Journal of Pediatric Psychology, Vol. 27, No. 1, 2002, pp. 5–18

Collecting and Managing Multisource and Multimethod Data in Studies of Pediatric Populations

Grayson N. Holmbeck,[1] PhD, ***Susan T. Li,***[1] PhD, ***Jennifer Verrill Schurman,***[2] MA, ***Deborah Friedman,***[1] BA, and ***Rachael Millstein Coakley,***[1] MA
[1]Loyola University of Chicago and [2]University of Cincinnati

Objective: **To provide recommendations for the collection and management of multisource and multimethod data in studies of children and adolescents with pediatric conditions.**
Methods: **We discuss limitations of single-source and single-method data collection strategies. We review strategies for collecting and managing multisource and multimethod data, including coverage of the literature on level of agreement across sources, strengths and weaknesses of various source and method aggregation strategies, and methods of examining discrepancies between sources.**
Results: **Multisource and multimethod data collection strategies enable researchers to rule out alternative explanations for their findings and pose research questions that would probably not be testable with single-source, single-method data sets.**
Conclusions: **We emphasize the utility of multisource and multimethod data and provide recommendations for future work.**

Key words: *multisource data; multimethod data; optimal informants; research methodology; discrepancies; congruence.*

Once an investigator has formulated a research question, multiple decisions need to be made regarding the research design. One set of decisions involves the nature of the data to be collected (Rosenthal & Rosnow, 1991). What sources or informants will provide the data (e.g., will one use child report? parent, teacher, or peer report? reports of medical personnel? raters of videotaped family interactions?)[1] and what methods will be used to collect the data? (e.g., will one collect information using questionnaires? observational methods? sociometric methods? medical chart reviews?). How one answers such questions is determined, in part, by the type of research conducted, as well as the nature of the hypotheses posed. Answers to these questions are critical because they will determine the type of data one will have available when at-

[1]We prefer the term *source* to *informant* since the latter typically refers only to people (e.g., children, peers, mothers, fathers, teachers, physicians), whereas the former can refer to people and additional "sources" of data (e.g., psychophysiological, medical chart). Also, certain distinctions between sources and methods are in need of clarification. Whereas some distinctions are clear (e.g., if one collects mother-reported questionnaire data, mothers are the *source* and the questionnaire is the *method*), others are not. For the purposes of this review, an observational procedure will be considered a *method,* whereas a rater of observational data will be considered a *source.* A medical chart review is considered a data collection *method;* the information contained in the medical chart is considered a *source.*

All correspondence should be sent to Grayson N. Holmbeck, Loyola University of Chicago, Department of Psychology, 6525 N. Sheridan Road, Chicago, Illinois 60626. E-mail: gholmbe@luc.edu.

tempting to answer a research question and will affect one's ability to rule out alternative explanations for one's findings.

Research in pediatric psychology varies widely with respect to the types of sources and methods employed. Although some of the limitations of single-source and single-method data have been noted (e.g., Bank & Patterson, 1992; La Greca & Lemanek, 1996), an extended discussion of such limitations, along with a thorough discussion of possible alternatives, is not available. Given that the use of multisource and multimethod data complicates data reduction, data analysis, and interpretation of findings (Jacob & Tennenbaum, 1988; La Greca & Lemanek, 1996), such issues will be reviewed in this article. The purpose of this article is to (1) review limitations of single-source and single-method data collection strategies, (2) discuss strategies for collecting and managing multisource and multimethod data, and (3) present a set of recommendations for investigators seeking to conduct multisource and multimethod research with pediatric populations.

Limitations of a Single-Source, Single-Method Design

A single-source and single-method study uses a single respondent (e.g., mother) and a single data collection method (e.g., questionnaires) to answer a research question of interest. Drotar (1997) found that in 16 of 57 studies he reviewed (28%), mothers were the sole reporters of data on family functioning *and* child outcome. In 79% of the 57 studies, questionnaires were the sole method employed. With respect to reporters of family functioning in family and health research, Patterson (1990) found that 80% of the studies she reviewed relied on a single informant, with questionnaire/interview strategies the most common method. Finally, Drotar (1997) found that only 5% of the studies he reviewed included an observational component, and Patterson (1990) found that only 15% of family and health studies used "outsider" ratings of family process.[2]

Although a single-source and single-method approach may be an appropriate data collection strategy in certain areas of research (see below), this strategy has inherent limitations. First, with such a methodological strategy, one is unable to rule out common method variance interpretations of the findings. Personality and cognitive characteristics of the informants (i.e., response biases; e.g., Youngstrom, Izard, & Ackerman, 1999), rather than associations between true score variance, may account for significant relations between variables. In fact, some studies in pediatric psychology have shown that correlations between *different* constructs from the perspective of the same source are higher than correlations between the *same* constructs from the perspective of different sources (e.g., Kronenberger, Carter, & Latta, 1997). Second, with a single-source/single-method design, one misses out on the opportunity to examine a number of interesting hypotheses. For example, when one has collected data from multiple sources, discrepancies between reports from different sources can become interesting predictor or criterion variables (e.g., discrepancies between family members with respect to their views of the family system; Paikoff, 1991). At a more basic level, the collection of multisource data provides a more complex view of the phenomena of interest and can include reports from understudied sources (e.g., teachers, fathers; La Greca & Lemanek, 1996). Third, as noted by La Greca and Lemanek (1996), exclusive reliance on certain sources may make it more difficult to assess certain types of functioning. For example, the assessment of a child's internal state (e.g., internalizing symptoms) may be more valid if self-reports are employed in addition to parent reports. Finally, when one collects data with a single method and fails to find an association between a predictor and an outcome, one will likely conclude (perhaps erroneously) that the predictor is not related to the outcome. For example, in a study of differential treatment of siblings in families who have children with cystic fibrosis, Quittner and Opipari (1994) found few differential treatment effects when using interview methodologies, yet found such effects when using a daily diary technique.

