonfather reports of depression, father bipolar
disorder, and child attachment were not included. The goals of the present paper were to (a) provide brief reviews of epidemiology research on paternal depression and the role of conflict as a mediator between parent and child psychopathology, (b) discuss factors that may contribute to inconsistencies in the literature, and (c) report and discuss results of a meta-analysis of this literature.

1.1. Epidemiology of paternal depression

According to epidemiological studies, twice as many women as men meet criteria for a diagnosis of major depression or dysthymia (Eaton et al., 1997; Kessler, McGonagle, Swartz, Blazer, & Nelson, 1993; Weissman, Bruce, Leaf, Florio, & Holzer, 1991) or subclinical depressive symptoms (Nolen-Hoeksema, 1987). This adult gender difference has been found not only in samples of European-Americans, but also among African-Americans and Latinos (Blazer, Kessler, McGonagle, & Swartz, 1994) and Europeans (Angst et al., 2002). Despite this gender difference, however, epidemiological research also indicates that a sizable number of men experience depression at some point in their lives. The National Comorbidity Survey (Blazer et al., 1994; Kessler et al., 1993) estimated the 30-day and lifetime prevalence of major depressive episode in adult males to be 3.8% and 12.7%, respectively. Subclinical depressive symptoms are even more common (Boyd & Weissman, 1981; Comstock & Helsing, 1976). Rates of depression in adult males may not be uniform across age, ethnic, and socioeconomic groups, however. Only one study to date, the National Comorbidity Survey (Blazer et al., 1994), has reported within group variation on rates of depression in adult males. Although these findings are limited to one sample and need to be replicated, they nevertheless are helpful in describing basic prevalence rates of depression in males.

Similar to patterns of depression in women, rates of depression in men younger than 45 years of age are greater than rates for males 45 years and older (Blazer et al., 1994). Thus, those in the age group at greatest risk for depression are also those in the age group likely to have children. This trend appears to be consistent across different ethnic groups, including African-Americans and Hispanics, and across different socioeconomic groups. Additionally, the age trend in depression rates is noteworthy in light of the greater risk for suicide in males, particularly those younger than 34 years (Blair-West, Cantor, Mellsop, & Eyeson-Annan, 1999).

The National Comorbidity Survey (Blazer et al., 1994) did note ethnic and socioeconomic status group differences in lifetime prevalence of depression in men. European-American and Hispanic men showed similar rates, whereas African-American men showed the lowest prevalence of depression (Blazer et al., 1994). Additionally, depression was more common in males with lower educational attainment and low incomes. Although studies of the relation between paternal depression and child psychopathology should examine the possible contribution of these demographic variables, few have done so. In summary, despite the greater prevalence of depression in women than men, males still experience significant amounts of depression, particularly during the child-rearing years, and therefore should be the focus of investigation.

1.2. Interpersonal context of depression

The interactions of depressed individuals with others not only have important consequences for those with the disorder, but also can affect the well-being of others, particularly the offspring of depressed parents (Coyne, 1976; Coyne, Burchill, & Stiles, 1991; Cummings & Davies, 1999; Gotlib & Hammen,
Parenting is a complex form of social interaction that can pose significant challenges to depressed parents (Coyne, Downey, & Boergers, 1991). The literature on offspring of depressed mothers provides ample illustration of this point. In comparison to well mothers, depressed mothers are more critical (Goodman, Adamson, Riniti, & Cole, 1994; Webster-Stratton & Hammond, 1988), conflictual (Ge et al., 1994; Gotlib & Goodman, 1999; Harnish et al., 1995), and hostile with their children (Conger et al., 1995; Harnish et al., 1995; Patterson & Bank, 1989; Patterson & Dishion, 1988).

Not only have conflictual parent–child interactions been associated with depression in mothers, but these interchanges also have been linked with child internalizing and externalizing symptoms (Conger et al., 1995; Ge et al., 1994; Harnish et al., 1995; Patterson & Bank, 1989). Whereas some parent–child conflict traditionally has been considered developmentally appropriate (Blos, 1962; Laursen & Collins, 1994), higher rates of parent–child conflict more often have been found in clinical samples of families and are associated with a variety of adjustment problems in youth (Patterson, Reid, & Dishion, 1992; Smetana, 1996). Moreover, mother–child conflict has been found to mediate associations between maternal depressive symptoms and child internalizing and externalizing symptoms (Conger et al., 1995; Ge et al., 1994; Harnish et al., 1995).

Similar interaction patterns likely occur for depressed fathers, but less is known about these processes. Patterson and Dishion (1988) proposed that fathers’ depressive symptoms strain personal resources and lead to more irritable and conflictual interactions with children. In tests of this model in normative and high-risk samples, respectively, Conger et al. (1995) and Kane et al. (2002) found that conflict mediated associations between depressive symptoms in fathers and adolescents’ externalizing symptoms. Similar relations between depressive symptoms in fathers and coercive discipline have been found in samples of younger children (Jacob & Johnson, 1997). Thus, similar to what has been found for depressed mothers, conflict also may mediate relations between father and child psychopathology (Conger et al., 1995; Kane, Garber, & Kaminski, 2002).

An important limitation of this literature, however, is its cross-sectional nature. As with much of the research on the relation between maternal depression, and child outcomes, the literature on depression in fathers and child outcomes is primarily correlational, and therefore no conclusions about the direction of causality can be made. Although depression in fathers may be associated with child psychopathology through mediating parenting behaviors, alternative hypotheses including a reversed direction of effects or third variables are plausible. Thus, possible explanation may include “child effects,” genetics, shared familial stressors, fathers’ impact on maternal functioning, modeling of cognitive styles or affective responses, or other unmeasured third variables.

1.3. Discrepancies in the literature

Fundamental to the literature is the assumption that paternal depression represents a form of “caretaking casualty” (Sameroff & Chandler, 1975) that may negatively impact children. Using this framework, several studies have found links between paternal depression, children’s socioemotional development, and father–child relationships (Compas, Phares, Banez, & Howell, 1991; Conger et al., 1995; Jacob & Johnson, 1997; Marchand & Hock, 1998). However, some studies have failed to find such relations (Forehand, Long, Brody, & Fauber, 1986; Ivens & Rehm, 1988). Several differences across studies may contribute to these discrepant findings, such as sampling characteristics of fathers and children, symptom versus diagnostic assessment of fathers and children, types of informants about
criterion data, and publication bias. Any of these variables may moderate relations between paternal depression and the effect size categories.

