Development and Psychopathology, **15** (2003), 403–430
Copyright © 2003 Cambridge University Press
Printed in the United States of America
DOI: 10.1017.S0954579403000221

Developmental differences in the phenomenology of depression

BAHR WEISS AND JUDY GARBER
Vanderbilt University

Abstract

Most researchers and clinicians now agree that children and adolescents are able to develop depressive disorders, and there also appears to be consensus that developmental level has relatively little influence on the phenomenology of the depression. The present paper examines the validity of this latter assumption from methodological, theoretical, and empirical perspectives. We first review reasons why there might be developmental differences in the symptoms that define depression, and then discuss the implications and significance if such differences do or do not exist. Next, we highlight methodological and design issues relevant to the appropriate evaluation of this question. Then, we propose that this broad developmental question actually is comprised of two subquestions—one focusing on symptoms and the other focused at the syndrome level—that have not yet been clearly differentiated in the field. Finally, after conducting a meta-analysis of the current empirical literature and reviewing its limitations, recommendations are made regarding future research in this area.

It is now generally accepted by both clinicians and researchers that children and adolescents can become depressed (Birmaher et el., 1996; Kovacs, 1989); there is, however, less consensus regarding the structure and the nature of the symptoms that comprise the depressive syndrome prior to adulthood. On the one hand, developmentalists (e.g., Carlson & Garber, 1986; Cicchetti & Toth, 1998) have suggested that the depressive syndrome may vary as a function of the cognitive, social, and biological level of the affected individual. On the other hand, current formal criteria for defining depressive disorders in children are based in large part on the adult criteria outlined in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV; American Psychiatric Association, 1994).

Specifically, major depressive disorder in adults, as defined by DSM-IV, is indicated by the presence of either depressed mood or anhedonia in conjunction with at least four other symptoms involving the vegetative, psychomotor, and cognitive domains (see Table 1). Dysthymic disorder represents a less intense, but more longstanding, variation of major depressive disorder. Only minor variations in the DSM-IV criteria are made for children; for example, for children and adolescents, irritability is considered a manifestation of dysphoric mood and the time requirement for a diagnosis of dysthymic disorder is 1, rather than 2, years.

Developmentalists have questioned the application of adult criteria to children and adolescents and have suggested that a developmental perspective may be more appropriate.

The authors would like to thank Thomas M. Achenbach, Carrie M. Borchardt, Rose Calderon, Stephen W. Hurt, Javad H. Kashani, Elizabeth McCauley, John C. Reid, Neal D. Ryan, and Frances Worchel for their willingness to provide additional information regarding their published data and John Weisz and Steve Hollon for their constructive feedback on previous versions of this paper. During completion of this work, Bahr Weiss was supported in part by National Institute of Mental Health grants (1-R01-MH54237 and 1-R01-MH58275), and Judy Garber was supported in part by grants from the William T. Grant Foundation (96173096) and the National Institute of Mental Health (1-R01-MH57822-01A1).

Address correspondence and reprint requests to: Bahr Weiss, Peabody MSC #512, Vanderbilt University, Nashville, TN 37203; E-mail: bahr.weiss@vanderbilt.edu.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

Table 1. *Combined DSM-IV symptom criteria for major depressive and dysthymic disorders*

1. Depressed or (for children or adolescents) irritable mood
2. Anhedonia
3. Appetite disturbance (decreased or increased appetite) or significant change in body weight (loss or gain)
4. Sleep disturbance (insomnia or hypersomnia)
5. Psychomotor agitation or retardation
6. Fatigue or low energy
7. Excessive guilt and/or feelings of worthlessness, low self-esteem
8. Impaired concentration or ability to make decisions
9. Recurrent thoughts of death, suicidal ideation, or suicide attempts
10. Hopelessness

According to the developmental viewpoint, the manner in which depression is experienced, as well as expressed, will depend in part on the individual's level of physiological, social, and cognitive development. Cicchetti and Schneider–Rosen (1984), for instance, asserted that the broad criteria used to define depression in adults may "need to be translated into age-appropriate guidelines for children, sensitive to developmental changes in the children's experience and expression of depression" (p. 7). That is, although there may be certain "core symptoms" that are invariantly a part of depression across development (Carlson & Kashani, 1988), there also may be other depressive symptoms that vary with developmental level.

To date, there have been a few more than a dozen empirical investigations relevant to this developmental question. Although there has not been a systematic review of this literature, the consensus of the field appears to be that the essential symptoms of depression are isomorphic across development (see e.g., Carlson & Kashani, 1988; Ryan et al., 1987). As Cicchetti and Toth (1998) noted, "most often the criteria associated with adult depression have been applied to children, and developmental considerations that may affect the etiology, course, and outcome of depression in children and adolescents have been minimized or disregarded entirely" (p. 222). In fact, the belief that symptoms of depression are essentially the same across development appears to have served, at least in part, as the rationale for such research endeavors as the modification and evaluation of adult psychotherapies for use with adolescents (e.g., Moreau, Mufson, Weissman, & Klerman, 1991) and longitudinal studies of childhood depression and suicide (Rao, Weissman, Martin, & Hammond, 1993). However, we agree with the position of Goodyer (1996), who stated that "The suggestion that the clinical presentation of major depression varies with age is far from resolved and more developmentally sensitive studies are required" (p. 407).

In the present review we examine this question systematically and suggest that the field may have reached a premature conclusion with regard to the developmental isomorphism of depressive symptoms. The present paper (a) outlines reasons why there might be developmental differences in the symptoms of depression; (b) discusses the implications and significance of such differences, if they do exist; (c) highlights methodological and data analytic matters essential for appropriate evaluation of this question; (d) proposes that this developmental question actually is comprised of two subquestions that have not been clearly differentiated in the field; (e) conducts a meta-analysis of current empirical studies and reviews this literature's limitations; and (f) recommends directions for future research.

Why There Might Be Developmental Differences in the Phenomenology of Depression

In considering why there might be developmental differences in depression, it is important first to discuss what we mean by "developmental continuity,"and, conversely, "developmental differences." There are at least two ways in which developmental continuity could be construed. The first involves continuity within the individual, that is, continuity in regard to whether an individual who is depressed at one developmental stage is depressed at a later developmental stage. This assumes that "depression" occurs at more than one developmental level. In fact, individuals with mood

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

disorders in childhood have been found to have recurrent depressive episodes during adolescence (e.g., Emslie et al., 1997; Kovacs et al, 1984) and adulthood (Harrington, Fudge, Rutter, Pickles, & Hill, 1990; Weissman et al., 1999). The second type of continuity involves continuity of the form or essential characteristics of the disorder across developmental level. This focuses on whether depression at different developmental stages is similar in terms of symptom pattern, correlates, and so on, irrespective of whether the same individuals are depressed at different points in time.

Because within-individual continuity depends on the stability of causal factors *for a particular individual* whereas continuity of form depends on stability of *the effects* of the causal factors, it is possible to have continuity of form without having within-individual continuity, and vice versa. Thus, although the same causal factors may be responsible for within-individual continuity and continuity of form, these two forms of continuity have some degree of independence, depending on (a) whether the causal factors show within-individual stability, and (b) whether the effects of the causal factors vary with developmental level.

Probably because research focusing on within-individual continuity requires longitudinal designs, most research on developmental differences in depression has been cross-sectional and focused on the continuity of form question, which is the focus of this review. It is important to emphasize, though, that conclusions regarding developmental continuity of the form of disorder are not necessarily applicable to developmental continuity within the individual. Research investigating continuity of form (which generally has been assessed using cross-sectional designs but also can be assessed using longitudinal designs) is not a substitute for within-individual continuity (which require longitudinal designs), but rather is an important and interesting question in its own right.[1]

1. It should be noted that if there are differences between depression in adults and nonadults, it is likely that these will be a matter of degree rather than absolute differences.

Differences in Expression

There are at least two ways in which depression may differ across the developmental level.[2] First, there might be "heterotypic continuity" (Kagan & Moss, 1962). That is, there may be developmental differences in how symptoms are expressed but the symptoms do not differ when considered as higher level constructs; this form of continuity could occur when investigating either within-individual continuity or continuity of form. An example of heterotypic continuity would be if depressed individuals of all developmental levels experienced anhedonia, but how the anhedonia was experienced and expressed varied with developmental level. Anhedonia in a young child might be characterized by lack of interest in playing with toys; in adolescents it might be manifested as a pervasive sense of boredom; in adults, it might be reflected in a lack of interest in sex. Similarly, although dysphoric mood may be a core symptom of depression for all ages, the manner in which it is expressed may vary. Whereas young children might be more likely to cry and look sad than to talk about their sadness (Poznanski, Cook, & Carroll, 1979), adolescents might be more likely to express their dysphoria as irritability and adults may verbalize their sadness directly. Although on the surface it might appear that these reflect variations in depressive symptoms, the core underlying constructs would be the same (e.g., with anhedonia, a loss of interest in age-appropriate activities that are normally reinforcing).[3]

Thus, a developmental perspective predicts that there probably would not be "homotypic" continuity (i.e., phenotypic or symptomatic con-

2. There also is the case that the wording of assessment instruments will vary developmentally because subjects' abilities to understand assessment questions will vary with age. Because this variation does not represent a true difference in phenomenology, it is not discussed further.

3. This issue is not unique to depression and, rather, can occur with regard to other forms of psychopathology (Loeber & Hay, 1997). For instance, when defining aggression developmentally, biting and hair pulling are considered expressions of aggression in toddlers, whereas verbal sarcasm or the use of weapons are more common forms of aggression in teens.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

sistency; e.g., Kagan & Moss, 1962) across development. Rather, it proposes that "continuity is manifest in lawful relations, and not in isomorphic behavior" (Cicchetti & Toth, 1991, p. 4), with the specific manifestations of some depressive symptoms expected to vary with development.

Differences in the Syndrome

The second and more theoretically interesting type of developmental difference is that the specific symptoms that comprise the syndrome of depression may differ developmentally; that is, a different combination of symptoms may characterize depression in different developmental groups. There are at least two ways in which this could happen: there might be developmental limitations in younger developmental groups' ability to experience or develop certain symptoms of depression, and the causes and/or consequences of depression might vary to some degree across development.

Developmental limitations. In regard to the first possibility, some theorists have suggested that young children are not yet developmentally capable of experiencing specific symptoms of depression. For example, early psychoanalytic writers (e.g., Gaylin, 1968; Rie, 1966) theorized that children could not experience depression because they lacked many of the necessary intrapsychic structures. Although this particular view is no longer held, the notion that developmental limitations might underlie age differences in the occurrence of some symptoms of depression has been suggested by other theorists. For instance, Cicchetti and Schneider–Rosen (1984) have questioned whether young children have the cognitive capacity and linguistic ability to develop, experience, and/or express certain depressive symptoms.

Supporting this perspective is evidence indicating that children undergo important cognitive developmental changes (Flavell, 1985) that could affect their experience of certain depressive symptoms. Such development influences their understanding of emotions (Harris, 1989; Schwartz & Trabasso, 1984), concept of the self (Hart & Damon, 1986; Harter, 1999), social cognitions (Fiske & Taylor, 1991), social perspective taking (Selman, 1980), and causal attributions (Friedlander, 1988; Ruble & Rholes, 1981). For example, young children's negative affect tends to result from a reaction to the immediate environment (Weiner & Graham, 1985). In contrast, preadolescents have the capacity to think abstractly about the self and future (Harter, 1983) and hence are more able to develop stable negative self-views and negative expectations about the future, increasing their ability to sustain negative emotions beyond the immediate situation (Harris, 1989; Piaget, 1971).