[2]We conducted a review of the *Journal of Pediatric Psychology* with respect to the types of sources assessed and methods used. Over a 6-year period, from 1994 to 1999, 57 of 201 papers (28%) were single-source, single-method; 27 (13%) were multisource, single-method; 19 (9%) were single-source, multimethod; and 98 (49%) were multisource, multimethod. Most of the 57 single-source, single-method studies employed either child or parent report (89%) and questionnaires (63%). Put another way, our review of the literature revealed that, regardless of method, 38% of the studies used a single source and, regardless of source, 42% of the studies used a single method. Even when multiple sources were used, they tended to be based on a combination of child reports, parent reports, and/or reviews of medical charts and physiological data. As a consequence, some sources were rarely used in the studies we reviewed. For example, across all studies, teachers were used in only 11 (5%) of the studies. Similarly, some methods were rarely used. Across all studies, observational measures were used in only 18% (n = 36) of the studies.

Despite the limitations of the single-source, single-method design, such a strategy may be appropriate (or even ideal) when addressing certain research questions. If one poses a very specific and focused research question, such a design may be appropriate. For example, if one seeks to determine the neuropsychological status (in comparison to national norms) of children with a particular pediatric condition, this type of design may be ideal. If one seeks to maximize the power to detect significant effects, a single-source and single-method study with a very large sample may be appropriate (Gray & Steinberg, 1999).

Moving Beyond Single-Source, Single-Method Data: Collecting and Managing Multisource and Multimethod Data

In this section, we discuss strategies for collecting data from multiple sources with multiple methods, as well as methods of handling differential response rates across sources.[3] The literature on "agreement" across sources and methods is reviewed, as well as the implications of such research for methods of managing data from multiple sources and methods. Strengths and weaknesses of the "optimal informant" strategy will be discussed (La Greca & Lemanek, 1996; Loeber, Green, Lahey, & Stouthamer-Loeber, 1989). Finally, we review the following methods for managing multisource and multimethod data: (1) data from individual sources/methods may be kept in their original, disaggregated form; (2) data from multiple individuals/methods can be aggregated by summing (standardized scores) across sources and methods; (3) scores can be aggregated via the use of latent variable models; and (4) discrepancies between sources or between methods may be calculated and used as predictor or outcome variables.

Collecting Multisource and Multimethod Data

Collecting multisource and multimethod data is resource-intensive; thus, selecting the appropriate informants for a given construct is an important methodological step. Typically, the nature of the construct determines, in part, the chosen respondents (La Greca & Lemanek, 1996). For family environment measures, for example, respondents often include mothers, fathers, a target child and siblings, and/or external observers. For measures of adjustment, the perspectives of target child, parents, mental health professionals, teachers, and peers may be assessed. If one seeks to collect multisource and multimethod data, there should also be a clear conceptual rationale for using each source and each method (La Greca & Lemanek, 1996). For example, teacher report of child adjustment could be used when examining whether mothers' views of the family environment are associated with adjustment in the child (assuming that associations between the two are of interest conceptually and contribute to existing knowledge in the field). In this way, one would be examining whether family relationships affect child adjustment across contexts (home → school), while at the same time reducing the problem of common method variance. Finally, developmental considerations should guide choices of sources and methods (see Table 1 in La Greca & Lemanek, 1996, for a list of optimal informants as a function of age of child and construct type).

One of the difficulties in collecting multiple reports involves the maintenance of consistent participation rates across informants. Some researchers have suggested that certain raters are particularly difficult to recruit and assess (e.g., fathers; Kazak, Segal-Andrews, & Johnson, 1995; Phares, 1992). Depending on the data analytic method used, complete data across all sources may be required or assumed by that analytic strategy. Indeed, problems arise when one attempts to conduct analyses (e.g., MANOVAs) that include variables based on responses of two types of informants (e.g., mothers and fathers), where the response rate is much lower for one type of informant. If the amount of missing data is large, one should consider running analyses separately for each respondent (see Holmbeck, Shapera, & Hommeyer, in press, for an example of this strategy).

Similarity of rated constructs across reporters is an important issue, as low agreement can result from poor construct measurement across sources. When measuring a construct with multiple sources, many studies simply rewrite items originally designed for one type of informant (e.g., children) for use with another type of informant (e.g., parents; see Cole, Hoffman, Tram, & Maxwell, 2000, for an

[3]We assumed that all statements and recommendations in this section refer to research that is multisource *and* multimethod. To improve readability, sometimes we refer only to sources *or* methods (e.g., "optimal informant"), even though the information provided could apply equally well to either.

example involving the Child Depression Inventory [CDI], and the Revised Children's Manifest Anxiety Scale [RCMAS]). It is often assumed that items will be appropriate and meaningful for both types of respondents. In cross-cultural research, where investigators seek to use similar measures across different ethnic populations, concerns over construct measurement equivalence are seriously considered (Knight, Tein, Shell, & Roosa, 1992). Unfortunately, measurement equivalence is often not considered when investigators seek to use similar measures across different informants. The cross-cultural literature may be particularly instructive for those seeking to conduct cross-informant research.