Investigations of clinical samples may differ significantly from normative samples in the magnitude of effect sizes. This difference could either increase or decrease the magnitude of observed effect sizes. For instance, samples that select fathers based on their clinical levels of psychopathology might be more likely to yield larger effect sizes due to shared genetic influences, more disturbed parenting, or both. However, the reverse also could be true. Clinical samples could have restricted ranges thus attenuating relations between paternal depression and child psychopathology. Similar processes may operate in clinical and risk samples of children. The greater distribution of child psychopathology in these samples relative to low-risk samples may increase the probability of uncovering significant associations whereas restriction of range in clinical groups may reduce effect sizes. Although the manner in which sampling strategies impact effect sizes initially may be unclear, it is reasonable to hypothesize that sample characteristics may account for systematic variation in any of the three outcome categories examined in the present paper.

In addition to sample characteristics, methodological factors may account for systematic variance in the results. First, some studies have assessed depression in fathers using categorical diagnostic criteria (Jacob & Leonard, 1986), whereas others have measured symptoms with self-report questionnaires (Compas et al., 1991; Webster-Stratton, 1988). Although using diagnoses can be informative about depressive disorders, power to detect effects may be compromised with such dichotomous data. Similar problems can occur with respect to the measurement of child outcomes. Although some studies have used questionnaires to assess child symptoms (Achenbach, 1991), several studies have assessed child psychopathology with diagnostic criteria (Cunningham, Benness, & Siegel, 1988; Ivens & Rehm, 1988).

The use of symptom rating scales may result in larger estimates compared to those obtained with dichotomous variables. Continuous measures of paternal and child symptoms may yield larger effects due to the greater variability found with this assessment strategy. However, the use of symptom rating scales may have inherent costs in measurement related to the “apples and oranges” criticism of meta-analysis. Whereas high scores on symptom rating scales, such as the Beck Depression Inventory (BDI), may indicate the presence of disorder, they also may indicate the presence of nonspecific negative affective symptoms (Gotlib, 1984). Not all individuals who score high on the BDI, for example, will necessarily meet criteria for a diagnosis of depression. Similar processes may operate when using any continuous scale to measure psychopathology in adults or children. Thus, the potential for greater variability in continuous measures of psychopathology may result in effects larger than those found with dichotomous measures of predictor and criterion variables. However, this sensitivity may be achieved at the cost of not measuring clinical levels of psychiatric disorder.

Another methodological factor that can impact individual study effects concerns the informant about child outcomes. A review of the extant literature shows that three primary informant categories have been utilized to measure child psychopathology: parent report, child self-report, and teacher report. The magnitude of relations between paternal depression and child symptoms has been found to differ as a function of reporters about child behavior (Conger et al., 1995; Ge et al., 1994). Whereas some studies have examined parents’ report of symptoms or conflict, others have used multiple indicators of child outcomes, including parent, self, and teacher reports. Studies assessing child outcomes based on parent report only may systematically differ from those using multiple raters of child functioning.

The use of meta-analysis can help clarify our understanding of factors related to variability in effect sizes. First, meta-analysis yields an overall estimate of effect size and a more powerful test of statistical
significance of this effect by aggregating estimates from different studies (Weiss, Catron, Harris, & Phung, 1999). Pooling estimates from independent studies allows us to determine if nonsignificant findings were a function of low sample size in individual studies. Second, it is possible to test sample and methodological factors as moderators of effects across different studies. Finally, by documenting significant effect sizes, we can begin to establish a more thorough understanding of the role of parental depression in child development and to develop models to test mediating and moderating processes.

2. Method

2.1. Literature search

The first step in obtaining relevant studies was to conduct searches of bibliographic databases (PsycInfo, ERIC, Medline, and Dissertation Abstracts). The following key words were used: father(s), paternal, depression, depressive, child, adolescent, psychopathology, internalizing, externalizing, conflict, and discipline. Searches were made of reports from 1975 to 2000. Subsequently, the reference sections of all obtained studies were scanned for other potential studies. In addition, relevant journals published between 1980 and 2000 (Child Development, Developmental Psychology, Journal of Abnormal Child Psychology, Development and Psychopathology, Journal of Clinical Child Psychology, Journal of Family Psychology, Journal of Child Psychology and Psychiatry, and Journal of Consulting and Clinical Psychology) were reviewed. Furthermore, qualitative literature reviews on relations between father and child forms of psychopathology were perused for potential study candidates. Finally, a list of references and a request for additional data and/or citations were sent to authors who published relevant articles or who conduct research on parental depression, child psychopathology, and paternal contributions to child development.

Several criteria were established initially for determining eligibility in the meta-analysis. To be included, studies had to assess either depressive symptoms or make a diagnosis of depression in fathers based on fathers’ report and assess child or adolescent internalizing symptoms or diagnoses, externalizing symptoms or diagnoses, or father–child conflict. Only studies printed in English were used. In an effort to minimize the “file drawer” problem (Orwin, 1994; Rosenthal, 1979) and obtain optimal estimates of effect sizes, both published and unpublished studies were eligible for inclusion in the present study. Reports from peer-reviewed or nonreviewed journals, conference presentations, dissertations, unpublished data and materials obtained directly from the author, and book chapters, including quantitative information, were acceptable. Studies were not excluded based on the quality of methodology as long as the appropriate information was included to calculate an effect size.

Together, 32 papers with 49 effect sizes met eligibility standards. However, 9 papers were excluded due to incomplete information necessary to calculate an effect size, thereby resulting in a final sample of 23 papers with 40 effect sizes. For example, some studies used general indicators of child psychopathology, such as the Total Problem Score on the Child Behavior Checklist (CBCL). Studies were excluded if the primary informant about fathers’ psychopathology was not the father himself (e.g., Beardslee et al., 1987; Pfiffner et al., 1999). Additionally, several other studies that examined the constructs of interest were excluded based on the presentation of analyses. These studies tested paternal depression as a predictor of one of the three effect size categories in a hierarchical multiple regression equation without reporting bivariate correlations. Because of the problems inherent in computing
standard errors of beta weights in multiple regression (Lipsey & Wilson, 2001), data from these studies could not be used. Additionally, several published studies presented results of interest on the same sample. Only one study using the sample was retained to avoid issues of statistical dependencies among effect sizes. The final sample consisted of 17 studies of paternal depression and child internalizing symptoms, 17 studies of paternal depression and child externalizing symptoms, and 6 studies of paternal depression and father–child conflict.