Thus, certain depressive symptoms may be affected by development, in particular cognitions such as guilt, low self-esteem, and hopelessness (Cicchetti & Toth, 1998). For instance, it is during early childhood that individuals develop certain prerequisites for the emergence of guilt, such as an awareness of standards regarding prohibited behavior and simple forms of perspective taking (Kagan, 1982; Zahn–Waxler, Chapman, & Cummings, 1984). Although young children might experience some rudimentary forms of guilt, particularly after a real-life transgression that is punished, they may not necessarily be capable of experiencing the kind of guilt associated with depression, which may involve other factors such as beliefs about control and personal responsibility (Weiner & Graham, 1985). It generally is during adolescence, with the advent of abstract thinking, that individuals develop increasing concern for social issues and moral principles (Piaget, 1971), which serves as ripe ground for the development of depressive cognitions about moral inadequacy and guilt. Thus, the experience of depressive guilt may be associated with changes in cognitive development, and therefore not apparent in young children.

Similarly, a number of cognitive changes do not occur until middle to later childhood with regard to the ability to describe the self in psychological rather than concrete terms (Damon & Hart, 1982; Selman, 1980), to conceive the self in terms of stable social characteristics and personality traits (Stipek & MacIver, 1989), and to use social comparisons for

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

the purposes of self-evaluation (Ruble & Frey, 1987). All of these abilities, which are not well developed during early childhood, are possible prerequisites for the kind of low self-esteem that characterizes depression (Harter, 1993). For example, the ability to make judgments about themselves in comparison with their peers and their ego ideal may increase children's likelihood of developing the negative, self-derogatory conclusions that characterize low self-esteem. Moreover, with the onset of formal operational thinking during adolescence, children are also able to construct a broader concept of their personality, which makes it easier for them to detect inconsistencies in the self that can lead to the feelings of dysphoria and low self-worth that characterize depressive disorders (Harter, 1993, 1999).

Finally, young children might not be developmentally capable of experiencing sustained hopelessness, given their level of cognitive development and their tendency to be present oriented. Rutter (1986) has suggested that younger children may be less prone to experience feelings of helplessness and hopelessness "because they do not view failure as implying a stable and lasting limitation on their performance" (p. 23). According to Piaget (1971; Singer & Revenson, 1996), the "future" is the last time concept children learn. Moreover, because young children tend to be "concrete" and less capable of abstractions that involve the projection of the self into the future, they might not be able to experience the sort of hopelessness that characterizes depression in adults. If true, cognitions of hopelessness could not contribute to the development or maintenance of depression (Abramson, Metalsky, & Alloy, 1989).

Thus, according to developmental theory and evidence, certain symptoms of depression might not occur in children because they are not yet capable of experiencing them. This, in turn, could result in developmental differences in the symptoms that characterize depression at different ages.

Different causes. A second reason why developmental differences in the syndrome of depression might occur is because there might be developmental variations in the causal processes underlying depression (Wickramaratne & Weissman, 1998). For example, the work of some developmental theorists (e.g., Cole & Turner, 1993; Nolen–Hoeksema, Girgus, & Seligman, 1992; Weisz, Southam–Gerow, & McCarty, 2001) suggests that the relations among cognitions, stress, and depression may change from middle childhood through early adolescence. For example, Nolen–Hoeksema et al. (1992) found that negative life events, but not attributional style, significantly predicted depressive symptoms in 8-year-old children, whereas in 11- to 12-year-old children a pessimistic explanatory style, both alone and in conjunction with negative events, significantly predicted depressive symptoms. This developmental difference may reflect developmental changes in the causes of depression.

Another possible reason for developmental differences in the causes of depression has been suggested by Post (1992), who proposed that repeated exposure to the same (or a similar) psychosocial stressor leads to increased biological sensitivity and pathological responsiveness to the stressor, as a result of "experience-dependent modifications of the genome at the level of transcriptional regulation" (Post & Weiss, 1998). According to this perspective, the relative importance of environmental versus biological etiologies for depression will vary as a function of the amount of experience an individual has had with depressogenic events. Further, because children (and, to a lesser extent, adolescents) will have had relatively few of these experiences, it will take fairly substantial environmental events to evoke depression, because the biological mechanism capable of evoking depression has not been sensitized. Hence, at younger developmental levels the etiology may be more environmental than biological. On the other hand, for adults who have experienced many depressogenic events and thus have developed biological sensitivity, it will take a relatively minor environmental stimulus to evoke depression.

If true, this would suggest that certain biological features of depression might be more likely to be present among individuals of higher developmental levels, due to an in-

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

creased importance of a direct biological etiology. Thus, there might be certain depressive symptoms that children are developmentally capable of experiencing (perhaps, for instance, appetite and sleep disturbances) but that are more a part of depression in older developmental groups because the symptoms are tied more strongly to a causal factor in these older groups.

Variations in causal relations

To understand how such differences in causal processes might arise, we next briefly review (a) why a specific (i.e., individual) symptom might be related to the broader collection of symptoms (i.e., the syndrome of depression); and (b) how developmental differences in the relation between a specific symptom and the syndrome of depression might arise.[4] To begin, a specific symptom such as anhedonia and the full syndrome of depression both may be caused by a common third variable; that is, they may share a mutual causal agent (see Figure 1a); if the relation between the symptom and the rest of the syndrome is sufficiently large, then the specific symptom will be considered part of the syndrome. For instance, the same biochemical imbalance might cause anhedonia as well as other symptoms of the syndrome (e.g., sleep problems; dysphoric mood; appetite change).

Second, an individual symptom might be related to the broader syndrome of depression because the symptom itself may be part of the causal process producing the depression (see Figure 1b). For example, the reason hopelessness is strongly correlated with depression in both adults (e.g., Beck, Steer, Beck, & Newman, 1993) and children (e.g., Asarnow & Guthrie, 1989) may be because hopelessness is part of the causal process for some forms of depression (Abramson et al., 1989). Thus, in this instance, a symptom of depression (e.g., hopelessness) would be related to the syndrome because it is part of the causal chain underlying depression.

A third reason an individual symptom might be associated with the depressive syndrome is because the symptom may be a consequence of depression (see Figure 1c). For instance, concentration problems, which are presently considered a symptom of depression (American Psychiatric Association, 1994), might be a result of other symptoms of depression (e.g., distracting cognitions; fatigue). In such a case, then, a specific symptom would be related to depression because it was a consequence of other symptoms of the depressive syndrome.

If any of the causal processes (represented by the arrows in Figures 1a–c) change or vary across development, then there likely would be developmental differences in the relations among depressive symptoms. For instance, if the etiology of depression were less biological in children than in adolescents or adults (Post, 1992; Post & Weiss, 1998), then the relation between symptoms directly resulting from these biological processes (e.g., sleep disturbance) and the rest of the syndrome of depression might be greater for older age groups. This would occur because with increasing development, the causal processes underlying depression would be more responsible (both in absolute terms and relative to the nondepressive causes of sleep disturbances) for intersubject variability in sleep disturbances. Compared to younger groups, in older groups a greater proportion of the variance in sleep disturbance would be accounted for by the causal processes underlying depression, and hence sleep disturbance would be more closely linked to depression among the older group.

Implications of Developmental Differences Versus Developmental Isomorphism

There are several reasons why it is important to determine the extent to which there is developmental continuity in depression. First, assessing developmental continuity represents a step in the validation process for the construct of depression in nonadults. Validating a construct such as depression in nonadults involves many pieces, such as investigating correlates, the course of the syndrome, and so

4. For simplicity, we discuss the relation between a single symptom and the syndrome. However, the same logic applies to multiple symptoms simultaneously correlated.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

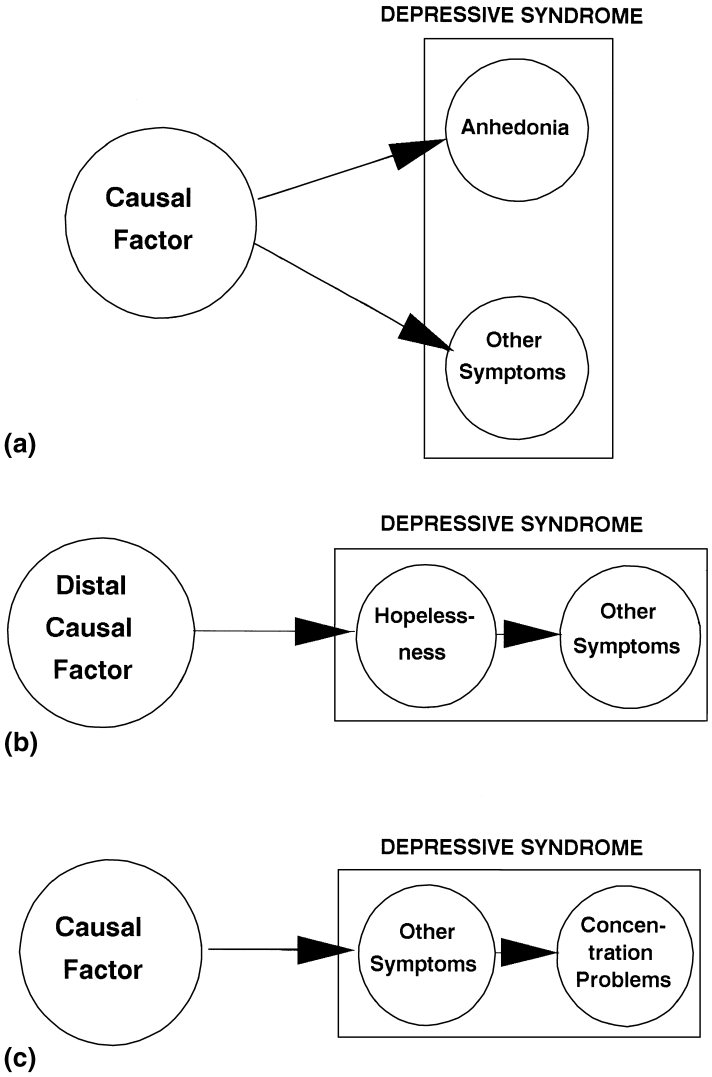


Figure 1. Different pathways for the relations between specific symptoms and the depressive syndrome.

forth, in relation to theory-based predictions. In the case of childhood depression, the theoretical predictions and criteria come at least initially from the adult depression literature. If we are to call it depression, then the degree of similarity of its structure to that of adult depression is one step in the validation process.

Second, the issue of developmental continuity is important in its own right (Cicchetti & Toth, 1998). Determining the extent to which and how psychopathology changes with development addresses a fundamental question regarding the structure and nature of psychopathology (Compas, Ey, & Grant, 1993). And finally, determining whether there are developmental differences in the symptoms of depression will be an important preliminary step in determining the extent to which the wealth of information accumulated about the etiology, assessment, and treatment of depression in adults might be applied to children. For instance, if it were found that certain symptoms (e.g., hopelessness) were less a part of depression in children than in adults, then this would reduce the likelihood of their

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

causal status for depression in children, suggesting that they should not be included as part of diagnostic criteria and that interventions focused on these symptoms might not be effective (Abramson et al., 1989).

Design Issues: How to Test the Developmental Difference Question

Although testing the developmental difference question at first might seem relatively straightforward, there are a number of design and analytic issues that should be considered if the hypothesis is to be appropriately evaluated. These factors include how a sample is defined and selected, how development is defined, the methodology used to assess symptoms, and issues related to statistical power and acceptance of the null hypothesis.

Population selection and sampling procedures

In the empirical literature, three types of samples have been used in investigations of developmental differences in depression: (a) diagnostically homogeneous samples (e.g., Ryan et al., 1987) restricted to subjects who have received an affective diagnosis (hereafter referred to as diagnosed samples); (b) diagnostically heterogeneous, clinic-referred samples (hereafter referred to as clinic samples; e.g., Achenbach & Edelbrock, 1983); and (c) normal or epidemiological samples (hereafter referred to as normal samples; e.g., Kashani, Rosenberg, & Reid, 1989). These populations vary along at least two dimensions that shape the conclusions that can be drawn from the data: (a) the number and type of inclusion and exclusion criteria; and (b) the severity of depressive symptoms in the sample.