Certain data collection strategies in studies of pediatric samples can result in ambiguous findings. Researchers often collect data from children when they come to a medical clinic. Given that some symptoms may be situation-specific (e.g., internalizing symptoms), and that children are often more distressed in medical settings than in nonmedical settings, this method of data collection may artificially inflate scores obtained from self-report questionnaires and interviews. This strategy can also increase discrepancies between reports of children and parents. An additional concern is cross-source contamination. If one conducts a mail survey of parents and children, one is unable to determine the degree to which family members discuss their responses or whether they completed each other's questionnaires. A home visit methodology is preferable since researchers can monitor the independent completion of measures.

If one samples multiple informants, but does so exclusively with questionnaires, one has conducted a multisource, single-method study. One problem with an exclusive focus on self-report data is that such information may be biased, particularly given the difficulty most individuals have in recalling past events over long time frames. An alternative strategy would be to supplement the multi-informant questionnaire data with additional methods (Quittner & DiGirolamo, 1998). By increasing the number of methods, one usually increases the number of sources. For example, if family observational data are used (i.e., an additional method), we would also be adding "coders" as an additional source of data. Of course, a new method does not always bring with it a new source. With the daily diary technique (see Quittner & DiGirolamo, 1998, for a description), for example, one adds a new method but the number of sources remains unchanged.

Agreement Across Sources/Methods

In general, correlations that assess the level of agreement between sources or between methods are in the low-to-moderate range (Achenbach, McConaughy, & Howell, 1987), although there are exceptions (Cook & Goldstein, 1993; Schwarz, Barton-Henry, & Pruzinsky, 1985). In the child psychiatry and psychopathology literatures, agreement on diagnoses and ratings of child mental health outcomes and symptomatology has been a targeted area of study (e.g., Cantwell, Lewinsohn, Rohde, & Seeley, 1997). In a widely cited meta-analytic review, Achenbach et al. assessed level of agreement across studies on internalizing and externalizing symptomatology as reported by children, parents, mental health workers, teachers, and peers. Correlations between similar types of informants (e.g., mother, father) were higher than correlations between different types of informants (e.g., mother, teacher) or self-other correlations, but all were in the low-to-moderate range. Studies investigating source agreement among other variables (e.g., family environment, parenting) also revealed associations in the low-to-moderate range (e.g., Cook & Goldstein, 1993; Jacob & Windle, 1999; Melby, Conger, Ge, & Warner, 1995; Sessa, Avenevoli, Steinberg, & Morris, 2001). Given the low level of agreement, a number of investigators have recommended the use of multisource and multimethod data (e.g., Achenbach et al., 1987; Cantwell et al., 1997).

This low-to-moderate informant agreement has been problematic for both clinicians and researchers. Clinically, decisions made regarding weighting of sources can have serious consequences for diagnosis and treatment planning at the individual level (Power et al., 1998). Consequently, some have attempted to determine a "gold standard" approach for combining information from various sources to make diagnoses for children and adolescents (see Kazdin, 1994, for a review). Low informant concordance also has important implications for research methodology. Researchers studying a particular disorder in a population of children or adolescents must decide on the most appropriate informant for their research (La Greca & Lemanek, 1996). If they attempt to circumvent this design problem by including multiple informants, they must decide how to combine the varying reports when conducting data analyses (Shaffer, Lucas, & Richters, 1999).

Some researchers have found that certain respondents are underreporters, whereas others are

overreporters (e.g., Cole, Martin, Peeke, Seroczynski, & Fier, 1999). In the area of child pain, for example, parents usually underestimate their child's level of pain (e.g., Chambers, Reid, Craig, McGrath, & Finley, 1998), although the degree to which parents under- or overreport appears to vary as a function of context (e.g., the type of procedure experienced by a child; Chambers, Braaksma, Craig, Bennett, & Huntsman, 1999). Parents are also more likely to underestimate their child's level of substance use (Deffenbaugh, Hutchinson, & Blankschen, 1993; O'Donnell et al., 1998). Studies that examine under- and overreporting of adjustment report mixed results, with some finding that parents report more symptoms than children (e.g., Loeber et al., 1991), and others finding the reverse (e.g., Cantwell et al., 1997).

Finally, several studies have attempted to identify individual and family factors that may moderate levels of agreement between sources (Holmbeck, 1997). Putative moderators of agreement for child symptomatology include type of symptom, gender of child and parent, child age, child ethnicity, parental psychopathology, clinic versus nonpatient status, family cohesion and adaptability, life stress events, social desirability of the symptom, family size, and familiarity with the target (e.g., Kazdin, 1994; Youngstrom, Loeber, & Stouthamer-Loeber, 2000). Achenbach et al. (1987) investigated several moderators of agreement and found consistent evidence across studies that specific variables such as child age and type of outcome influenced agreement levels.