2.2. Data preparation

A coding scheme was developed to extract relevant data from the research reports. Coded information included study descriptors (e.g., year of publication and type of publication), sample characteristics (e.g., ethnicity of participants, child age, SES, and normative vs. clinical samples of fathers and children), information on depression measurement (e.g., symptom rating scales and diagnostic criteria), information on dependent variables (e.g., symptom rating scales and diagnostic criteria), informant (self, parent, or teacher), and effect size (e.g., magnitude and type of statistic). Most studies reported bivariate correlations between the variables of interest. However, for studies that did not report correlations, Wilson’s (1996) Excel program for effect size calculation was used to convert the appropriate information into correlations. For example, two studies reported means and standard deviations of child internalizing and externalizing symptoms within samples of depressed and nondepressed fathers. These data were converted to correlations.

The correlation coefficient \( r \) was used as the effect size indicator in the present review. Given the difficulty with calculating standard errors of correlation coefficients, each \( r \) was converted to Fisher’s \( z \) as suggested by Lipsey and Wilson, (2001) and converted back to \( r \) following analyses. In the present study, only one effect size per construct per study was calculated to avoid issues of statistical dependencies among effect sizes. Consequently, for studies reporting analyses between fathers’ depression and multiple indicators of the same child outcome, the computed effect sizes and sample sizes were combined into a weighted average. For all cases, positive effect sizes indicate worse outcomes for children. Despite efforts to gather all published and unpublished studies and/or data, it is possible that not all studies were included in the present study. Influences of possible excluded studies are considered later.

Interrater reliability was assessed for 25% of the studies across two coders. Following Orwin’s (1994) guidelines, the appropriate indices of interrater reliability were determined for each item depending upon that item. For items including author’s affiliation, publication year, and publication type, the percent agreement rate was computed. The agreement per item ranged from 85% to 100% with an average across all study descriptors equaling 97%. Next, for items, such as socioeconomic status, race, and other nominal or ordinal items, \( \kappa \) was calculated to assess reliability while controlling for agreement by chance. These \( \kappa \) values ranged from .74 to 1.0. Finally, reliabilities for effect sizes were computed as the correlation between the two coders. Pearson’s \( r \) for the effect sizes between the two coders equaled .94.

Separate categories of dependent variables were derived based on the theoretical similarity of measures. Three outcome categories were examined: internalizing symptoms [standardized measures of internalizing, such as the CBCL, or psychiatric assessment of depression or anxiety], externalizing symptoms [standardized measures of externalizing, such as the CBCL, or psychiatric assessment of conduct problems], and father–child conflict (observations or self-report questionnaires of father–child conflict and/or fathers’ harsh discipline practices). Effect sizes based on related constructs, such as total
problems as assessed on the CBCL, which are composed of or correlate with the indices of interest were not included. Agreement between coders on the outcome categories was high (κ=.90).

3. Results

Descriptive statistics for each effect size category (internalizing, externalizing, and conflict) are presented in Tables 1–3. Analyses of the distributions of effect sizes for internalizing and externalizing symptoms indicated the presence of one sample size outlier per effect size category and one correlation outlier in the internalizing category. To prevent analyses from being disproportionately affected by extreme cases, all three outliers were winsorized to values two standard deviations above the mean for each effect size category.

Each effect size value was computed by weighting the observed effect size by its inverse variance weight (in this case, the sample size minus 3; \(n - 3\)). Confidence intervals were computed based on the standard error of the mean effect size and the critical value from a \(z\) distribution associated with \(\alpha=.05\) (Shadish & Haddock, 1994). Mean effect sizes were computed under a random effects assumption that provides a more conservative test of statistical significance. The noniterative method of moments was used to derive the variance component needed to compute parameter estimates (Lipsey & Wilson, 2001).

As shown in Table 4, the mean effect sizes for the three construct categories were .24 for internalizing symptoms, .19 for externalizing symptoms, and .20 for conflict. All three mean effect sizes were significant at \(p < .05\) and did not differ statistically as evidenced by their overlapping confidence intervals. By Cohen’s (1988) standards, these effect sizes are considered modest.