Each of these three populations has its own general form of inclusion and exclusion criteria (or lack thereof). To obtain a diagnosed sample it is, of course, necessary to specify in advance the criteria that define "depressed" children and adolescents. This selection process makes developmental comparisons problematic. When subjects are selected using certain criteria (e.g., DSM-IV depression criteria) that then serve as the dependent variables in a group comparison, the likelihood of finding group differences in these criteria invariably will be decreased. This is because one is, to some degree, equating the groups when subjects are selected using the same criteria. The extent to which the selection procedure equates the groups will depend on base rates, the form of the criteria, and whether the comparisons are based on categorical or continuous variables. This will be true even when using a polythetic approach to classification criteria in which multiple combinations of symptoms are possible.

For instance, suppose one inclusion criteria was the presence of sadness. This would result in the two developmental groups having the same percentage of subjects with the symptom (i.e., 100%), and to compare the two groups on the frequency of sadness would be meaningless. Even comparing the groups on the severity of sadness would be problematic (although less so than a frequency comparison), because subjects with no or low levels of sadness would have been consistently eliminated from both groups, thus tending to equate the group means.

A similar process will occur when polythetic criteria are used. Suppose, in this instance, the inclusion criteria are the presence of sadness or anhedonia; that is, a subject must be either sad or anhedonic. The application of these criteria means that the more sad or anhedonic a subject is, the more likely he or she will be included in the sample, regardless of whether he or she is a child or an adolescent and regardless of the frequency of these symptoms in the population (assuming that the frequency is not zero). Consequently, this inclusion criterion will increase the frequency of sadness and anhedonia in both developmental groups, thus tending to equate them for these variables. Of course, this matching will not ensue to the same extent as with the single inclusion criterion, but it will nonetheless occur.

To a lesser degree, this same problem will occur with heterogeneous clinic samples. The selection criteria for a clinic sample will be the presence of behaviors that concern parents or teachers sufficiently to cause them to bring the child to a clinic. These criteria will be less

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

explicit and less consistent across subjects than those used to create a diagnosed sample, but they still will tend to equate groups being compared, consequently decreasing the likelihood of finding developmental differences.

Another complication with clinic samples is that the problem behaviors that concern the person responsible for referral (e.g., in children, the parents; in adults, the individual him- or herself) may vary as a function of age. Suppose, for example, that although abnormal levels of fatigue occurred with equal frequency in children and adolescents, parents were more concerned by fatigue when it occurred in young children. If this were true, then it is likely that children, in comparison to adolescents, would be referred more often for severe fatigue; thus, in a clinic sample, children would show higher rates of fatigue. Rather than representing developmental differences in this depressive symptom, however, these different rates of fatigue across ages would actually reflect varying patterns of parental concern about the symptom.

This might seem to suggest that attempts to answer the question of developmental differences in the symptoms of depression should be restricted to normal, unselected community samples. However, a major shortcoming of community samples is that the frequency with which children experience the various symptoms of depression will be considerably lower than found among clinic or diagnosed samples. If too few children in the total sample exhibit depression, the likelihood of identifying developmental effects will be decreased, because the distribution will be skewed and the range of the variables restricted, thereby decreasing power.

Thus, all three of these populations have both advantages and disadvantages. What perhaps may be the best approach for sample selection is first to obtain subjects from some population (e.g., children of depressed parents; a mental health clinic) where there is likely an increased risk for depression, whatever "depression" may be for a particular developmental group. Then minimal screening criteria rather than strict diagnostic criteria can be applied to eliminate from the sample those subjects who clearly are irrelevant to the purposes of the study (e.g., "normals" from the former sample; children strictly with conduct problems from the latter). Once a set of symptoms has been identified using such high risk samples, then the generalizability of the results could be tested further by using clinic and community samples.

Defining developmental groups

There are at least three different methods for dividing a sample developmentally: pubertal status, cognitive developmental level, and chronological age or grade in school. Ideally, which method investigators choose should depend on their theory regarding the cause of the hypothesized developmental differences. For example, one possible explanation for developmental differences in the phenomenology of depression might be that they are the result of biological/hormonal changes that occur at puberty (Angold & Worthman, 1993). Researchers who hold such a view should divide their sample according to pubertal status, because (insofar as the pubertal theory is valid) this division would give the most homogeneous and meaningful groups. On the other hand, investigators who hypothesize that developmental differences result from differences in cognitive level should divide their sample according to cognitive indices of development (e.g., Piagetian stage, mental age; see Kovacs & Paulauskas, 1984).

Unlike the first two, the third basis for dividing the sample—age or grade—is not related to an a priori conceptualization about the processes underlying developmental differences. Chronological age or grade in school cannot in and of themselves be direct causes of developmental differences but rather must serve as markers of underlying, correlated processes (e.g., social experience, cognitive and pubertal development; Rutter, 1986). Thus, one drawback of using age or grade is that they are indirect indices of development. Chronological age and grade, however, have the advantage of being easy to measure, thereby making large samples feasible. In addition, because using age or grade as the basis for dividing the sample does not involve an a priori judgment as to the reason for differences, it may

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

be most appropriate early in the process of exploring developmental differences, when the investigator does not yet have a clear theory regarding the mechanisms underlying such differences. In fact, under certain circumstances, categorization by age may produce a more powerful test for dividing the sample than other methods. Suppose, for example, that one investigator divided the sample by pubertal status and another divided it by age. If it turned out that developmental differences were actually related to cognitive level, then the investigator dividing by age would have a more powerful design, if cognitive level were more highly correlated with age than with pubertal status.

What then might be the best approach for defining developmental groups? Perhaps it is dividing a sample using multiple methods, with the investigator conducting separate developmental comparisons based on each method. Kovacs and Paulauskas (1984), for instance, divided their sample by pubertal status, as well as by cognitive stage, and then conducted two sets of analyses using the two different operationalizations of developmental level. Although it would require a greater expenditure of research energy, such an approach allows for a comparison of the hypotheses upon which the developmental groupings are based; that is, the theory underlying the method that produces the largest number of developmental differences would receive more empirical support.

Assessment methodology

Depressive phenomenology can be assessed by a number of different methods, including self-report questionnaires, clinical interview, and parent or teacher report. Behavioral observation of depressive symptoms also is possible, although this method is less well developed and the many internal, subjective symptoms of depression may be relatively difficult to assess via short-term behavioral observation (Garber & Kaminski, 2000). The strengths and weaknesses of these various assessment methodologies have been well reviewed (e.g., Curry & Craighead, 1993; Kazdin, 1987). Therefore, the present discussion focuses on assessment issues specific to developmental comparisons.

Self-report, although providing a direct account of the important subjective symptoms of depression, is problematic in that the age of the informant (i.e., the child) is confounded with the age of the subject (i.e., the child). Hence, if younger children are less able to comprehend the self-report instrument, one may find developmental differences that reflect subjects' differential ability to read and understand the questionnaire rather than differences in the manifestation of the symptoms themselves. Similarly, younger children also may be less able to introspect and conceptualize their symptoms, and hence report them; thus, cognitive ability may create artifactual developmental differences in symptoms.

Other informants (e.g., parents) are not subject to this form of confounding, although they may have their own unique biases. For example, maternal report of children's symptoms may be correlated with the mother's own level of psychological distress (e.g., Fergusson, Lynskey, & Horwood, 1993; Renouf & Kovacs, 1994). If there is bias in maternal report, then this could influence the results of developmental comparisons in at least two ways. First, parent–child concordance could interact with age such that the correlation between mothers' and offspring's reports might be larger for adolescents than for children, because adolescents may verbalize their feelings more than younger children (Edelbrock, Costello, Dulcan, Conover, & Kala, 1986). Thus, mothers of adolescents might report more of some symptoms because they are more aware of them, not because adolescents actually have more of these symptoms than children.

Second, there could be the appearance of age differences in depressive symptoms based on maternal report if the level of maternal depression varies as a function of the age of the child. That is, if depression is more common in mothers of younger children (e.g., Chilcoat & Breslau, 1997), then their report of symptoms in their young children might be more influenced by their own symptoms than would the report of mothers of older children. Consequently, both of these factors could create bias that could influence the results of developmental comparisons.

Optimally, then, multiple methods and informants should be used in studies examining

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

this developmental question (Achenbach, McConaughy, & Howell, 1987). One approach that might be particularly useful would be to contrast the information provided by different sources to determine if developmental differences in depressive symptoms occur as a function of type of informant. Another approach would be to combine the information obtained from multiple sources using various statistical (e.g., confirmatory factor analysis; see Bollen & Long, 1993) or clinical methods (Loeber, Green, Lahey, & Stouthamer–Loeber, 1991; Reich & Earls, 1987).

Comparisons using previously published data

Another design issue important to consider is the use of previously published data as a developmental comparison group. In several studies reviewed below (i.e., Carlson & Kashani, 1988; Carlson & Strober, 1979; Friedman, Hurt, Clarkin, Corn, & Aronoff, 1983; Mitchell, McCauley, Burke, & Moss, 1988), investigators compared their own data from one developmental group (e.g., adolescents) to data published by other authors from a second developmental group (e.g., adults). Thus, the data for the two developmental groups were collected at different sites under different conditions, and sometimes with different assessment instruments. The advantage of this strategy is that it allows investigators to increase the number of developmental comparisons they can make with their data. One drawback of this approach, however, is that it confounds assessment and site differences with the developmental level of the subjects. Thus, this strategy could either produce apparent developmental differences when there actually are none or apparent developmental isomorphism when there are in fact differences.

The report by Friedman et al. (1983) illustrates this problem. These authors compared their adult sample to two adolescent samples, their own adolescent sample collected in the same study and an adolescent sample reported by Strober, Green, and Carlson (1981). The two adult–adolescent comparisons produced quite different results. When Friedman et al. compared their adult sample to their own adolescent sample, only 1 of 30 developmental comparisons was significant. In contrast, when they compared their adult sample to Strober et al.'s (1981) adolescent sample, 9 of 28 developmental comparisons were significant. These contrasting results appear to have been a function of important differences between the adolescent samples of Friedman et al. and Strober et al. Of the 9 depressive symptoms that were significantly different between the adults in Friedman et al. and the adolescents in Strober et al., seven symptoms also were significantly different between the two different adolescent samples (i.e., Friedman et al. and Strober et al.). Given these significant findings, it is quite possible that the significant developmental differences between the Friedman et al. (1983) adult sample and the Strober et al. (1981) adolescent sample were largely due to site or methodological differences between the two studies rather than a true developmental effect.

A similar problem can occur within studies, when developmental groups differ on important clinical or demographic characteristics other than developmental level. For example, in the study by Mitchell et al. (1988), their preadolescent sample contained 47% inpatients whereas their adolescent sample contained no inpatients. Few significant differences were found, which might suggest that younger and older children exhibit the same levels of the symptoms. However, if inpatients have higher levels of symptoms than outpatients, which is a reasonable assumption, then the inpatient versus outpatient differences may have run in opposite directions and canceled each other out. Thus, studies comparing depressive symptoms of individuals at different ages should be sure that their samples do not differ in important ways other than the developmental variable of interest.

Statistical power and acceptance of the null hypothesis

In any investigation, obtaining a minimal level of statistical power is critical. This may be particularly true in regard to the search for developmental differences in depression, because the null hypothesis is of interest in its own right; that is, determining that there are

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

in fact no developmental differences has important theoretical and clinical implications. Such an enterprise, of course, puts researchers in the statistically precarious position of trying to "prove" the null hypothesis. Acceptance of the null hypothesis can never be final, and even tentative acceptance requires ruling out the alternative hypothesis that null findings represent low statistical power or poor research design, rather than a lack of meaningful population differences. Thus, researchers seeking to support a null hypothesis need to make comparisons with a relatively high degree of statistical power and methodological rigor.

Unfortunately, a number of the studies making developmental comparisons of depressive phenomenology have failed to obtain minimally acceptable power. For instance, Stehouwer, Bultsma, and Blackford's (1985) sample consisted of 50 subjects. With a sample of this size, the probability of finding a half standard deviation difference (a moderate-sized difference) between groups is less than 0.50 (Cohen, 1988). A sample size of 125, split approximately evenly between two developmental groups, would produce a minimally acceptable power of 0.80.