The "Optimal Informant" Strategy

Why is there low-to-moderate agreement among sources and among methods (Kazdin, 1994)? It may be that (1) each informant or method provides a unique, but meaningful perspective on child behavior, because such behavior differs across settings (Hart, Lahey, Loeber, & Hanson, 1994; Shaffer et al., 1999); (2) different informants and methods have different types of biases; or (3) some raters and methods are more valid and reliable than others for certain symptoms and behaviors (Hart et al., 1994; Shaffer et al., 1999). The latter perspective has led some to search for "optimal informants" (e.g., Loeber et al., 1989), particularly where multiple informants and methods are possible (e.g., see Table 1 in La Greca & Lemanek, 1996).

One approach to determining the "best reporter" for a construct is to isolate the person in the best position to rate a construct because of the environmental domain that he or she occupies (e.g., Kellam, 1990). Rational criteria are used to make such decisions. For example, teachers are considered the best reporters of classroom behavior since they are the natural reporters of behavior in the classroom domain. In contrast, direct observation may be considered the most valid measure of parental discipline (Dishion, 1990). One advantage of the "best reporter" approach is that it eliminates the need for multiple reports of the same construct. A disadvantage of this approach is that it ignores the possibility that other raters could contribute unique and important information.

An empirical alternative to this rational approach of choosing an optimal informant or method involves the assessment of conditional probabilities for agreement across sources (Loeber et al., 1989; Stanger & Lewis, 1993). With this strategy, one assesses the conditional probability that one informant will endorse a particular symptom, given that another informant has reported the symptom. Research using this approach has revealed, for example, that children provide little unique information beyond parent and teacher report of hyperactivity or oppositional behavior. On the other hand, children seem to offer important unique information for the assessment of conduct problems (Loeber et al., 1989; Power et al., 1998). Despite the findings in this line of research, some continue to recommend that all reporters should be treated equally (Piacentini, Cohen, & Cohen, 1992).

Handling Multisource and Multimethod Data

As it is not always possible to identify an optimal reporter for a given construct, many investigators collect multisource and multimethod data. As discussed below, there are several ways to handle such data. Table I details the advantages and disadvantages of each of these approaches.

Keeping Source/Method Data Disaggregated

The simplest solution to managing multisource and multimethod data is to leave the data in its original disaggregated form. The advantage of this strategy is that it allows the investigator to examine the effects of each source or method separately (Table I). For example, with disaggregated variables, one can

Table I. Methods of Handling Multisource and Multimethod Data

Method of Handling Data	Advantages	Disadvantages
Keeping source/method data disaggregated	1. One can examine effects of each source/method separately 2. One can examine associations between variables based on different sources/methods 3. Useful in small *n* studies variance 4. Useful when *r*s among predictors are low	1. Large number of analyses may be necessary (increases chance of Type I errors) 2. Does not distinguish between shared and nonshared source/method
Source/method aggregation via summing	1. Decreases number of analyses 2. Useful in small *n* studies when latent variable modeling is not possible 3. Allows one to manage effects of missing data	1. Does not distinguish between shared and nonshared source/method variance 2. Effects of individual sources/methods cannot be examined 3. Sources/methods must be at least moderately correlated
Source/method aggregation via latent variable modeling	1. Extracts shared perspective across sources/methods 2. Model fit indices are computed 3. A theory-driven data management strategy	1. Large *n* is needed 2. Divergent perspectives are considered error 3. Sources/methods must be moderately correlated
Discrepancies between sources or between methods	1. One can examine impact of discrepancies on variables of interest 2. One can examine discrepancies as predictors or outcomes 3. A theory-driven data management strategy	1. One must choose among multiple (and sometimes nonoverlapping) data analytic strategies 2. Findings are sometimes difficult to interpret, depending on the data analytic strategy employed

determine whether mother, father, and child reports of the family environment are differentially related to teacher report of child adjustment. One can offset the impact of common method variance by examining only associations between variables assessed from the perspective of different informants. In sum, this disaggregation strategy is useful, particularly with small studies, where latent variable modeling is not possible, or when there are low correlations among informants. Unfortunately, this strategy often results in a large number of analyses, which increases the chance of Type I errors. This approach also does not distinguish between shared and nonshared method variance.

Source/Method Aggregation via Summing

A different approach to analyzing multisource data is to aggregate or combine data across sources based on a simple linear model (Schwarz et al., 1985). By combining reporters, the researcher captures multiple perspectives while improving predictive power. This approach can be used with any number of sources and in studies with small sample sizes. It is also possible to differentially weight various raters in the linear combinations to emphasize the perspective of one rater over another.

An advantage of this approach is that aggregation can simplify the data analysis by reducing the number of analyses and, therefore, the Type I error rate (Table I). In addition, the composite variable formed from the linear combination should have a higher reliability than any of the individual scales (Cook & Goldstein, 1993; Schwarz et al., 1985). As suggested by Cook and Goldstein, however, "one of the shortcomings of aggregating over the reports of multiple informants is that the ability to distinguish variance due to the unique perspective of a rater and the perspective common to all raters is lost" (p. 1378).

To aggregate, one typically collapses across measures only when these measures are associated to some degree (indicating some shared variance across informants). To determine whether measures across informants are associated, a variety of approaches is available. If one has a large enough sample (see Bentler & Chou, 1987, and Tabachnick & Fidell, 1996, for sample size recommendations), a confirmatory factor analytic strategy may

be useful, provided one has a theory for how the measures should covary. Because samples sizes are typically quite small in studies of pediatric populations, the investigator may wish to examine a correlation matrix for evidence of significant associations. In this case, the investigator can establish a correlational criterion (e.g., $r \geq .40$; Chassin, Pitts, & Prost, in press) to designate when measures from different sources or methods can be aggregated. Once aggregation is justified, variables based on different metrics can be standardized (i.e., converted to *z*-scores) and then summed.