**Table 1**

<table>
<thead>
<tr>
<th>Publication</th>
<th>Data source</th>
<th>Total (N)</th>
<th>Average age of children (years)</th>
<th>Effect size ((r))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Billings &amp; Moos, 1985</td>
<td>Journal</td>
<td>43</td>
<td>13.5</td>
<td>.26</td>
</tr>
<tr>
<td>Carro et al., 1993</td>
<td>Journal</td>
<td>70</td>
<td>3.0</td>
<td>.32</td>
</tr>
<tr>
<td>Compas et al., 1991</td>
<td>Journal</td>
<td>27</td>
<td>12.0</td>
<td>.35</td>
</tr>
<tr>
<td>Forehand et al., 1986</td>
<td>Journal</td>
<td>46</td>
<td>13.5</td>
<td>.07</td>
</tr>
<tr>
<td>Forehand &amp; Smith, 1986</td>
<td>Journal</td>
<td>33</td>
<td>12.8</td>
<td>.29</td>
</tr>
<tr>
<td>Gallimore &amp; Kurdek, 1992</td>
<td>Journal</td>
<td>35</td>
<td>13.6</td>
<td>.28</td>
</tr>
<tr>
<td>Ge et al., 1995</td>
<td>Journal</td>
<td>451</td>
<td>12.7</td>
<td>.10</td>
</tr>
<tr>
<td>Ivens &amp; Rehm, 1988</td>
<td>Journal</td>
<td>38</td>
<td>10.0</td>
<td>.23</td>
</tr>
<tr>
<td>Jacob &amp; Leonard, 1986</td>
<td>Journal</td>
<td>45</td>
<td>14.0</td>
<td>.32</td>
</tr>
<tr>
<td>Kane et al., 2002</td>
<td>Manuscript</td>
<td>78</td>
<td>11.8</td>
<td>.35</td>
</tr>
<tr>
<td>Marchand &amp; Hock, 1998</td>
<td>Journal</td>
<td>46</td>
<td>4.0</td>
<td>.36</td>
</tr>
<tr>
<td>Norrell, 1984</td>
<td>Dissertation</td>
<td>24</td>
<td>15.0</td>
<td>.27</td>
</tr>
<tr>
<td>Reider, 1991</td>
<td>Dissertation</td>
<td>90</td>
<td>4.4</td>
<td>.60</td>
</tr>
<tr>
<td>Renk et al., 1999</td>
<td>Journal</td>
<td>150</td>
<td>13.7</td>
<td>.23</td>
</tr>
<tr>
<td>Schaughency &amp; Lahey, 1985</td>
<td>Journal</td>
<td>41</td>
<td>NA</td>
<td>.20</td>
</tr>
<tr>
<td>Webster-Stratton, 1988</td>
<td>Journal</td>
<td>85</td>
<td>4.3</td>
<td>.09</td>
</tr>
<tr>
<td>Weissman, 2000</td>
<td>Unpublished data</td>
<td>35</td>
<td>13.1</td>
<td>.07</td>
</tr>
</tbody>
</table>
Next, analyses were conducted to test the hypothesis that variation in effect sizes within the three construct categories was related systematically to study characteristics. First, the Q statistic (Hedges & Olkin, 1985), a test of homogeneity of variance under fixed effects assumptions, was computed for each outcome category. Hedges$^Q$ is distributed as a chi-square with $k/C0$ degrees of freedom, where $k$ is the number of effect sizes. The $Q$ statistic indicates whether the variance in the sample of effect sizes can be accounted for by the sampling error associated with the individual studies included in the computation of the mean effect size. A significant $Q$ indicates that sampling error alone is insufficient to account for the variance in the effect size distribution and that other factors may explain this variability. The $a$ level for the following $Q$ tests was set at $p < .05$. The $Q$ statistics for both the internalizing and externalizing construct categories were significant at the .05 $a$ level (see Table 4).

Results indicated that the assumption of homogeneity of variance within the internalizing and externalizing categories must be rejected. Therefore, the variation in the distributions of effect sizes is greater than would be expected by chance alone. However, the $Q$ statistic for conflict was nonsignificant.

<table>
<thead>
<tr>
<th>Publication</th>
<th>Data source</th>
<th>Total $N$</th>
<th>Average age of children (years)</th>
<th>Effect size ($r$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Billings &amp; Moos, 1985</td>
<td>Journal</td>
<td>43</td>
<td>13.5</td>
<td>.35</td>
</tr>
<tr>
<td>Carro et al., 1993</td>
<td>Journal</td>
<td>70</td>
<td>3.0</td>
<td>.32</td>
</tr>
<tr>
<td>Compas et al., 1991</td>
<td>Journal</td>
<td>27</td>
<td>12.0</td>
<td>.35</td>
</tr>
<tr>
<td>Conger et al., 1995</td>
<td>Journal</td>
<td>75</td>
<td>11.8</td>
<td>.20</td>
</tr>
<tr>
<td>Conger et al., 1995</td>
<td>Journal</td>
<td>196</td>
<td>12.7</td>
<td>.02</td>
</tr>
<tr>
<td>Cunningham et al., 1988</td>
<td>Journal</td>
<td>26</td>
<td>9.8</td>
<td>.00</td>
</tr>
<tr>
<td>Dean &amp; Jacobson, 1982</td>
<td>Journal</td>
<td>90</td>
<td>8.6</td>
<td>.20</td>
</tr>
<tr>
<td>DeKlyen et al., 1998</td>
<td>Journal</td>
<td>38</td>
<td>4.0</td>
<td>.07</td>
</tr>
<tr>
<td>Forehand et al., 1986</td>
<td>Journal</td>
<td>46</td>
<td>13.5</td>
<td>−.12</td>
</tr>
<tr>
<td>Jacob &amp; Leonard, 1986</td>
<td>Journal</td>
<td>45</td>
<td>14.0</td>
<td>.16</td>
</tr>
<tr>
<td>Marchand &amp; Hock, 1998</td>
<td>Journal</td>
<td>46</td>
<td>4.0</td>
<td>.18</td>
</tr>
<tr>
<td>Kane et al., 2002</td>
<td>Manuscript</td>
<td>78</td>
<td>11.8</td>
<td>.42</td>
</tr>
<tr>
<td>Reider, 1991</td>
<td>Dissertation</td>
<td>90</td>
<td>4.4</td>
<td>.48</td>
</tr>
<tr>
<td>Renk et al., 1999</td>
<td>Journal</td>
<td>150</td>
<td>13.7</td>
<td>.21</td>
</tr>
<tr>
<td>Schaughency &amp; Lahey, 1985</td>
<td>Journal</td>
<td>41</td>
<td>NA</td>
<td>.07</td>
</tr>
<tr>
<td>Webster-Stratton, 1988</td>
<td>Journal</td>
<td>85</td>
<td>4.3</td>
<td>−.02</td>
</tr>
<tr>
<td>Weissman, 2000</td>
<td>Unpublished data</td>
<td>35</td>
<td>13.1</td>
<td>.00</td>
</tr>
</tbody>
</table>

Table 2
Studies examining relation between fathers’ depression and child/adolescent externalizing symptoms/diagnoses

<table>
<thead>
<tr>
<th>Publication</th>
<th>Data source</th>
<th>Total $N$</th>
<th>Average age of children (years)</th>
<th>Conflict ($r$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conger et al., 1995</td>
<td>Journal</td>
<td>75</td>
<td>11.8</td>
<td>.23</td>
</tr>
<tr>
<td>Conger et al., 1995</td>
<td>Journal</td>
<td>196</td>
<td>12.7</td>
<td>.21</td>
</tr>
<tr>
<td>Kane et al., 2002</td>
<td>Manuscript</td>
<td>81</td>
<td>11.8</td>
<td>.39</td>
</tr>
<tr>
<td>Powers, 1984</td>
<td>Dissertation</td>
<td>36</td>
<td>10.0</td>
<td>.18</td>
</tr>
<tr>
<td>Webster-Stratton, 1988</td>
<td>Journal</td>
<td>85</td>
<td>4.3</td>
<td>.01</td>
</tr>
</tbody>
</table>
This finding indicates that the assumption of homogeneity of the distribution of effect sizes within the conflict category cannot be rejected. Hence, the variation in the distribution of effect sizes is not greater than would be expected solely by chance.