Another factor influencing statistical power is the transformation of continuously distributed assessments of symptoms to dichotomous ("presence/absence") data. In general, when a continuously distributed variable is dichotomized, statistical power is reduced (Cohen, 1990) and the probability of making a Type II error is increased. For diagnostic purposes, it may be useful to convert continuous ratings to dichotomies; for purposes of addressing the question of developmental differences, however, such a division probably is not desirable.

Dichotomization of the independent variable (i.e., developmental level) may similarly decrease power. This problem is relatively easy to remedy. Rather than converting age into a categorical variable (i.e., children versus adolescents), it may simply be left continuous. A general linear model could then be used for continuous independent variables (e.g., Neter, Wasserman, & Kutner, 1989) in analyses, instead of the analysis of variance (ANOVA) approach that has been used in most studies examining this issue. Cognitive level, insofar as it is conceptualized as a continuous process, could be treated similarly. Even when it is not, assessments of cognitive level generally involve dichotomization of continuous variables (see Kovacs & Paulauskas, 1984). Which approach produces greater power is then an empirical question that can be answered by analyzing cognitive level as a continuous or categorical variable. Finally, although pubertal status is more clearly categorical, when assessed via the five-stage Tanner (1962) system, it still involves dichotomization and justifiably might be treated as three categories (i.e., prepubertal, transitional, and pubertal) rather than two. Such strategies could serve to increase statistical power.

Two other points related to acceptance of the null hypothesis are important to note. First, some investigators (e.g., Borchardt & Meller, 1996) failed to provide full information in their published reports on their nonsignificant findings, which makes it difficult for readers to determine the extent to which the null (or alternative) hypothesis is supported. Second, investigators (e.g., Kovacs & Paulauskas, 1984) sometimes have performed joint tests, wherein they simultaneously tested a number of conceptually related depressive symptoms for developmental differences. Although this approach has much to recommend it in that it tests for developmental differences in systems of related symptoms, it also would be useful to report simple univariate tests in order to facilitate the integration of findings across investigations.

When is enough enough?

Finally, it is important to consider the question of when there are enough differences in the individual symptoms to warrant concluding that there are developmental differences in the overall manifestation of depression. This question might seem prosaic. However, review of the conclusions drawn by the studies comprising this literature indicates that what constitutes enough differences in the individual symptoms varies widely across research groups and that, far from being prosaic, this question in fact is pivotal. For instance, Worchel, Nolan, and Wilson (1987) found significant developmental differences for 5 of 27

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

(19%) items on the Children's Depression Inventory, and concluded that the results of their study "suggest that there are some developmental differences" (p. 413) in depressive symptomatology. In contrast, Carlson and Kashani (1988) found significant differences for 12 of 17 (71%) symptoms, more than three times the rate reported by Worchel et al. (1987), and concluded that "while there appear to be several specific developmental modifications, at least on the basis of this diverse sampling of age groups from different studies, the basic symptom picture of serious depressive disorder was relatively unchanged regardless of age" (pp. 1225). Similarly, Ryan et al. (1987) concluded that there were "relatively few" differences between children and adolescents, although over a third of the core depressive symptoms taken from DSM-III-R showed significant developmental differences. Thus, at present there does not appear to be a consistent standard for judging whether there are overall developmental effects.

What, then, might provide a consistent standard against which to judge the extent of differences when they are observed? The most obvious one is conventional statistical significance. At a minimum, a simple chi-square test can be used to evaluate whether the number of significant developmental effects differs from chance. A more desirable strategy would be to perform a multivariate test, evaluating all symptoms simultaneously.

Review of Empirical Studies

Before presenting this empirical review, it is important to note that underlying the question of developmental differences in depression are two subquestions that must be delineated clearly in order to appropriately conduct the empirical review. This is true because these two questions, the first of which focuses on individual symptoms of depression and the second on the syndrome of depression, have different methods of analysis and different potential implications.

Developmental differences in depression: Symptom levels versus syndromes

In the depression literature, a distinction has been made between the symptom, syndrome, and disorder levels of depression (Compas et al., 1993). The symptom level focuses on depression as an individual symptom (i.e., sadness); the syndrome level focuses on depression as a cluster of related symptoms; the disorder level focuses on depression as a discrete psychopathological entity (e.g., major depressive disorder).

This distinction has important implications for the study of developmental differences in depression.[5] At the symptom level, researchers have investigated whether the mean level or prevalence of depressive and depression-related symptoms varies developmentally, using ANOVA-type designs. An example of this approach is provided by Carlson and Kashani (1988), who compared the frequency of depressive symptoms across four different developmental groups. In contrast, at the syndrome level, researchers have used factor analytic techniques to compare the relations among depressive symptoms across developmental groups in order to determine whether the syndrome of depression varies developmentally. An example of this approach is provided by Ryan et al. (1987), who compared the factor structures of K-SADS symptom ratings for groups of prepubertal versus pubertal children.

In general, the distinction between the symptom and syndrome levels of inquiry has not been well established in this literature. However, this distinction has important implications for the conclusions that can be drawn from these different kinds of studies. Investigations of developmental differences at the symptom level provide information relevant to assessment and epidemiology. They can tell us about developmental differences in how "depressed" children present and so on. In contrast, studies at the syndrome level can provide information about relations among symptoms, and thus may yield clues about developmental differences in the etiology of depression.

Review of studies at the symptom level. To cumulate the results of studies comparing developmental groups at the symptom level, we

5. To date, there appear to be no studies that have made developmental comparisons at the disorder level; therefore, this discussion focuses on the symptom and syndrome levels.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

Table 2. *Study characteristics*

Study	Sample Size	Age Range	Female (%)	Population Group	Developmental Definition	Assessment Methodology	Informant
Achenbach	1300/1300	4–16	50	Nor/Cli	Age	PC	P
AHQC	1300/1300	4–16	50	Nor/Cli	Age	PC	P
Baker	100	16–76	69	Dia	Age	SI	S, F
Borchardt	56	5–19	36	Dia	Age	CR	CR
Carlson[a]	28			Dia	Age		
Friedman	26/27	15–45		Dia	Age	SI	S
Garber	23/84/30	7–13	100	Cli	Age	SI, PC, TC	S, P, T
Kashani	70/70/70	8–17	50	Nor	Age	SI	S, P
Kovacs	28/52, 64/16[b]	8–13	48	Dia	Pub/Cog	SI	S
Kovacs 92	633/633	7–16	53	Nor	Age	SR	S
Links	1586/1035	4–16	49	Nor	Age	PC	P
Messer	500/500/500	6–15	0	Nor[c]	Age[d]	SR	S
Mitchell[e]	45/50	7–17	48	Dia	Age	SI	S, P
Ryan	95/92	6–18	46	Dia	Pub	SI	S, P
Smucker[f]	615/369	8–16	52	Nor	Age	SR	S
Stehouwer	25/25	13–54	100	Dia	Age	SR	S
Strober	40	12–17	70	Dia	Age	SI	S
Weiss	515/515	8–16	40	Cli	Age	SR	S
Woodruff	54	21–65	61	Dia	Age	SI	S
Worchel	304	6–17	49	Nor	Age	SR	S

Note: Sample sizes are for separate developmental groups within a particular study. Population: Nor, normal; Cli, Clinic; Dia, Diagnosed. Defining developmental groups: Age, age; Pub, pubertal status; Cog, cognitive developmental level. Assessment methodology: CR, hospital chart review; PC, parent checklist; SI, structured interview; SR, self-report; TC, teacher checklist. Informant: F, family; P, parent; S, subject; T, teacher. Achenbach: Achenbach & Edelbrock (1983); AHCQ, Achenbach et al. (1991); Baker, Baker et al. (1971); Borchardt, Borchardt & Meller (1996); Carlson, Carlson & Strober (1979); Friedman, Friedman et al. (1983); Garber, Garber (1984); Kashani, Kashani et al. (1989); Kovacs, Kovacs & Paulauskas (1984); Kovacs 92, Kovacs (1992); Links, Links et al. (1989); Messer, Messer et al. (1995); Mitchell, Mitchell et al. (1988); Ryan, Ryan et al. (1987); Smucker, Smucker et al. (1986); Stehouwer, Stehouwer et al. (1985); Strober, Strober et al. (1981); Weiss, Weisz et al. (1989); Woodruff, Woodruff et al. (1967); Worchel, Worchel et al. (1987).

[a]The Carlson and Kashani (1988) comparison, which is reviewed here but not reported in Table 2, contrasted the Baker et al. (1971) adult sample and the Ryan et al. (1987) adolescent sample.

[b]The first sample size is prepubertal versus pubertal, and the second sample size is "not abstract" versus "primarily abstract" cognitive developmental level.

[c]Messer et al.'s (1995) sample was a mixed sample with 50% selected as being at high risk for antisocial behavior and the other 50% randomly selected from public schools.

[d]The developmental level was based on grade.

[e]The figures do not include the Baker et al. (1971) sample, to which the children and adolescents were compared.

[f]Gettysburg sample.

used meta-analytic techniques (e.g., Cooper & Hedges, 1994). Specifically, for each study we computed an effect size estimate, which represented the magnitude of the effect of developmental level on the various depressive symptoms assessed in the study (e.g., anhedonia; sleep disturbance). There were three study inclusion criteria for the meta-analysis. The first was that a study had to compare at least two developmental groups on depressive or depressive-related symptom(s). The second criterion was that the analyses had to provide information regarding individual relations, rather than joint (i.e., combined) relations among symptoms (Kovacs & Paulauskas, 1984); this criterion was included because different authors have used different symptom clusters and thus it would not be possible to cumulate across studies. That is, studies had to univariately compare developmental groups on individual symptoms (e.g., guilt) rather than multivariately compare developmental groups on "depressive cognitions" (e.g., hopelessness, guilt, and low self-esteem). This meant that many instrument validation reports (e.g., Knight, Hensley, & Waters, 1988) were not

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

included, because they did not report developmental item-level analyses. Finally, comparisons needed to involve developmental groups containing at least 15 subjects, so that within-study parameter estimates would be minimally stable (Cooper & Hedges, 1994). There were no age-range restrictions. To identify relevant studies published in the literature from 1970 to 2000, we (a) conducted a computer literature search, (b) manually reviewed 12 journals that publish psychopathology research, and (c) reviewed bibliographies of studies identified in the first two approaches.

Sixteen studies passed these criteria, representing 20 comparisons; Table 2 describes the general characteristics of the samples and methods.[6] For a particular symptom (e.g., anhedonia) to be included in this review, it had to have been assessed by at least 3 of these studies; 29 symptoms passed this criterion (see Table 3).

For our effect size estimate, we used Cohen's (1988) *d*, which in this instance equals the mean symptom level for the more developmentally advanced group, minus the mean for the less advanced group, divided by the pooled standard deviation; thus, a positive effect size indicated that the more advanced group exhibited higher levels of the symptom. Because in small samples *d* is a biased estimate of the population parameter, we used Hedges and Olkin's (1985) adjustment for small sample bias. When means and standard deviations were not provided in the published report, we used the formulas provided by Smith, Glass, and Miller (1980) for computing effect sizes from *t* statistics, and so on. When full information was not available or could not be derived from the published report, data were requested from the author(s); all the authors we contacted provided the requested information. When data that were originally continuous were dichotomized, we requested the original means and standard deviations from the authors. When they were unavailable, we used probit analysis (Finney, 1971) to calculate effect sizes.

Variability across studies is inherent in any type of literature review, and we therefore used several standard techniques to address different aspects of this variability. First, in our analyses we weighted by the inverse of the variance of the parameter estimate for the effect size (Hedges & Olkin, 1985), which controls for differences in sample sizes and differences in reliability of the effect size estimates. Then, before computing the mean effect size for each symptom and testing whether it differed significantly from zero (i.e., whether there was a significant effect for developmental level on the symptom), we first determined separately for each of the 29 symptoms whether the various effect sizes across the different studies in our meta-analysis could be represented by a common effect size, their mean. Toward this end, we conducted Hedges and Olkin's (1985) Q statistic for each of the 29 symptoms. The Q statistic, which represents the residual sums of squares when a standardized variable (e.g., effect sizes) is used as a dependent variable, determines whether there is significant heterogeneity across the effect sizes by assessing whether the observed variability of effect sizes (across studies) is greater than expected by chance, given a single population effect size underlying the sample effect sizes.