In the case of missing values for one or more informants, one could (1) use only those participants who have complete data across all informants (i.e., the listwise deletion method; in our view, this is the least desirable strategy, particularly in small pediatric studies); (2) use all available data from all possible sources (i.e., the available case or pairwise deletion method; one disadvantage of this strategy is that aggregated scores will be based on input from different informants across participants); or (3) when data from multiple informants ($\geq$3) are available, one could make decisions about the number of available informants that are required to compute an aggregate total for each participant.

The issue of missing data is particularly salient in research utilizing multisource, multimethod data. Indeed, the potential for missing data increases with the number of sources and methods. In addition to the listwise deletion and available case methods, other common missing data techniques include mean substitution/replacement (of which there are several types; Raaijmakers, 1999), simple or multiple regression imputation, and more complex approaches such as multiple imputation and maximum likelihood estimation (Arbuckle, 1996; Little & Schenker, 1995; Raaijmakers, 1999; Schafer, 1997; see Acock, 1997, for a very readable discussion of most available strategies). A pivotal assumption underlying many of these techniques is that the data are "missing at random" (MAR) or "missing completely at random" (MCAR; Roth, 1994). This can become a stumbling point for pediatric studies in which the availability of data from a particular source or method is confounded with subtypes of the study population (e.g., when medical test information is more complete for those with more severe forms of an illness). Current techniques are being evaluated for use with nonrandomly missing data (Arbuckle, 1996; Kromrey & Hines, 1994; Raaijmakers, 1999), but these developments lag behind applications for data sets in which MAR and MCAR assumptions hold. At what point does missing data become a severe problem? Some authors suggest that when incomplete cases comprise only a small fraction of the data set (e.g., 5%), complete case deletion is a relatively efficient method (Schafer, 1997). However, when missing data occur in as much as 20% or more of the cases, other techniques are warranted (Roth, 1994). Of course, the best strategy is to prevent attrition and the occurrence of missing data in the first place (see Mason, 1999, for several helpful suggestions).

Source/Method Aggregation via Latent Variable Modeling

To better capture the "shared perspective" discussed in the previous section, some have advocated the use of latent variable modeling (e.g., Bank, Dishion, Skinner, & Patterson, 1992; Bank & Patterson, 1992; Bray, Maxwell, & Cole, 1995; Cook & Goldstein, 1993; Jacob & Windle, 1999). With this approach, multiple reports are used as measured indicators of a latent construct. Systematic variance common to all reporters is retained in the construct (and is assumed to be measured without error), and any nonshared variance becomes error variance (made up of the informants' unique perspectives *and* measurement error; Table I).

This perspective assumes that the most valid underlying construct is the one shared across reporters and methods. However, raters often hold divergent perceptions of the same phenomenon. In our view, this is a limitation of a latent modeling approach since any unique variance that a single reporter would contribute, which is not shared by other reporters, is not included in the latent variable (and is treated as "error"). Someone who might be considered the best reporter of a construct using a rational approach may have the lowest loading on the latent variable and be virtually excluded from the analysis due to low correlations with other sources.

A latent variable modeling approach to multiple reporter data is most appropriate when sources or raters are moderately correlated. Problems develop when reports are weakly correlated (e.g., a stable latent construct may not be identifiable). The construct formed may essentially be a single reporter construct, despite the inclusion of multiple sources, due to the high loading of a strong source and low

loadings for other sources. When this occurs for *both* predictors and outcomes, the systematic variance of one reporter may account for the predictor → criterion relationship (a relationship that may not be shared across raters). As a solution to this problem, Bank et al. (1992) recommend that independent and dependent variables in the same analysis be assessed with nonoverlapping sources or methods. However, obtaining enough reports to have different, nonoverlapping reporters across predictor and outcome latent variables may be difficult (Cook & Goldstein, 1993). An additional problem with latent variable modeling is the necessity of having a large sample, a problematic requirement for researchers studying pediatric populations.

Discrepancies Between Sources or Between Methods

An additional use of multisource data is to examine the divergent views of multiple informants as variables of interest (e.g., Paikoff, 1991). Rather than viewing such differences of opinion as error (i.e., deviations from the "true score"), these differing perspectives may be an important window on the functioning of a child or family system (Table I).

Types of Discrepancies. Both between-person (e.g., parent perceptions vs. child perceptions) and within-person (e.g., discrepancies between child reports of ideal vs. actual self-concept) have been examined (Holmbeck & O'Donnell, 1991; Paikoff, 1991). Put another way, studies have focused on multisource, single-target (i.e., between-person) discrepancies or single-source, multitarget (i.e., within-person) discrepancies. Investigators have also examined the "meaning" of various types of discrepancies by examining associations between discrepancies and child adjustment outcomes (Paikoff, 1991). Moreover, discrepancies can themselves be examined as outcomes. For example, Collins, Laursen, Mortensen, Luebker, and Ferreira (1997) report that certain types of discrepancies are more likely at certain ages and at certain developmental transition points. It would be interesting to know what sorts of conditions produce more discrepancies between parents and children or, alternatively, what variables produce more congruence.