Next, analyses were conducted to determine systematic influences on the effect sizes within the internalizing and externalizing categories given the significant $Q$ found for each of these categories. First, an analog of a one-way ANOVA as delineated by Hedges and Olkin (1985) was conducted to assess whether demographic variables accounted for variation in child outcomes. This analysis partitions the total homogeneity $Q$ for each outcome category into the portion explained by demographic variables ($Q$ between) and the residual within groups portion of variance ($Q$ within). Separate analyses were conducted for each demographic variable on each outcome. No significant effects of socioeconomic status, ethnic composition, or gender composition of the studies were noted.

Next, weighted regression for meta-analysis as outlined by Hedges and Olkin (1985) was conducted to model the variability in the internalizing and externalizing effect size distributions. Separate weighted

### Table 4
Mean effect sizes, confidence intervals, and $Q$ statistics for each child outcome domain

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of independent effect sizes and total sample ($k, N$)</th>
<th>Under random effects model</th>
<th>Under fixed effects model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean effect size</td>
<td>95% Confidence interval</td>
</tr>
<tr>
<td>Internalizing</td>
<td>(17, 1284)</td>
<td>.24**</td>
<td>.15–.34</td>
</tr>
<tr>
<td>Externalizing</td>
<td>(17, 1157)</td>
<td>.19**</td>
<td>.11–.28</td>
</tr>
<tr>
<td>Conflict</td>
<td>(6, 565)</td>
<td>.20**</td>
<td>.09–.30</td>
</tr>
</tbody>
</table>

* $p<.05$.
** $p<.01$.

### Table 5
Results of weighted regression analyses for internalizing and externalizing effect size outcome categories

<table>
<thead>
<tr>
<th>Outcome category</th>
<th>Predictor variables</th>
<th>Beta</th>
<th>S.E. Beta</th>
<th>$\beta$</th>
<th>Model $Q$ ($df$)</th>
<th>Residual $Q$ ($df$)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internalizing</td>
<td>Fathers’ measure of depression</td>
<td>−.20</td>
<td>.15</td>
<td>−.31</td>
<td>15.88**</td>
<td>8.42 ns (12)</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>Sampling method</td>
<td>−.21</td>
<td>.07</td>
<td>−.70**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Child’s measure of internalizing</td>
<td>−.31</td>
<td>.14</td>
<td>−.45*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Informant</td>
<td>−.25</td>
<td>.15</td>
<td>−.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Externalizing</td>
<td>Fathers’ measure of depression</td>
<td>−.09</td>
<td>.16</td>
<td>−.17</td>
<td>5.33 (4)</td>
<td>11.85 ns (12)</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>Sampling method</td>
<td>−.11</td>
<td>.07</td>
<td>−.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Child’s measure of externalizing</td>
<td>.01</td>
<td>.15</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Informant</td>
<td>−.12</td>
<td>.10</td>
<td>−.31</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p<.05$.
** $p<.01$. 
regressions were conducted in which internalizing and externalizing each served as the criterion variable to be predicted from the study variables. As shown in Table 5, it was possible to account for a significant portion of variance in the distribution of effect sizes comprising the internalizing category. For internalizing, the sampling method \( (\beta = -0.70, p < 0.01) \) and the measure of child internalizing symptoms \( (\beta = -0.45, p < 0.05) \) emerged as significant predictors. Larger effect sizes were found in studies using community samples and continuous measures of child internalizing symptoms. Measures of fathers’ depression and informants about child outcomes were not significant predictors of variation, however. The study predictor variables accounted for a significant portion of the variance such that a nonsignificant \( Q \) residual emerged. Thus, the study variables contributed to the heterogeneity in effect sizes, and the random variation between studies for the internalizing category was explained by the model. For externalizing, the study variables could not account for the systematic variation in the distribution of effect sizes. Therefore, variables not measured in the present study account for the heterogeneity in effect sizes. Possible variables are considered later in the discussion.

Although meta-analysis has several distinct advantages over qualitative reviews, the methodology is not without potential biases. Foremost is the potential upward bias of the mean effect size that may occur from systematic biases in available research. In an effort to address the “file drawer” problem, researchers (Orwin, 1983; Rosenthal, 1979) have suggested reporting the fail-safe \( N \) to gauge the magnitude of the impact of unobtained studies on the results. Following the formula established by Orwin (1983), the necessary number of additional studies required to make the mean effects sizes of internalizing, externalizing, and conflict nonsignificant was generated. Using Cohen’s (1988) standard of .10 as an effect size too small to warrant consideration, 62 studies of internalizing, 47 studies of externalizing, and 20 studies of father–child conflict, all with null effects, would be required to reduce these mean effect sizes to nonsignificant. However, the size of the fail-safe \( N \) and the number of studies obtained after a rigorous search, particularly in light of small body of literature, suggest that the mean effect size for each outcome category is fairly robust.

4. Discussion

Recent qualitative reviews (Phares, 1997; Phares, Duhig, & Watkins, 2002) have argued that depression in fathers is associated with a variety of negative emotional and behavioral outcomes in children. In an effort to build on these reviews, the primary purpose of this meta-analysis was to assess the direction and magnitude of associations between depression in fathers and three outcomes: child internalizing and externalizing symptoms, and father–child conflict. Consistent with Phares (1997, 2002), results from the present review indicated that depression in fathers is significantly positively associated with internalizing and externalizing symptoms in children and father–child conflict. Although modest in the proportions of variance accounted for, each mean effect size was statistically significant and similar to those found by Beck (1999) for relations between maternal depression and child behavior problems.

Second, the magnitude of effects varied systematically as a function of study characteristics. Meta-analytic techniques were used to assess the variability in each effect size category and to determine the role of study characteristics, such as sampling methods and types of measures, in accounting for significant variation. Tests of the homogeneity of variance revealed significant variation in the
internalizing and externalizing distributions, and study characteristics were tested as moderators of the relation between paternal depression and children’s internalizing and externalizing problems. For internalizing, symptom rating scales and community samples were associated with larger effect sizes than diagnoses and clinical samples of children and/or parents, respectively. For externalizing, study characteristics did not account for heterogeneity, thus indicating that factors other than those examined here account for this variability.