Of the 20 symptoms included in DSM-IV (American Psychiatric Association, 1994) depressive disorders diagnostic criteria, 17 showed significant heterogeneity, as did 6 of the 9 "associated" symptoms (e.g., anxiety).[7] This significant variability implies that there were developmental effects for these 23 symptoms but that the developmental effect for each of them was not consistent across the various

6. One study that otherwise fit these criteria was not included in this review. Inamdar, Siomopoulos, Osborn, and Bianchi (1979) compared their data for a sample of depressed adolescents to data for adults published by Woodruff, Murphy, and Herjanic (1967). However, Inamdar et al. (1979) systematically excluded reporting comparisons of symptoms that occurred in less than 50% of any of the samples, thereby potentially biasing their results.

7. When determining whether data fulfill assumptions (e.g., homogeneity of variance), one typically sets a relatively high alpha (Kirk, 1982), because one is attempting to establish a null hypothesis of no effect. We acted similarly in this instance, using $\alpha = .10$ for the homogeneity tests.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

studies in this sample. This follows because if the true population effect size for a particular symptom were zero (i.e., if there were no developmental effects for that particular symptom), then there would not have been significant variability across the various studies in the sample. Stated differently, if there were no developmental effects for a particular symptom, then all of the studies would have effect sizes close to zero, with little variability around a mean of zero.

Next, for the 23 symptoms that showed significant variability, we recomputed the *Q* statistic, this time controlling for three potential study-level moderators of the effect of developmental level on symptom level: gender composition (percentage of females in the study sample), type of sample (normal, clinic, diagnosed), and informant (three categories had sufficient sample sizes to be tested: parent, self-report, interview). Thus, for each of the 23 symptoms, these three variables were included in models predicting the effect size (which represents the developmental effect). Then we again tested the *Q* statistic (i.e., the residual sums of square) to determine whether the residual variability was greater than would be expected by chance, given a single population effect size. Among these 23 symptom domains, 19 (15 of the DSM symptoms and 4 of the associated symptoms) continued to show significant variability (see Table 3).

This significant variability has two implications. First, it would not be appropriate to compute mean effect sizes for each of the 19 symptoms, because the degree of variability within each of the symptoms indicates that a single mean for each of the symptoms could not adequately represent the effect sizes (Hedges & Olkin, 1985). Consequently, for these symptoms it was not appropriate to test whether the mean effect size differed from zero. Second, this significant variability implies that for these 19 symptoms, there were developmental effects but that the developmental effect for each of the symptoms was not consistent across the various studies in this sample, even when controlling for gender composition, type of sample, and informant.

For each of these symptom domains we assessed the effects of the three potential moderaters (gender composition, type of sample, and informant) in separate models, to determine their individual contributions to effect size variability. Of these 77 analyses (only symptom domains showing significant effect size variability were included in this analysis and for some symptom domains, tests of these predictors were not possible because of empty cells), 10 produced significant effects ($p < .05$): 3 involving the gender composition of the sample, 5 involving the type of sample, and 2 involving the informant. Predominantly female samples (those with 50% or more females) showed an essentially zero effect size (absolute value less than 0.10) for somatic concerns, externalizing problems, and hallucinations (i.e., among females, there were minimal developmental differences for these three symptom areas). In contrast, predominantly male samples showed an effect size of –0.25 for somatic concerns (i.e., among males, older developmental groups were reported to show 0.25 standard deviations less somatic concerns than younger developmental groups), –0.18 for externalizing problems, and –0.36 for hallucinations. Thus, the levels in these three apparently unrelated associated symptom areas did not vary significantly for females across developmental level but decreased with increasing developmental level for males.

In regard to type of sample, the effect size for hypersomnia for normal samples was essentially zero (absolute value less than 0.10), 0.25 for clinic samples, and 0.50 for diagnosed samples. For verbal sadness, the effect sizes for normal and diagnosed samples were essentially zero, whereas the effect size for clinic samples was 0.27. For psychomotor retardation, the effect size for normal samples was essentially zero, whereas the effect sizes for diagnosed and clinic samples were 0.31 and 0.33, respectively. For anorexia, the effect size was –0.14 for normal samples, essentially zero for clinic samples, and 0.48 for diagnosed samples. For social withdrawal, the effect size for normal samples was essentially zero, whereas the effect sizes for diagnosed and clinic samples were 0.13 and 0.14, respectively. To summarize, for a subset of symptom areas, particularly vegetative symptom areas, diagnostic and clinic samples tended to show

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

Table 3. *Effect size (ES) heterogeneity and mean by symptom*

Symptom	*N*	*Q*	Mean ES
1. Agitation	11	16.76****	
2. Anhedonia	15	3.17	0.11****
3. Anorexia	15	12.52**	
4. Appetite problems	5	5.07**	
5. Concentration	13	18.82****	
6. Fatigue	18	28.30****	
7. Guilt	16	14.08*	
8. Hopelessness	14	1.62	0.19***
9. Hyperexia	10	20.99****	
10. Hypersomnia	12	5.03	0.16***
11. Insomnia	14	9.24*	
12. Irritable	14	10.67*	
13. Retardation	15	16.48**	
14. Sadness (nonverbal)	12	17.40***	
15. Sadness (verbal)	13	12.05**	
16. Self-esteem	19	62.67****	
17. Sleep problems	10	1.82	−.04
18. Suicide	20	40.51****	
19. Weight gain	6	1.48	0.28****
20. Weight loss	6	15.34***	
21. Anxiety	19	81.38****	
22. Delusions	5	23.71****	
23. Externalizing	15	11.99	−.04
24. Hallucinations	10	1.98	−.03
25. Increased energy	3	0.18	−.12**
26. Social withdrawal	16	8.06	0.09***
27. Somatic complaints	13	15.84**	
28. Somatic concerns	7	0.63	0.02
29. Worse in morning	4	34.23****	

Note: Q, *Q* statistic for testing the heterogeneity of effect sizes (Hedges & Olkin, 1985).
*$p < .10$. **$p < .05$. ***$p < .01$. ****$p < .001$. For the probability that the heterogeneity differs significantly from chance expectations or that the mean effect is not equal to zero.

larger (and usually positive) effect sizes than normal samples.

In regard to informant, the effect size for parents was essentially zero for hallucinations and somatic concerns. In contrast, the effect size for interviews was –0.36 for hallucinations and 0.33 for somatic concerns; the effect size for somatic concerns for self-report was –0.28.

Finally, for each symptom that did not show significant variability and thus could be represented by a common effect size, we performed a weighted (by the inverse of the variance) *t* test to assess whether the average effect size for the symptom differed significantly from zero. Although these tests were based on relatively few observations (the average test involved about 12 observations), four of the five DSM symptoms and two of five associated symptoms showed significant developmental effects (see Table 3). Of these six effects, five domains (anhedonia, hopelessness, hypersomnia, weight gain, and social withdrawal) showed higher symptom levels among the more developmentally advanced subjects and one (increased energy) showed decreased symptom levels among the more developmentally advanced subjects. Taken together, these effects may suggest that higher levels of some vegetative or motivational symptoms may be found among more developmentally advanced subjects, although the pattern is far from conclusive.

Limitations. One limitation of the symptom

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

level studies just reviewed is that they tested, in essence, the main effect for developmental level on the various depressive symptoms. That is, these studies were based on a one-way ANOVA model, with developmental group as the independent variable and symptom level as the dependent variable (item level = developmental group). This main effect approach, however, fails to consider the possibility that age-related changes in symptom levels could reflect normal developmental change unrelated to depression (Digdon & Gotlib, 1985) or may be related to changes in syndromes other than depression (e.g., changes in concentration problems may be related to changes in attention-deficit/hyperactivity disorder).

To illustrate this potential complication, suppose, for example, that an investigator found that adolescents experience more fatigue than children. Although fatigue is an important symptom of depression, such a finding by itself leaves unaddressed the issue of whether this increase is related to depression. Rather, it is possible that this increase reflects the fact that in general, adolescents experience more fatigue than children. The use of a diagnosed sample does not resolve this problem, because developmental differences produced by this analytic approach could still represent normal development, even in a diagnosed sample.

On the other hand, this strategy could produce nonsignificant results, even if developmental differences did exist, if normal developmental change and depression-related change ran in opposite directions. For example, if fatigue increased in general with age but decreased with age in relation to depression, then it would be possible to find nonsignificant differences for this symptom. Moreover, the relative importance of a symptom vis-à-vis depression might change with age, even if the level did not. That is, children and adolescents normatively might experience the same level of fatigue, yet only in adolescents might this be part of the broader, depressive syndrome.

A related problem is that an individual symptom might increase, not because there has been a developmental change in the pattern of depression but rather because there has been a change in the overall severity of depression for all of the depression symptoms. If this were the case, there would be no change in the pattern of symptoms, but the individual symptoms would show developmental effects.

These issues suggest that the relative levels of the symptom in depressed versus nondepressed subjects as well as in younger versus older subjects should be taken into account. This could be accomplished by including depression status (i.e., depressed or nondepressed if using categorical variables; level of depression if using continuous variables) as a factor in the ANOVA model:

$$\begin{aligned}\text{item} = {} & \text{developmental group} \\ & + \text{Depression status} \\ & + \text{Developmental group} \\ & \quad \text{by Depression status.}\end{aligned}$$

In this approach, the interaction between depression status and developmental group would be the effect of interest. This factor represents the extent to which a developmental difference in the level of a symptom was specific to depression or was part of a more general change unrelated to depression.

In fact, our moderator analyses involving the type of sample represent a similar form of analysis. In our meta-analysis, because the effect size for a particular symptom represents developmental effects for a particular symptom, the effects of type of sample test whether the effect of development differs as a function of the type of sample, which in the above equation is represented by the Developmental Group × Depression Status (type of Sample) interaction. Table 4, which lists mean effect sizes for normal and diagnostic samples, illustrates this. However, one drawback to making these comparisons at the across-study (i.e., meta-analytic) level is that the number of observations may be very low (see Table 4).

In addition, both within- and across-study approaches may contain an element of tautology. As part of such an analysis one must first provide categorizations vis-à-vis depression for both children and adolescents. Yet it is the validity of such a categorization that is un-

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

Table 4. *Mean effect sizes and sample sizes by sample type*

Symptom	Normal	Diagnosed
Agitation	–0.43 (3)	0.17 (6)
Anhedonia	0.06 (6)	0.14 (6)
Anorexia	–0.14 (6)	0.48 (6)
Anxiety	0.01 (8)	0.12 (5)
Concentration	–0.10 (3)	–0.18 (6)
Fatigue	0.10 (7)	0.52 (6)
Guilt	–0.00 (7)	–0.04 (4)
Hallucinations	–0.01 (3)	–0.22 (4)
Hopelessness	0.09 (5)	0.25 (6)
Hyperexia	0.15 (3)	0.12 (4)
Hypersomnia	0.04 (5)	0.50 (4)
Insomnia	–0.12 (5)	0.15 (6)
Irritable	0.06 (6)	0.01 (4)
Retardation	0.08 (5)	0.31 (6)
Sadness (nonverbal)	0.10 (4)	–0.37 (4)
Sadness (verbal)	0.06 (6)	0.01 (4)
Self-esteem	0.06 (8)	–0.23 (6)
Social withdrawal	0.02 (7)	0.13 (4)
Somatic complaints	0.01 (5)	–0.21 (4)
Suicide	0.09 (9)	–0.46 (6)

Note: The sample sizes are in parentheses.

clear and is the raison d'être for the analysis. Further, as noted above, one runs the risks that are associated with defining an affective sample; that is, one's selection criteria may serve to equate the "depressed" children and adolescents, thus changing the meaning and significance of the interaction test.