A Developmental Perspective on Source and Method Discrepancies. Although discrepancies between sources/methods can occur because of biases and measurement error or because some reporters and methods are superior to others, a developmental perspective can also help to explain the occurrence of source/method discrepancies. For example, a child's level of social and cognitive maturation may affect parent-child concordance across childhood and adolescence, although few studies have examined such associations (Kazdin, 1994). In general, one might expect a linear function for parent-child concordance, such that children will agree more with parents as their cognitive level comes closer to the cognitive level of their parents. The picture becomes more complex, however, when we consider social maturation issues. Adolescents may be less likely than younger children to report behaviors and events to their parents (Collins et al., 1997; Verhulst & van der Ende, 1992). Thus, although adolescents may be more able to recognize internal experiences with age, they also may be less likely to communicate this information to their parents during this period of development (Edelbrock, Costello, Dulcan, Conover, & Kalas, 1986). Moreover, parents have fewer opportunities to observe the behavior of their adolescent offspring, making parents increasingly reliant on their child's verbal report of symptoms. Consistent with expectations based on cognitive and social development factors, Verhulst and colleagues (Verhulst, Althaus, & Berden, 1987; Verhulst & van der Ende, 1992) found that parent-child concordance with regard to internalizing symptoms increases from childhood to adolescence but then decreases with age during the adolescent period. Consequently, parent-child concordance over time may exist as a curvilinear, rather than a linear, function.

A different hypothesis might be advanced if we examined discrepancies in reports of parent-child interaction instead of internalizing symptoms (Collins, 1990; Collins et al., 1997; Holmbeck, 1996). For example, with respect to pediatric populations, discrepancies between parent and child reports of who is responsible for making certain that the child adheres to certain aspects of medical regimens are likely to peak during early adolescence (Anderson, Ho, Brackett, Finkelstein, & Laffel, 1997). Thus, we have different developmental predictions depending on the nature of the discrepancy examined. Such differences in the developmental patterning of congruence and incongruence suggests that future work should focus on clarifying the relationship between a child's stage of development and parent-child concordance by including multiple markers of development (both cognitive and social) and by examining child age as a continuous variable.

Methods for Computing Discrepancies and Data An-

alytic Strategies. The most common strategy for computing discrepancy scores is to use difference scores (e.g., mother minus child; Ohannessian, Lerner, Lerner, & von Eye, 1995; Welsh, Galliher, & Powers, 1998) or absolute values of difference scores (e.g., Dekovic, Noom, & Meeus, 1997; Feinberg, Howe, Reiss, & Hetherington, 2000). These scores are then used as predictors or outcomes in data analyses. Although a number of scholars have written extensively concerning the utility of difference scores (e.g., Carlton-Ford, Paikoff, & Brooks-Gunn, 1991; Gottman & Krokoff, 1990; Rogosa & Willett, 1983; Rovine, 1994), there are still problems with the use of difference scores when examining divergent viewpoints across informants (Griffin, Murray, & Gonzalez, 1999). In other words, even though difference scores may not be flawed mathematically for certain types of statistical analyses, they may not be useful or meaningful for all purposes.

Suppose we compute a difference score between mothers' and adolescents' views of who is in charge of certain medical adherence tasks (a hypothetical 10-item scale, with total scores that range from 10 to 40, where higher scores indicate that the adolescent is more in control). If we subtract adolescent perceptions from mother perceptions, difference scores for this variable can range from −30 to +30, with positive scores indicating that the mother sees the adolescent as more in control of adherence tasks than does the adolescent and a negative score indicating that the adolescent sees himself or herself as more in control than does the mother. The problem with this distribution of difference scores is twofold (Griffin et al., 1999): (1) incongruence is highest at the extremes of the continuum, with congruence being highest at the 0 point; in other words, we have a curvilinear distribution ranging from incongruence (e.g., −30) to congruence (e.g., 0) to incongruence (e.g., + 30); and (2) congruence scores are difficult to interpret because a score of 0 can result from the difference of two high scores (40–40 = 0) or the difference of two low scores (10–10 = 0); thus, very different mother-adolescent dyads can yield identical scores. To deal with problem 1, one could assess the predictive utility of a quadratic difference score term (the difference score squared; e.g., Welsh et al., 1998), which would allow one to determine if incongruence scores (at the two ends of the continuum of difference scores) are more or less predictive than congruence scores (at the middle of the continuum of difference scores). Unfortunately, interpretation of such quadratic terms becomes ambiguous when the quadratic curve does not bend near 0 (the point of congruence at the middle of the continuum). Findings based on absolute values of difference scores are somewhat easier to interpret because they address problem 1. With such scores, the continuum ranges from congruence to incongruence. On the other hand, problem 2 is not addressed with absolute value scores. Moreover, any information on the direction of incongruence (mother-high/adolescent-low vs. mother-low/adolescent-high) is lost with this strategy.