4.1. Internalizing symptoms

Several different explanations may account for the greater association between depression in fathers and internalizing psychopathology in children in community samples. In the present study, most community samples of parents and children were normative, nonclinical samples. However, community samples of children at high risk for developing psychopathology also were included. For instance, the sample used by Kane et al. (2002) included community recruited high-risk adolescent offspring of mothers with a history of depression and for comparison, a sample of adolescents of mothers with no history of depression. Although recruited from the community, samples like these would be expected to yield larger estimates of effect sizes given the risk for psychopathology compared to normative community samples.

In contrast, restriction of range in clinical samples may have yielded smaller effects. Selecting participants at the clinically significant end of the continuum may have obscured more subtle variation in both paternal and child psychopathology. The greater homogeneity of psychopathology in clinical samples of fathers and children may have reduced variability, thus leading to smaller effect sizes for studies in which these samples were utilized. Although some researchers have developed statistical procedures for correcting for attenuation due to range restriction in meta-analysis (Hunter & Schmidt, 1990, 1994), these methods were not utilized in the present study. As Rosenthal (1994) notes, such adjustments must be applied with great caution, as they are based on uncertain assumptions and estimates of other parameters, and hence run the risk of inadvertently increasing bias to the extent to which these assumptions and estimates are themselves inaccurate. Thus, the process of correcting for range restriction may be more problematic than the artifact itself.

Effect sizes also were greater in studies using symptom measures of internalizing problems in children. This finding is especially intriguing because the level of severity captured by diagnoses of depression in children might be expected to relate to greater impairments in paternal functioning. However, in contrast to diagnoses, continuous measures of child internalizing symptoms may be more sensitive to variability in child functioning, thereby leading to correlations of greater magnitude. Rather than dichotomizing participants as diagnosed or not, continuous measures of internalizing problems do not reduce to zero scores for participants who fall between clinical significance and no disorder. Hence the middle range of symptoms is preserved, thereby increasing variability and the magnitude of the correlations.

In contrast to the findings for child symptom measures, methods for assessing paternal depression were not related to heterogeneity. Symptom rating scales and diagnoses of depression in fathers were associated with child and adolescent internalizing symptoms to the same degree. Consequently, restriction of range in assessment of depression in fathers does not appear to be related to the magnitude of association between child and father functioning.

An alternative explanation may be that differences between clinical and subclinical depressive symptoms in fathers are not related to differences in children’s functioning. That is, subclinical levels of
depressive symptoms may be sufficient for the occurrence of internalizing symptoms in children. In other words, there may be a threshold effect, such that subclinical levels of depressive symptoms in fathers may be sufficient to show a relation to children’s internalizing symptoms. However, it is important to keep in mind that data from the present study were correlational and do not imply a direction of effect from parent to child.

Type of informant about children’s symptoms did not significantly account for heterogeneity. Thus, paternal depression correlated with child outcomes about comparably when reported by multiple and different informants despite the fact that different reporters about children’s symptoms typically share only a small to moderate percent of variance, particularly on reports of internalizing symptoms (Achenbach, McConaughy, & Howell, 1987).

Methodological reasons also may explain the lack of variation as a function of informants, including the limited variability of data from different informants and the manner in which different types of informants were coded. Given the statistical dependencies among multiple effect sizes for the same construct from the same study, only one effect size per construct per study was obtained. In studies utilizing multiple informants about child outcomes, this process necessitated the calculation of a weighted average effect size for a specific construct. The effect of any one informant may have been reduced to form the average of different reporters. For example, the effect size from the Ge et al. (1994) study was the weighted mean of parent, observation, and self-reports whereas effect sizes from the Gallimore and Kurdek (1992) and the Forehand and Smith (1986) studies were parent report only. Interestingly, the weighted average of self, parent, and observational indices of internalizing symptoms in the Ge et al. study yielded an effect size estimate of a magnitude similar to that of parent report in that study alone. Averaging of parent and other reports of internalizing symptoms in Ge et al. may have resulted in a more valid estimate of the population correlation, but at the same time may have reduced variability related to specific informants. Such consequences of creating weighted mean effect sizes for studies in the internalizing category are not limited to the paper by Ge et al.

4.2. Externalizing symptoms

With regard to child externalizing symptoms, study characteristics did not account for the heterogeneity in the studies. Thus, other variables not assessed in the present review likely explain differences among studies. Two possible moderators are marital discord and child physical abuse. The quality of interactions between parents as well as parents’ physically abusive discipline have been found to be risk factors for child behavior problems (Cummings & Davies, 1994; Emery, 1982; Emery, Fincham, & Cummings, 1992; Erel & Burman, 1995; Grych & Fincham, 1990) and correlates of parental depression (Downey & Coyne, 1990; Hops, 1992; Phares et al., 2002). Downey and Coyne (1990) concluded that families in which high levels of interparental conflict co-occur with parental depression also have children with greater externalizing symptoms. Although milder forms of negative parent–child interactions, such as conflict, have been found to be related to parental depression (Conger et al., 1995; Dodge, Bates, & Pettit, 1990, 1994; Patterson, 1996; Patterson & Dishion, 1988), in the absence of marital discord or physical abuse the relation between parental depression and child externalizing symptoms may be weaker (Downey & Coyne, 1990; Emery, 1982; Fendrich, Warner, & Weissman 1990). Most studies examining relations between marital discord or physical abuse and child outcomes have been with mothers, although similar processes may operate for depressed fathers. The possible
contribution of paternal depression to child externalizing symptoms in the context of marital conflict or physical abuse should be studied further.

4.3. Father–child conflict

In the conflict domain, sampling error alone was sufficient to account for the variation in effect sizes in these studies. Two explanations for this finding are possible. First, there simply may be no systematic variation in effect sizes. Thus, different sampling and measurement strategies would be expected to yield effect sizes of similar magnitude and direction.