Review of studies at the syndrome level. In contrast to the ANOVA-type approach used to assess symptom-level differences, correlational strategies such as factor analysis assess the impact of developmental level on relations among symptoms and may be less subject to the limitations of the symptom level approach just noted, at least insofar as one is interested in questions of structure rather than epidemiology. If used appropriately, correlational strategies will be less dependent on a priori notions, and hence less likely to be tautological. They involve relatively few preconceived notions regarding what is or is not part of depression and approach the issue from a syndromal level by assessing relations among symptoms rather than levels of individual symptoms.

The simplest correlational approach involves comparison of item–total score correlations across age groups. If a symptom correlates more highly with the total score (i.e., depression) for, say, the adolescents than for the children then this particular symptom is more involved in the depression of the adolescents than the children. This approach, which is similar (although not identical) to a factor analysis restricted to a single general factor, can be applied to data not originally intended for this form of developmental comparison (see below, Smucker, Craighead, Craighead, & Green, 1986).

A more sophisticated correlational approach than the item–total score strategy involves developmental comparisons of factor structures. A few researchers (e.g., Ryan et al., 1987) have used factor analysis to explore such differences by comparing the results of separate (for different developmental groups) factor analyses of ratings of depressive symptoms. Such an approach avoids the necessity of a strict a priori assumption of what constitutes depression in the different developmental groups, because it is the empirical analysis of the data that reveals this. However, developmental differences in multifactor solutions may be subtle, highly complex, and difficult to interpret (see, e.g., Weiss et al., 1991).

More readily interpretable results can be produced by restricting the factor solution to a single, general factor upon which all items potentially load (see e.g., Weiss et al., 1992). This general factor will represent what is most common among the items for each group; insofar as the items have been chosen appropriately (i.e., the majority being depression related), loadings on the general factor will represent the extent to which each item is related to a general syndrome of depression. By comparing the loadings across developmental groups, one can evaluate whether the extent to which each symptom is part of a general syndrome of depression varies as a function of developmental level.

Whereas we were able to identify 16 studies that provided information regarding the symptom prevalence/level question, only 5 studies were located that could provide correlational data on developmental differences in

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

depression. The first of these was based on data that Achenbach and Edelbrock (1983)[8] provided in their manual for the parent-report Child Behavior Checklist (CBCL). As part of their reliability and validity analyses, these authors conducted factor analyses of the CBCL for different age and sex groups. Although the male children produced a "depressed" factor that seemed to focus on affective and cognitive components of depression, the male adolescents failed to produce a depression factor. The factor that they produced most similar to depression was an "uncommunicative" factor that focused on shy, withdrawn behavior. The female children similarly produced a "depressed" factor focusing on affective and cognitive symptoms, whereas the female adolescents produced a "depressed withdrawal" factor that focused more on withdrawn behavior than on cognitive symptoms. To summarize, depression in children appeared to involve cognitive symptoms such as guilt and low self-esteem whereas depression in adolescents appeared to involve withdrawn behavior, at least as based on parent-report CBCL data. It is important to note also that for none of these child and adolescent groups did "depression" include any of the vegetative symptoms (e.g., sleep disturbance) that are an important part of adult depression (American Psychiatric Association, 1994).

Smucker et. al. (1986) reported normative and reliability data for the CDI. As part of this analysis, they reported the item–total score correlations separately for grades 3–6 and 7–9. We tested whether these item–total score correlations differed by grade. Across 26 CDI items, only 1 item, "I do not have any friends," differed significantly. Younger children produced a higher item–total score correlation, suggesting that this item was more a part of depression for younger than older children; however, given the number of tests, this single significant result should be evaluated with caution. Thus, analyses of Smucker et al.'s (1986) data support developmental continuity of the symptoms that comprise depression. However, the generalizability and significance of these results may be limited by the relatively narrow age range of the comparison (a mean difference of 3.5 years between age groups); that is, the narrow age range may have been too small to discern developmental effects.

Ryan et al. (1987) compared the factor structures of K-SADS symptom ratings for a group of prepubertal versus pubertal children who had received affective disorder diagnoses. The authors stated that the two groups produced "essentially the same" factor structures, with both groups producing five factor solutions with several similar factors, although the child results were somewhat less well articulated (i.e., the factors were somewhat less internally consistent). However, the separate child and adolescent factor structures were not provided in the article, and it may be necessary to approach this "essentially the same" conclusion with some caution. In their conclusions regarding symptom prevalence findings, these authors stated that there were "relatively few" differences between children and adolescents vis-à-vis symptom-level differences, although in their sample there were significant developmental differences for over one third of the DSM depressive symptoms.

In the fourth study, Weiss et al. (1992) used confirmatory factor analysis to compare the CDI factor structure for 515 children versus 515 adolescents. They found that 10 of 26 CDI item loadings on a general depression factor showed significant developmental differences. Items relating to guilt and to externalizing problems tended to be more a part of depression for the children, whereas for the adolescents, vegetative items (anhedonia and fatigue), affective items (sadness, loneliness, irritability), and items relating to concerns about the future tended to be more a part of depression for the adolescents. Further, four of nine symptom categories showed significant developmental differences in their correlations with the general factor. Paralleling the item-level analyses, externalizing behavior and guilt were more strongly related to depression in children than in adolescents, whereas affec-

8. More recently, Achenbach (1991) provided updated factor analyses of the CBCL with a larger sample. These analyses were not included here, because the Achenbach analyses were intentionally structured to remove developmental differences.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

tive symptoms (e.g., sadness) and concerns about the future were more strongly related to depression in adolescents than children. Thus, these authors found a substantial number of developmental differences at the syndrome level of depression, some of which paralleled the findings of Achenbach and Edelbrock (1983); for example, cognitive symptoms such as guilt were more a part of depression for children than adolescents. As with all studies using clinic-referred samples, it is possible that the apparent group effects may be due to group differences in the symptoms that elicit clinic referral. However, although it is clear how differences in clinic referral patterns might influence differences in mean levels of symptoms, it is less obvious how differences in clinic referral patterns might influence differences in relations among symptoms.

In the fifth study, Messer and colleagues (1995) compared the factor structure for a 13-item depression inventory among a large sample of boys in Grades 1 through 10. The authors concluded that a single-factor model most parsimoniously explained the structure of the questionnaire at all grade levels, and they reported that the rank ordering of the loadings was fairly consistent from Grade 1 to Grade 10 (r = .64) and highly consistent across shorter time periods (mean $r \approx .81$). The authors also reported that the mean factor loading showed a consistent and moderately large increase across grade (increasing about 30% from .61 in the first grade to .80 in the sixth grade). This suggests that coherence of the depression factor increased with grade-level, perhaps because the older children had a greater ability to develop a "depressed" self-schema (Messer et al., 1995). Thus, whereas the authors' results appear to support developmental continuity of depression in regard to the relative importance of symptoms, their results do suggest developmental differences in regard to the coherence or internal consistency of the factor. In interpreting these results, it is important to note that the sample only included males, which may be an important limitation to generalizability given that the prevalence for depression in females is approximately twice the prevalence for males, at least postpubertally (Angold & Rutter, 1992). It also is important to note that the depression measure did not include certain biological symptoms (e.g., sleep; appetite problems), which also may limit generalizability.

In summary, the results of these five studies were not consistent. Whereas two studies suggested substantial developmental differences, two others suggested relatively few differences and one study provided some support for both the developmental difference and developmental isomorphism positions. Important factors influencing the generalizability of the specific studies include a narrow age range, potential subjectivity in evaluating comparisons, use of clinic-referred samples, and use of a sample containing males only.

Limitations. The main limitation of the syndromal approach is not so much inherent to the approach but rather a function of methodological choice. If a questionnaire does not include a symptom or symptom area, then it cannot produce a developmental difference in that area. If a questionnaire only includes symptoms based on adult criteria for depression, then the results of the analyses will be circular in that other symptoms specific to childhood depression will not be identified. For instance, although the CBCL covers a wide range of symptoms, it does not include items assessing anhedonia, and thus if there were developmental differences in this symptom area, they would not be identified if the CBCL were used.

Discussion

The two broad goals of this paper were to examine the validity of the hypothesis that developmental level has relatively little influence on the phenomenology of depression and provide suggestions regarding the most appropriate and effective ways in which this question can be addressed. Contrary to what appears to be the current consensus, the results of this review suggested that it may be premature to conclude that depression is developmentally isomorphic at either the symptom or syndrome level. Although not large in an absolute sense, the number of studies in this area had the potential to provide at least preliminary support for one position or the other, if

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

results were consistent. However, the available evidence was not sufficiently consistent with either hypothesis to reach a firm conclusion. In regard to our symptom-level review, of 29 depressive symptoms evaluated, 19 showed significant variability across studies in the meta-analysis, even after the effects of gender, type of sample, and informant were controlled. This indicates that for these 19 symptom areas there were developmental effects but that some as-yet unknown, between-study factors influenced the magnitude of these effects. Moreover, although sample sizes were small (in the meta-analysis, the average comparison contained about 12 studies), of the 10 symptoms that did not show significant variability, 6 showed significant developmental effects. Five of these six tests (anhedonia, hopelessness, hypersomnia, weight gain, and social withdrawal) indicated higher levels of symptoms among more the developmentally advanced, with some indication that the symptoms showing this effect tended to be motivational or vegetative. Diagnostic and clinic samples tended to show larger (and usually positive) effect sizes, indicating that researchers are less likely to find developmental differences in normal samples, although this was by no means an absolute finding, given that 19 of the tests of sample type failed to produce a significant effect.

At the syndrome level, three studies specifically examined syndromal differences and two more studies were amenable to secondary data analysis in the current review. The results of these studies were not consistent: two studies appeared to support developmental isomorphism, two studies appeared to support developmental differences in depression, and one provided some support for each position.

It is important to note that these results in no way suggest that it is not possible to identify children who fit adult diagnostic criteria for depression (e.g., Ryan et al., 1987). However, this fact by itself provides little evidence for the patterning of depression in children or adolescents. This is because the presence of children fitting adult criteria does not preclude the possibility that there are other symptoms, unique to children, that are an important part of childhood depression, nor does the existence of children fitting adult criteria provide any information about whether the number of children fitting the adult pattern is greater than that expected due to chance covariance of symptoms.[9]

One consequence of the field's apparent acceptance of developmental isomorphism for depression (e.g., Carlson & Kashani, 1988; Ryan et al., 1987) is both an empirical and conceptual neglect of this question, and many of the studies we reviewed were conducted over a decade ago. The results presented here indicate, however, that there is reason to continue to conduct studies to address the question of whether there are developmental differences in the phenomenology of depression.

In interpreting our results, it is important to consider the potential impact of what Rosenthal (1979) called the "file drawer problem," which refers to the tendency for studies producing null findings to be rejected for publication and thus end up in a file drawer rather than in the literature. If reviewers and editors in fact have been biased toward rejecting studies with nonsignificant findings, then the frequency of significant differences found in this review may have been inflated. It obviously is difficult to judge accurately the extent to which this problem influenced our results, but several factors suggest that this potential problem may have had a less than substantial impact on our sample. First, many of the studies reviewed here focused on other issues in addition to developmental differences in individual symptoms. This should have lessened the impact of the aforementioned publication bias, because significant findings relevant to the other issues could foster publication. And second, because the studies in our meta-analysis focused on many symptoms, it was possible for a study to have a large number of nonsignificant results yet still have enough

9. The same note of caution may need to be sounded for other syndromes as well. According to DSM-IV, adult anxiety disorders can be diagnosed in children with only minor developmental modifications (Allen, Leonard, & Swedo, 1995). Here again, the fact that children can be identified who fit adult criteria does not necessarily mean that anxiety disorders are phenotypically similar in children and adults or that they even represent syndromes or disorders in children.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

significant findings to warrant publication (e.g., Worchel et al., 1987). Thus, although it is not possible to be certain, it seems unlikely that the file drawer problem substantially influenced the results reviewed here.