Thus, an alternative is needed. A strategy that has proved useful in examining divergent perspectives as independent variables is to test, with regression analyses, the significance of the interaction of the two perspectives in predicting outcomes (Holmbeck & O'Donnell, 1991). Such a strategy preserves the continuous nature of the independent variables and avoids problems inherent in using difference scores. If we continue with the earlier hypothetical example and assume that we are interested in predicting an objective outcome measure, such as metabolic control (e.g., glycosylated hemoglobin; Anderson et al., 1997), one would enter the mother and adolescent main effects in the first step of a regression, followed by their interaction (the product of mother and adolescent report) in the second step. One would then be able to determine the independent contributions of each report to metabolic control as well as examine whether divergent views are associated with higher or lower metabolic control. When all variables are normally distributed and when the independent variables are uncorrelated, a significant interaction with no accompanying main effects would indicate (depending on the direction of the interaction) that the congruence or discrepancy groups yielded the highest values on the outcome (i.e., a crossed interaction; Holmbeck & O'Donnell, 1991; see Aiken & West, 1991, and Holmbeck, 1997, 2002, for detailed discussions regarding interpretation of interaction effects).

To more fully interpret the significant interaction, it is sometimes useful to isolate four mother-adolescent congruence/incongruence groups (i.e., mother-high/adolescent-high, mother-low/adolescent-high, mother-high/adolescent-low, mother-low/adolescent-low) via median split (or based on a cutoff determined rationally) and then examine the significance of group differences. How one forms the four groups is critical; one needs to be thoughtful about how the cutoff is established. When me-

dian split strategies are applied to two independent variables (e.g., mother report of a family relationship variable and child report on the same variable), the amount of power available to detect significant effects decreases *and* the rate of false statistical significance increases. This is particularly problematic when the two independent variables are highly correlated (Maxwell & Delaney, 1993).

Another post-hoc interpretive strategy is to use difference scores to generate one's congruence/incongruence groups. That is, one could divide the sample into the following three groups: (1) the 25% of the sample with the lowest difference scores (usually negative), (2) the 25% with the highest scores (usually positive), and (3) the middle 50% with roughly equal scores. These percentages would need to be adjusted if the distribution of difference scores was not normally distributed about 0. One could then form two groups from the middle 50% (e.g., the 25% who are mother-high/adolescent-high and the 25% who are mother-low/adolescent-low). This latter post-hoc strategy, in conjunction with the initial regression approach, would make maximum use of the data in its continuous form and would also allow one to interpret significant interactions by creating groups true to the underlying difference scores (while, at the same time, differentiating between high-high and low-low families).

Instead of examining main effects and interactions, an alternative regression approach to discrepancies is to compute residuals for the regression of one informant's scores onto another informant's scores (Cole et al., 1999; Griffin et al., 1999). That is, if one were interested in examining discrepancies between mother and child report, the child report variable could be used as the dependent variable and the mother report variable could be entered as an independent variable in the first step of a multiple regression equation. The residual (i.e., "error") that remains could be considered a form of "discrepancy" between the two raters. For example, a given family's residual may represent the degree to which one reporter's (e.g., the child's) rating of adolescent control deviates from the "best fitting" regression line that represents the association between the two informants' ratings (e.g., mother and child report). If the values for the two reporters are displayed graphically (with one on the *x*-axis and one on the *y*-axis), such a deviation is represented by the vertical distance between a given data point and the regression line. A positive residual would indicate that the child's rating was greater than expected; a negative residual would indicate that the child's rating was less than expected. These residuals could then be used as outcomes or predictors. One drawback of this strategy is that the residuals may bear little relation to the actual "raw" difference scores. Finally, Griffin et al. (1999) offer an additional alternative, based on a piece-wise linear regression approach, that may be useful in certain situations.

When using divergent views as *dependent* measures, one could use a repeated measures ANOVA or MANOVA procedure, whereby mother and adolescent report are used as repeated measures within-subjects variables for the same construct (Fitzmaurice, Laird, Zahner, & Daskalakis, 1995). Continuing with the example above, if one employed an additional between-subjects measure as an independent variable (e.g., gender), a significant interaction between independent variables (e.g., gender $\times$ reporter) would indicate that differences between mother and adolescent perceptions of who is in charge of certain adherence tasks varies as a function of whether the child is a boy or a girl.

Conclusions

Given concerns with single-source, single-method research strategies and in light of our overview of methods for collecting and managing multisource and multimethod data, we provide recommendations that may guide future work in this field:

1. If possible and if relevant to one's research question, consider using multiple sources and multiple methods to assess each construct of interest. Selection of sources and methods should be theory-driven and should be based on developmental considerations. Use of multisource and multimethod data is particularly important when conducting correlational/regression-oriented research (La Greca & Lemanek, 1996), because it allows one to examine associations between predictors and outcomes that do not share common method or source variance.

2. Consider the possibility that multiple versions of the same measure (i.e., parent and child versions of a parenting behavior measure) may be nonequivalent. Strategies are available for examining factorial invariance across reporters (e.g., Cole et al., 2000).

3. Attempt to reduce "cross-informant contamination" where one informant monitors or influences the responses of another informant. Depending on the research question, it may be preferable to collect psychosocial adjustment data in non-

clinic settings, given that stress may be temporarily elevated in clinic settings (i.e., "setting contamination").

4. When selecting measures, one should initially decide whether an "optimal informant/method" approach will be used *or* whether a multisource and multimethod approach will be used.

5. If one selects a multisource and multimethod approach, different methods of aggregating across sources/methods are available. The advantages and disadvantages of each should be considered (Table I).

6. Consider the relevance of examining discrepancies between sources/methods either as predictors or outcomes. If this strategy is of interest, try to avoid the use of difference scores (given the interpretation problems that result when such scores are used). Alternative strategies were reviewed.