Alternatively, too few studies may exist as yet to capture the potential range of correlations. This argument seems more probable at this point. In contrast to the other effect size categories, a small number of studies \( (n = 6) \) were available to estimate the effect size between paternal depression and conflict. Although variation existed in sampling strategies and in the use of questionnaire methods versus observations of conflict, perhaps a larger sample of studies would have produced systematic variation beyond that explained by sampling error.

Additionally, there are several theoretical reasons to expect heterogeneity of variance in the internalizing and externalizing categories and homogeneity of variance for conflict. Previous research indicates that higher rates of parent–child conflict are found in clinical samples of children and adolescents (Kashani, Burbach, & Rosenberg, 1988; Puig-Antich, Kaufman, Ryan, & Williamson, 1993), clinical samples of parents (Smetana, 1996), and families of low socioeconomic status (Dodge et al., 1994; Pinderhughes, Dodge, Bates, Pettit, & Zelli, 2000). Given the limited parenting repertoires of depressed parents and greater use of coercive discipline in low-income families, fathers in these samples would be expected to utilize more negative and conflictual interaction strategies with children than normative groups of fathers. However, the variability in the clinical and socioeconomic status of fathers in studies comprising the conflict category was limited.

Whereas mother–child conflict may be more frequent than father–child conflict (Montemayor, 1982), Forehand et al. (1986) speculated that father–child conflict may be more harmful to children’s behavioral development than conflict with mothers. Both Forehand et al. and Johnson (1987) found that conflict with fathers was associated with child behavior problems and delinquency. Moreover, delinquency was negatively related to the degree of closeness in the father–child relationship (Johnson, 1987).

With regard to internalizing symptoms, Ge et al. (1994) and Cole and McPherson (1993) showed that conflict was related to current levels of child depressive symptoms. Similarly, Jacob and Johnson (1997) revealed that depressed fathers’ positive affect expressions and verbal interactions with children were negatively related to both externalizing and internalizing symptoms, and father’s approval was inversely related to child depressive symptoms. Thus, evidence from a variety of studies suggests that father–child interactions are related to behavioral and emotional symptoms in children.

4.4. Potential mediators

Support also has been found for components of a mediational model of father–child conflict (Conger et al., 1995; Ge, Conger, Lorenz, Shanahan, & Elder, 1995; Kane et al., 2002). Although limited given their cross-sectional designs, these studies provide initial support for the hypothesis that both maternal and paternal depression may impact children’s psychopathology partially through disrupted parent–child relationships. Although limited in heterogeneity, effect sizes in the conflict studies were consistently
positive. These individual studies, and the weighted mean effect size derived from the present meta-analysis, provide support for a potential mediational model of conflict in the relation between fathers’ depression and psychopathology in their offspring. Although not a focus of the present review, empirical evidence of an association between family conflict and children’s externalizing and internalizing psychopathology exists in the literature (e.g., Conger et al., 1995; Ge et al., 1994; Patterson et al., 1992). Conflict may be linked with externalizing and internalizing symptoms through inappropriate modeling of coercive and hostile behavior (Patterson & Dishion, 1988; Patterson et al., 1992), feelings of helplessness and loss of control related to an unpredictable environment (Ge et al., 1994), or emotional dysregulation (Cummings & Davies, 1999; Davies & Windle, 1997). However, additional prospective research is needed to determine how robust and under what circumstances this mediational model is supported across time.

An alternative hypothesis is that children’s genetic vulnerabilities contribute to their psychopathology. Family studies have found that offspring of depressed parents have higher rates of depressive disorders (Beardslee, Versage, & Gladstone, 1998; Downey & Coyne, 1990) and other psychiatric diagnoses compared to children of nondepressed parents (Weissman, Prusoff et al., 1984). Overall, the risk for developing depression in first degree relatives of individuals diagnosed with depression is significantly greater (20–25%) than for the general population (7%) (Goodman & Gotlib, 1999; Nurnberger, Goldin, & Gershon, 1996). Risk for the development of psychopathology in children is even greater when both parents, compared to one parent, meet criteria for a psychiatric disorder (Dierker, Merikangas, & Szatmari, 1999; Merikangas, Weissman, Prusoff, & John, 1988; Phares et al., 2002).

There also may be genetic effects on environments (e.g., parenting, marital relationship; Plomin, 1994). For example, studies of adopted twins reared apart have found that monozygotic twins were more similar than dizygotic twins on measures of the family environment and quality of parenting (O’Connor, Hetherington, Reiss, & Plomin, 1995; Plomin, Reiss, Hetherington, & Howe, 1994). Thus, genetics may contribute to ecological processes that are associated with parent and child psychopathology. Plomin (1994) noted, however, that although some variability in environments can be attributed to genes, a significant proportion cannot be accounted for by inherited characteristics and thus must be attributed to properties of the environment. Moreover, genetic vulnerabilities occur within environmental contexts that mediate and/or interact with inherited dispositions to produce phenotypes (Rutter, 2000). Thus, although genetics may account for a significant proportion of the variance between parent and child psychopathology, as well as patterns of parent—child interactions, environmental processes are necessary for the expression of inherited vulnerabilities (Rutter, 1999; Rutter & Sroufe, 2000).

4.5. Conclusions and future directions

In summary, the significant effect sizes found in the present review are particularly compelling given the lack of research on samples of depressed fathers. The majority of studies were not designed to assess paternal depression and its correlates, but instead were developed to measure correlates of maternal depression or child psychopathology in community and clinical samples. Additional research on clinical samples of depressed fathers may further document parent and child functioning at clinically significant levels. But at present, such studies are limited and likely difficult to accomplish given the lower base rate of depression in men compared to women. Given patterns of assortative mating (McLeod, 1993; Merikangas & Spiker, 1982), the use of high-risk community samples of offspring of depressed mothers may be one fruitful strategy for studying the effects of paternal depression as well.
At the correlational level, depression in mothers and fathers is related to maladaptive child outcomes. The majority of studies in the field, including those in the present review, however, have assessed concurrent relations between paternal and child functioning (Compas et al., 1991; Ge et al., 1994; Jacob & Leonard, 1986). Limited longitudinal research suggests that paternal depression is associated with psychopathology in offspring. Both Carro, Grant, Gotlib, and Compas (1993) and Billings and Moss (1985) found that initial levels of depression in fathers were related to child psychopathology at a 1-year follow-up. However, the significant longitudinal relations that have been found must be interpreted cautiously because initial levels of child psychopathology were not controlled in these studies. The current state of the literature precludes conclusions regarding paternal depression as a risk factor for future psychopathology in children. Nevertheless, results from the present meta-analysis suggest that research on the risk for psychopathology in children of depressed fathers is warranted.