We had anticipated that cognitive factors would be more closely tied to depression at higher developmental levels, because younger children might lack the cognitive ability to develop and sustain depressogenic cognitions. However, at least in some samples they apparently do not. Contrary to expectation, results of two syndromal-level studies (Achenbach & Edelbrock, 1983; Weiss et al., 1992) suggested that some cognitive factors, in particular guilt, were more a part of depression at lower developmental levels. One possible explanation for this finding is that, although children apparently have the ability to experience guilt, their ability to control such cognitions—particularly under conditions of stress—may be relatively limited (relative to adolescents), thus making these cognitions of guilt more depressogenic. This interpretation is post hoc, however, and requires validation.

The current review and its results provide several recommendations for future investigations. First, investigators should be clear about the question they are asking and the goals of their study. As we noted above, in the past many investigators have focused on the symptom-level question of "how do two groups of subjects of different developmental levels who fit diagnostic criteria for a mood disorder differ vis-à-vis depressive symptoms?" This question is different from what is probably the more important and interesting syndrome-level question: "how does the syndrome of depression differ in two groups of different developmental levels?"

In this regard, it is important to note that there is an unusual opportunity in this area, in that there already exist a number of data sets that could provide new information about the syndrome-level question. Investigators who have made developmental comparisons at the symptom level can reanalyze their data to determine whether there are developmental differences at the syndrome level (i.e., in relations among symptoms). In this regard, we recommend that investigators use single-factor models, which may produce more interpretable results. Before undertaking such analyses, however, investigators first should determine if their data are appropriate for factor analytic comparisons (i.e., whether the sample sizes are sufficiently large).

In addition, it will be optimal to use more than one method for defining developmental groups (see, e.g., Kovacs & Paulauskas, 1984). This will require that the developmental level of the groups be contiguous or nearly contiguous; that is, the developmental levels of the groups must not be so far apart that different methods of defining developmental groups will yield the same groups. Along the same lines, developmental comparisons will be more powerful if they are planned and conducted using a theoretical perspective on developmental differences (e.g., cognitive, biological).

In regard to assessment methodology, we suggest using multiple informants who can either be compared or combined using statistical or clinical best-estimate strategies and investigators should use continuous measures as much as possible. It will also be important to use psychopathology measures that are as broad as possible, in order to cover not only all of the "standard" (i.e., adult) symptoms of depression but also symptoms that may be specific to children.

If it ultimately is concluded there are developmental differences in depression, it will be important to determine at what point the effects occur. Because biological and cognitive changes occur over a period of time rather than as a step function (Harter, 1983, 1999; Tanner, 1962), it is likely that developmental differences in depression also will occur over a period of time . However, we believe that there are two likely points around which the change might potentially begin: (a) puberty, when a host of biological and social changes begin (e.g., Angold & Rutter, 1992), or (b) around age 18, when many individuals experience major life changes such as leaving high school and home to go to work or college (Rao, Hammen, & Daley, 1999).

A number of limitations of the developmental analyses conducted so far should be highlighted. First, in future studies it will be

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

important to consider the fact that developmental effects may not be linear. That is, for some symptom areas the developmental differences between children and adolescents may not be the same as developmental differences between adolescents and adults. For instance, it is possible that biological difference may be more likely to be found in comparisons of children and adolescents, because of the biological differences that occur at puberty. Given this and similar possibilities (e.g., symptoms related to leaving one's parents' home at age 18), the fact that the majority of the studies reviewed here focused on child versus adolescent comparisons represents a limitation of the current literature. Although these obviously represent valuable comparisons, it will be important to increase the number of comparisons involving adults, because adult depression represents that "standard" to which depression in younger individuals is being compared. Similarly, the fact that most developmental comparisons have relied on self-reports of depressive symptoms also represents a limitation in that developmental level of the subject is then confounded with developmental level of the informant. This may be unavoidable, however, given the internal, subjective nature of most depressive symptoms. It does suggest, however, that developmentally sensitive interview assessments may be a particularly useful assessment approach.

Another important issue that should be highlighted is that in most if not all studies the effect of comorbidity has not been directly taken into account. Depression, conceptualized as a symptom, syndrome, or disorder, has been found to be moderately comorbid with a number of other psychopathologies, including anxiety and conduct problems (e.g., Brady & Kendall, 1992; Cole & Carpentieri, 1990; Puig–Antich, 1982). Thus, it is unclear to what extent the results of developmental analyses of depression may have been influenced by depression's co-occurrence with other disorders and symptoms (Angold & Costello, 1993). Specifically, at the item level, even among samples selected based on the presence of a diagnosis of depression, it is possible that developmental increases or decreases in symptom levels may be unrelated to depression but rather related to a developmental change in a correlated, comorbid condition. Similarly, it is possible that developmental changes in the syndrome/factor structure may be due to changes in comorbid conditions rather than depression itself.

To address this issue, it may be useful to include comorbid conditions as covariates in the developmental analyses. For instance, overall level of anxiety might be included as a covariate in the item-level model. Similarly, one might partial out an overall anxiety score from the covariance matrices used to conduct factor analyses. However, because the symptom item covariance partialed out in these analyses actually may be caused by depression rather than the comorbid condition, both the partialed and unpartialed analyses should be considered, which would add another layer of interpretative complexity.

Another limitation of the developmental analyses that have been conducted to date, in particular the developmental factor comparisons, is that they have assumed that all members of a particular developmental group have the same structure for depression. Probably because it is a basic assumption of exploratory factor analysis that all subjects have the same factor structure, developmental comparisons have implicitly assumed that each developmental group was composed entirely of one type or factor structure of depression. It is possible, however, that all members of a developmental group do not share the same structure. This could occur if development did not affect every individual equally (i.e., if there were moderators of the effects of development) or if there are distinct subtypes of depression. If this were the case, then development could be associated with differences in the prevalence of different syndromes, rather than across the board differences in factor structures for two developmental groups. Another issue related to factor analysis is that one must consider whether the depression scale is part of a broad or focused assessment instrument, as this may influence how informants respond (e.g., depression scales that are part of a broad instrument like the CBCL may tend to produce more heterogenous depression scales).

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

It should be noted that, although this paper has focused on relatively traditional conceptualizations of depression, there are other models of depression that we should consider. For instance, in the tripartite model proposed by Clark and Watson (1991), depression is defined by the concurrent presence of high negative affectivity (i.e., high levels of nonspecific, general distress, such as sadness, anxiety, and anger) and low positive affect (i.e., symptoms of anhedonia, such as a lack of enjoyment or involvement in activities that ordinarily are pleasurable). The ideas discussed in the current paper, such as potential sources of developmental differences for depression, could apply as well to this tripartite model. However, because developmental research has almost exclusively utilized traditional definitions of depression, we focused on these more traditional conceptualizations. This does present an interesting avenue for future research, however.

To conclude, if it turns out that there are no reliable differences between child and adult forms of depression, this would support further investigation of whether other aspects of depression in adults are applicable to children (e.g., etiology, treatment). If, on the other hand, it turns out that differences do exist, then a number of steps should follow. Most important would be longitudinal investigations to determine whether these differences represent heterotypic continuity or different psychopathological entities, to assess whether the developmental differences are the result of differences in the causes or consequences of depression, and to identify the specific factors that account for the developmental differences (e.g., physiological maturation, cognitive development, linguistic ability). First, however, the fundamental question of whether there are developmental differences in the symptoms that comprise the syndrome of depression remains to be answered.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*, 358–372.

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4–18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: Queen City Printers.

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: Parents' reports for normative and clinical samples. *Monographs of the Society for Research in Child Development*, *56*, 1–120.

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.

Allen, A. J., Leonard, H., & Swedo, S. E. (1995). Current knowledge of medications for the treatment of childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 976–986.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Angold, A., & Costello, E. J. (1993). Depressive comorbidity in children and adolescents: Empirical, theoretical, and methodological issues. *American Journal of Psychiatry, 150,* 1779–1791.

Angold, A., & Rutter, M. (1992). Effects of age and pubertal status on depression in a large clinical sample. *Development and Psychopathology, 4,* 5–28.

Angold, A., & Worthman, C. W. (1993). Puberty onset of gender differences in rates of depression: A developmental, epidemiologic and neuroendocrine perspective. *Journal of Affective Disorders, 29,* 145–158.

Asarnow, J. R., & Guthrie, D. (1989). Suicidal behavior, depression, and hopelessness in child psychiatric inpatients: A replication and extension. *Journal of Clinical Child Psychology, 18,* 129–136.

Baker, M., Dorzab, J., Winokur, G., & Cadoret, R. J. (1971). Depressive disease: Classification and clinical characteristics. *Comprehensive Psychiatry, 4,* 354–365.

Beck, A. T., Steer, R. A., Beck, J. S., & Newman, C. F. (1993). Hopelessness, depression, suicidal ideation, and clinical diagnosis of depression. *Suicide and Life Threatening Behavior, 23,* 139–145.

Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., Kaufman, J., Dahl, R., Perel, J., & Nelson, B. (1996). Childhood and adolescent depression: A review of the past 10 years: Part I. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1427–1439.

Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models.* Newbury Park, CA: Sage Publications.

Borchardt, C. M., & Meller, W. H. (1996). Symptoms of affective disorder in pre-adolescent vs. adolescent inpatients. *Journal of Adolescence, 19,* 155–161.

Brady, E., & Kendall, P. (1992). Comorbidity of anxiety and depression in children and adolescents. *Psychological Bulletin, 111,* 244–255.

Carlson, G., & Strober, M. (1979). Affective disorders in

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

adolescence. *Psychiatric Clinics of North America, 2,* 511–526.
Carlson, G. A., & Garber, J. (1986). Developmental issues in the classification of depression in children. In D. M. Rutter, C. E. Izard & P. B. Read (Eds.), *Depression in young people: Developmental and clinical perspectives* (pp. 399–434). New York: Guilford Press.
Carlson, G. A., & Kashani, J. H. (1988). Phenomenology of major depression from childhood through adulthood: Analysis of three studies. *American Journal of Psychiatry, 145,* 1222–1225.
Chilcoat, H. D., & Breslau, N. (1997). Does psychiatric history bias mothers' reports? An application of a new analytic approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 971–979.
Cicchetti, D., & Schneider–Rosen, K. (1984). Toward a transactional model of childhood depression. In D. Cicchetti & K. Schneider–Rosen (Eds.), *Childhood depression* (pp. 5–28). San Francisco, CA: Jossey–Bass.
Cicchetti, D., & Toth, S. L. (1991). A developmental perspective on internalizing and externalizing disorder. In. D. Cicchetti & S. L. Toth (Eds.), *Internalizing and externalizing expression of dysfunction* (pp. 1–19), Hillsdale, NJ: Erlbaum.
Cicchetti, D., & Toth, S. L. (1998). The development of depression in children and adolescents. *American Psychologist, 53,* 221–241.
Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316–336.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45,* 1304–1312.
Cole, D. A., & Carpentieri, S. (1990). Social status and the comorbidity of child depression and conduct disorder. *Journal of Consulting and Clinical Psychology, 58,* 748–757.
Cole, D. A., & Turner, J. E. (1993). Models of cognitive mediation and moderation in child depression. *Journal of Abnormal Psychology, 102,* 271–281.
Compas, B. E., Ey, S., & Grant, K. E. (1993). Taxonomy, assessment, and diagnosis of depression during adolescence. *Psychological Bulletin, 114,* 323–344.
Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York: Russell Sage.
Curry, J. F., & Craighead, W. E. (1993). Depression. In T. H. Ollendick, & M. Hersen (Eds), *Handbook of child and adolescent assessment* (pp. 251–268). Boston, MA: Allyn & Bacon.
Damon, W., & Hart, D. (1982). The development of self-understanding from infancy through adolescence. *Child Development, 53,* 841–864.
Digdon, N., & Gotlib, I. H. (1985). Developmental considerations in the study of childhood depression. *Developmental Review, 5,* 162–199.
Edelbrock, C., Costello, A. J., Dulcan, M. K., Conover, N. C., & Kala, R. (1986). Parent–child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry, 27,* 181–190.
Emslie, G. J., Rush, A. J., Weinberg, W. A., Gullion, C. M., Rintelmann, J., & Hughes, C. W. (1997). Recurrence of major depressive disorder in hospitalized children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 785–792.
Fergusson, D. M., Lynskey, M. T., & Horwood, L. J. (1993). The effect of maternal depression on maternal ratings of child behavior. *Journal of Abnormal Child Psychology, 21,* 245–269.
Finney, D. J. (1971). *Probit analysis* (3rd ed.). Cambridge: Cambridge University Press.
Fiske, S., & Taylor, S. (1991). *Social cognition*. Reading, MA: Addison–Wesley.
Flavell, J. H. (1985). *Cognitive development*. Englewood Cliffs, NJ: Prentice–Hall.
Friedlander, S. (1988). Learned helplessness in children: Perception of control and causal attributions. In D. C. Morrison (Ed.), *Organizing early experience: Imagination and cognition in childhood* (pp. 33–53). Amityville, NY: Baywood Publishing.
Friedman, R. C., Hurt, S. W., Clarkin, J. F., Corn, R., & Aronoff, M. S. (1983). Symptoms of depression among adolescents and young adults. *Journal of Affective Disorders, 5,* 37–43.
Garber, J. (1984). The developmental progression of depression in female children. In D. Cicchetti & K. Schneider–Rosen (Eds.), *Childhood depression* (pp. 29–58). San Francisco, CA: Jossey–Bass.
Garber, J., & Kaminski, K. M. (2000). Laboratory and performance-based measures of depression in children and adolescents. *Journal of Clinical Child Psychology, 29,* 509–525.
Gaylin, W. (1968). *The meaning of despair: Psychoanalytic contributions to the understanding of depression.* New York: Science House.
Goodyer, I. M. (1996). Physical symptoms and depressive disorders in childhood and adolescence. *Journal of Psychosomatic Research, 41,* 405–408.
Harrington, R., Fudge, H., Rutter, M., Pickles, A., & Hill, J. (1990). Adult outcomes of childhood and adolescent depression, I: Psychiatric status. *Archives of General Psychiatry, 47,* 465–473.
Harris, P. L. (1989). *Children and emotion: The development of psychological understanding*. Oxford: Blackwell.
Hart, D., & Damon, W. (1986). Developmental trends in self-understanding. *Social Cognition, 4,* 388–407.
Harter, S. (1983). Developmental perspectives on the self system. In E. M. Hetherington (Ed.), *Handbook of child psychology: Vol. 4. Socialization, personality, and social development* (pp. 275–385). New York: Wiley.
Harter, S. (1993). Causes and consequences of low self-esteem in children and adolescents. In R. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 87–116). New York: Plenum Press.
Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
Inamdar, S. C., Siomopoulos, G., Osborn, M., & Bianchi, E. C. (1979). Phenomenology associated with depressed moods in adolescents. *American Journal of Psychiatry, 136,* 156–159.
Kagan, J. (1982). The emergence of self. *Journal of Child Psychology and Psychiatry, 23,* 363–382.
Kagan, J., & Moss, H. A. (1962). *Birth to maturity: A study in psychological development*. New York: Wiley.
Kashani, J. H., Rosenberg, T. K., & Reid, J. C. (1989). Developmental perspectives in child and adolescent

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

depressive symptomatology. *American Journal of Psychiatry, 146,* 871–875.
Kazdin, A. E. (1987). Assessment of childhood depression: Current issues and strategies. *Behavioral Assessment, 9,* 291–319.
Kirk, R. E. (1982). *Experimental design* (2nd ed.). Monterey, CA: Brooks/Cole.
Knight, D., Hensley, V. R., & Waters, B. (1988). Validation of the Children's Depression Scale and the Children's Depression Inventory in a prepubertal sample. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 29,* 853–863.
Kovacs, M. (1989). Affective disorders in children and adolescents. *American Psychologist, 44,* 209–215.
Kovacs, M. (1992). *Children's Depression Inventory: CDI manual*. North Tonawanda, NY: Multi-Health Systems.
Kovacs, M., Feinberg, T. L., Crouse–Novak, M., Paulauskas, S. L., & Finkelstein, R. (1984). Depressive disorders in childhood: II. A longitudinal study of the risk for a subsequent major depression. *Archives of General Psychiatry, 41,* 643–649.
Kovacs, M., & Paulauskas, S. L. (1984). Developmental stage and the expression of depressive disorders in children: An empirical analysis. In D. Cicchetti & K. Schneider–Rosen (Eds.), *Childhood depression* (pp. 59–80). San Francisco, CA: Jossey–Bass.
Links, P. S., Boyle, M. H., & Offord, D. R. (1989). The prevalence of emotional disorder in children. *Journal of Mental and Nervous Disorder, 177,* 85–91.
Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer–Loeber, M. (1991). Differences and similarities between children, mothers, and teachers as informants on disruptive child behavior. *Journal of Abnormal Child Psychology, 19,* 75–95.
Loeber, R., & Hay, D. (1997). Key issues in the development of aggression and violence from childhood to early adulthood. *Annual Review of Psychology, 48,* 371–410.
Messer, S. C., Angold, A., Costello, E. J., Loeber, R., van Kammen, W., & Stouthamer–Loeber, M. (1995). *Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor comparison and structure across development.* Unpublished manuscript.
Mitchell, J., McCauley, E., Burke, P. M., & Moss, S. J. (1988). Phenomenology of depression in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 12–20.
Moreau, D., Mufson, L., Weissman, M. M., & Klerman, G. L. (1991). Interpersonal psychotherapy for adolescent depression: Description of modification and preliminary application. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 642–651.
Neter, J., Wasserman, W., & Kutner, M. H. (1989). *Applied linear regression models* (2nd ed.). Homewood, IL: Irwin.
Nolen–Hoeksema, S., Girgus, J. S., & Seligman, M. E. P. (1992). Predictors and consequences of childhood depressive symptoms: A 5-year longitudinal study. *Journal of Abnormal Psychology, 101,* 405–422.
Piaget, J. (1971). *The construction of reality in the child.* New York: Ballantine.
Post, R. M. (1992). Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *American Journal of Psychiatry*, *149*, 999–1010.
Post, R. M., & Weiss, S. R. B. (1998). Sensitization and kindling phenomena in mood, anxiety, and obsessive-compulsive disorders: The role of serotonergic mechanisms in illness progression. *Biological Psychiatry, 44,* 193–206.
Poznanski, E. O., Cook, S. C., & Carroll, B. J. (1979). A depression rating scale for children. *Pediatrics, 64,* 442–450.
Puig–Antich, J. (1982). Major depression and conduct disorder in prepuberty. *Journal of the American Academy of Child Psychiatry, 21,* 118–128.
Rao, U., Hammen, C., & Daley, S. E. (1999). Continuity of depression during the transition to adulthood: A 5-year longitudinal study of young women. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 908–915.
Rao, U., Weissman, M. M., Martin, J. A., & Hammond, R. W. (1993). Childhood depression and risk of suicide: A preliminary report of a longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 21–27.
Reich, W., & Earls, F. (1987). Rules for making psychiatric diagnoses in children on the basis of multiple sources of information: Preliminary strategies. *Journal of Abnormal Child Psychology, 15,* 601–616.
Renouf, A. G., & Kovacs, M. (1994). Concordance between mothers' reports and children's self-reports of depressive symptoms: A longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 208–216.
Rie, H. E. (1966). Depression in childhood: A survey of some pertinent contributions. *Journal of the American Academy of Child Psychiatry, 5,* 653–685.
Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.
Ruble, D. N., & Frey, K. S. (1987). Social comparison and self-evaluation in the classroom: Developmental changes in knowledge and function. In J. C. Masters & P. William (Eds.), *Social comparison, social justice, and relative deprivation: Theoretical, empirical, and policy perspectives* (pp. 81–104). Hillsdale, NJ: Erlbaum.
Ruble, D. N., & Rholes, W. S. (1981). The development of children's perceptions and attributions about their social world. In J. H. Harvey, W. Ices, & R. F. Kidded (Eds.), *New directions in attribution research* (Vol. 3, pp. 3–36). Hillsdale, NJ: Erlbaum.
Rutter, M. (1986). The developmental psychopathology of depression: Issues and perspectives. In M. Rutter, C. E. Lizard & P. B. Read (Eds.), *Depression in young people: Developmental and clinical perspectives* (pp. 3–30). New York: Guilford Press.
Ryan, N. D., Puig–Antich, J., Ambrosian, P., Rabinovich, H., Robinson, D., Nelson, B., Iyengar, S., & Twomey, J. (1987). The clinical picture of major depression in children and adolescents. *Archives of General Psychiatry, 44,* 854–861.
Schwartz, G. E., & Trabasso, R. (1984). Children's understanding of emotions. In C. Izard, J. Kagan, & R. Zajonc (Eds.), *Emotions, cognitions, and behavior* (pp. 409–437). New York: Cambridge University Press.
Selman, R. L. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. New York: Academic Press.
Singer, D. G., & Revenson, T. A. (1996). *A Piaget primer: How a child thinks*. New York: International Universities Press.
Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The*

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press

benefits of psychotherapy. Baltimore, MD: Johns Hopkins University Press.
Smucker, M. R., Craighead, W. E., Craighead, L. W., & Green, B. J. (1986). Normative and reliability data for the Children's Depression Inventory. *Journal of Abnormal Child Psychology, 14,* 25–39.
Stehouwer, R. S., Bultsma, C. A., & Blackford, I. T. (1985). Developmental differences in depression: Cognitive–perceptual distortion in adolescent versus adult female depressives. *Adolescence, 20,* 291–299.
Stipek, D. & MacIver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60,* 521–528.
Strober, M., Green, J., & Carlson, G. (1981). Phenomenology and subtypes of major depressive disorder in adolescence. *Journal of Affective Disorders, 3,* 773–782.
Tanner, J. M. (1962). *Growth at adolescence* (2nd ed.). Oxford: Blackwell Scientific.
Weiner, B., & Graham, S. (1985). An attributional approach to emotional development. In C. E. Izard & J. Kagan (Ed.), *Emotions, cognition, and behavior* (pp. 167–191). New York: Cambridge University Press.
Weiss, B., Weisz, J. R., Politano, M., Carey, M., Nelson, W. M., & Finch, A. J. (1991). Developmental differences in the factor structure of the Children's Depression Inventory. *Psychological Assessment, 3,* 38–45.
Weiss, B., Weisz, J. R., Politano, M., Carey, M., Nelson, W. M., & Finch, A. J. (1992). Relations among self-reported depressive symptoms in clinic-referred children versus adolescents. *Journal of Abnormal Psychology, 101,* 391–397.
Weissman, M. M., Wolk, S., Goldstein, R. B., Moreau, D., Adams, P., Greenwald, S. Klier, C. M., Ryan, N. D., Dahl, R. E., & Wickramaratne, P. (1999). Depressed adolescents grown up. *Journal of the American Medical Association, 17,* 7–13.
Weisz, J. R., Southam–Gerow, M. A., & McCarty, C. A. (2001). Control-related beliefs and depressive symptoms in clinic-referred children and adolescents: Developmental differences and model specificity. *Journal of Abnormal Psychology, 110,* 97–109.
Wickramaratne, P. J., & Weissman, M. M. (1998). Onset of psychopathology in offspring by developmental phase and parental depression. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 933–942.
Woodruff, R. A., Murphy, G. E., & Herjanic, M. (1967). The natural history of affective disorders—I. Symptoms of 72 patients at the time of index hospital admission. *Journal of Psychiatric Research, 5,* 255–263.
Worchel, F., Nolan, B., & Wilson, V. (1987). New perspectives on child and adolescent depression. *Journal of School Psychology, 25,* 411–414.
Zahn–Waxler, C., Chapman, M., & Cummings, E. M. (1984). Cognitive and social development in infants and toddlers with a bipolar parent. *Child Psychiatry and Human Development, 15,* 75–85.

https://doi.org/10.1017/S0954579403000221 Published online by Cambridge University Press