An issue that should be examined in future studies is the extent to which paternal depression predicts child outcomes over and above maternal depression. Given the abundance of empirical support for the role of maternal depression as a risk factor for child psychopathology (Beardslee et al., 1998; Goodman & Gotlib, 1999; Gotlib & Goodman, 1999), tests of the incremental contribution of paternal depression are particularly important. Marchand and Hock (1998) found that after controlling for maternal depressive symptoms, paternal depression predicted children’s behavioral and emotional problems. In a sample of adolescents of depressed mothers, Kane et al. (2002) found that fathers’ depressive symptoms were associated with adolescent externalizing and internalizing symptoms after controlling for mother’s history of mood disorders. Thus, studies are beginning to show the possible unique contribution of paternal depression to negative child outcomes, although more tests of this hypothesis, particularly in clinical samples of fathers and children, are needed.

Additionally, paternal depression may indirectly increase the risk for psychopathology in offspring of depressed mothers. In their review, Goodman and Gotlib (1999) suggested that the mental health of fathers may exacerbate the effects of maternal depression on child outcomes through genetic or environmental means. Although limited, some evidence supports this hypothesis. For instance, both Weissman, Prusoff et al. (1984) and Dierker et al., (1999) found that children with two depressed parents were at significantly greater risk for disorder than children with only one depressed parent. Thus, the combined effects of paternal and maternal depression, rather than either parent alone, may be linked with worse child outcomes.

Similar to the literature on maternal depression, studies examining the timing of children’s exposure to paternal depression are limited. To our knowledge, no studies have investigated this “sensitive period” hypothesis. The evidence from the maternal depression literature is equivocal. Whereas some studies have found that earlier exposure (i.e., prenatally or in infancy) to mothers’ depression was related to greater child psychopathology over time (Alpern & Lyons-Ruth, 1993; Essex, Klein, Miech, & Smider, 2001), other studies have found that timing of maternal depression was not associated with worse child outcomes once other characteristics of the mothers’ depression (e.g., chronicity and severity) were considered (Brennan et al., 2000; Martin & Garber, 2003). Thus, the developmental age of children’s first exposure to maternal depression is not consistently related to worse child outcomes. If the maternal depression literature is any indication, timing of children’s exposure to depression in fathers may not predict child problems beyond other characteristics of fathers’ depression. Nevertheless, future studies should conduct explicit tests of this hypothesis.

As noted previously, other processes in addition to parent–child conflict may be related to paternal depression, such as marital discord, which has been found to be related to both maternal depression (Biglan et al., 1985; Downey & Coyne, 1990; Hops et al., 1987) and emotional and behavioral problems
in children (Cummings & Davies, 1994; Cummings & Davies, 1999; Emery, 1982; Grych & Fincham, 1990; Hetherington & Clingempeel, 1992). Given that depression compromises personal resources and can lead to negative interpersonal interactions (Coyne, 1976; Coyne, Downey et al., 1991; Patterson & Dishion, 1988), dysfunctional patterns of problem solving and interactions with spouses may be used by depressed fathers as well, which in turn may result in maladjustment in children. However, cross-sectional and longitudinal evidence of associations between depression in fathers and marital discord, much less the mediating or moderating role of interparental conflict in the relation between paternal and child psychopathology, is lacking (Phares, 1997).

Finally, results from this meta-analysis may have implications for clinical intervention. The development of child psychopathology may be associated with the mental health functioning of fathers. Thus, family or parent-based interventions for child problems may benefit from including fathers in therapy and by addressing their mental health needs in addition to those of mothers. Empirical evidence suggests that parent–child interactions may differ as a function of the psychiatric functioning and/or presence of each parent. Dumas and Gibson (1990) found that children of depressed mothers were less compliant and more conflictual with their fathers than with mothers, whereas children of nondepressed mothers were more compliant with their mothers than fathers. Additionally, Hops and Seeley (1991) reported that parent–child interactions varied as a function of the presence of the other parent. Mothers and fathers were more aggressive and less distressed when interacting with children in the presence of the other parent. Thus, involving both parents in therapy may impact multiple family subsystems and lead to better outcomes.

The almost exclusive focus on mothers in empirical and review papers on parental depression and child psychopathology has neglected the potentially important contribution of depression in fathers to child development (Forehand et al., 1986; Phares, 1996, 1997). This question is particularly intriguing given evidence indicating that children of two parents with psychiatric diagnoses are at greater risk for psychopathology than children with a single disordered parent or nondisordered parents (Dierker et al., 1999; Merikangas et al., 1988). Although sex differences in adult depression have long been documented (Boyd & Weissman, 1981; Nolen-Hoeksema, 1990; Weissman & Klerman, 1977; Weissman, Leaf, Holzer, Myers, & Tischler, 1984), a significant number of men experience the disorder and greater numbers experience subclinical elevations in symptoms.

Overall, depression in fathers is a significant positive correlate of child psychopathology and conflict. Future research attention to the role of mediating mechanisms and characteristics of the family environment, both cross-sectionally and longitudinally, may be helpful in documenting correlates and processes linking paternal mood disorders to child functioning. Efforts to identify these mechanisms while controlling for the more established risk factor of maternal depression would be helpful steps toward demonstrating whether paternal depression is a unique risk factor for child psychopathology.

Acknowledgements

This work was supported in part by grants from the National Institute of Mental Health. (R01-MH57822), NICHD (P30HD15052), and the William T. Grant Foundation (96173096). Peter Kane was supported in part from an NIMH training grant (T32-MH18921). We extend our appreciation to Bahr Weiss and Mark Lipsey for their consultation regarding the analyses. We also thank Kathleen M. Kaminski for her assistance with coding.
References


Kane, P., Garber, J., & Kaminski, K. M. (2002). Relationship quality as mediators and moderators of the association between fathers’ depressive symptoms and adolescents’ externalizing and internalizing symptoms (Unpublished manuscript).