Acknowledgments

Completion of this manuscript was supported by a grant (12-FY01-0098) from the March of Dimes Birth Defects Foundation.

Received November 30, 1999; revisions received December 1, 2000, and May 25, 2001; accepted May 28, 2001

References

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.

Acock, A. C. (1997). Working with missing values. *Family Science Review, 10,* 76–102.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Anderson, B., Ho, J., Brackett, J., Finkelstein, D., & Laffel, L. (1997). Parental involvement in diabetes management tasks: Relationships to blood glucose monitoring adherence and metabolic control in young adolescents with insulin-dependent diabetes mellitus. *Journal of Pediatrics, 130,* 257–265.

Arbuckle, J. L. (1996). Full information estimation in presence of incomplete data. In G. A. Marcoulides & R. E. Schumaker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum.

Bank, L., Dishion, T., Skinner, M., & Patterson, G. R. (1992). Method variance in structural equation modeling: Living with "glop." In G. R. Patterson (Ed.) *Depression and aggression in family interaction.* Hillsdale, NJ: Lawrence Erlbaum.

Bank, L., & Patterson, G. R. (1992). The use of structural equation modeling in combining data from different types of assessment. In J. C. Rosen & P. McReynolds (Eds.), *Advances in psychological assessment* (vol. 8; pp. 41–74). New York: Plenum.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods and Research, 16,* 78–117.

Bray, J. H., Maxwell, S. E., & Cole, D. (1995). Multivariate statistics for family psychology research. *Journal of Family Psychology, 9,* 144–160.

Cantwell, D. P., Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1997). Correspondence between adolescent report and parent report of psychiatric diagnostic data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 610–619.

Carlton-Ford, S. L., Paikoff, R. L., & Brooks-Gunn, J. (1991). Methodological issues in the study of divergent views of the family. In R. L. Paikoff (Ed.), *New directions for child development: Shared views in the family during adolescence* (no. 51, pp. 87–102). San Francisco: Jossey-Bass.

Chambers, C. T., Braaksma, D. N., Craig, K. D., Bennett, S. M., & Huntsman, E. (1999, April). *Parental factors influencing agreement between parent and child reports of acute pain.* Paper presented at the Seventh Florida Conference on Child Health Psychology, Gainesville, FL.

Chambers, C. T., Reid, G. J., Craig, K. D., McGrath, P. J., & Finley, G. A. (1998). Agreement between child and parent reports of pain. *Clinical Journal of Pain, 14,* 336–342.

Chassin, L., Pitts, S. C., & Prost, J. (in press). Heavy drinking trajectories from adolescence to young adulthood in a high risk sample: Predictors and substance abuse outcomes. *Journal of Consulting and Clinical Psychology.*

Cole, D. A., Hoffman, K., Tram, J. M., & Maxwell, S. E. (2000). Structural differences in parent and child reports of children's symptoms of depression and anxiety. *Psychological Assessment, 12,* 174–185.

Cole, D. A., Martin, J. M., Peeke, L. A., Seroczynski, A. D., & Fier, J. (1999). Children's over- and underestimation of academic competence: A longitudinal study of gender differences, depression, and anxiety. *Child Development, 70,* 459–473.

Collins, W. A. (1990). Parent-child relationships in the transition to adolescence: Continuity and change in interaction, affect, and cognition. In R. Montemayor, G. R. Adams, & T. P. Gullotta (Eds.), *From childhood to adolescence: A transitional period?* (pp. 85–106). Newbury Park, CA: Sage.

Collins, W. A., Laursen, B., Mortensen, N., Luebker, C., & Ferreira, M. (1997). Conflict processes and transitions in parent and peer relationships: Implications for autonomy and regulation. *Journal of Adolescent Research, 12,* 178–198.

Cook, W. L., & Goldstein, M. J. (1993). Multiple perspectives on family relationships: A latent variables model. *Child Development, 64,* 1377–1388.

Deffenbaugh, K. B., Hutchinson, R. L., & Blankschen, M. P. (1993). Substance use among youth (grades 4–12) in rural Indiana: Students' reported use vs. parents' perceptions of students' use. *Journal of Alcohol and Drug Education, 39,* 19–33.

Dekovic, M., Noom, M. J., & Meeus, W. (1997). Expectations regarding development during adolescence: Parental and adolescent perceptions. *Journal of Youth and Adolescence, 26,* 253–272.

Dishion, T. J. (1990). The family ecology of boys' peer relations in middle childhood. *Child Development, 61,* 1289–1307.

Drotar, D. (1997). Relating parent and family functioning to the psychological adjustment of children with chronic health conditions: What have we learned? What do we need to know? *Journal of Pediatric Psychology, 22,* 149–165.

Edelbrock, C., Costello, A. J., Dulcan, M. K., Conover, N. C., & Kalas, R. (1986). Parent-child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry, 27,* 181–190.

Feinberg, M. E., Howe, G. W., Reiss, D., & Hetherington, E. M. (2000). Relationship between perceptual differences of parenting and adolescent antisocial behavior and depressive symptoms. *Journal of Family Psychology, 14,* 531–555.

Fitzmaurice, G. M., Laird, N. M., Zahner, G. E. P., & Daskalakis, C. (1995). Bivariate logistic regression analysis of childhood psychopathology ratings using multiple informants. *American Journal of Epidemiology, 142,* 1194–1203.

Gottman, J. M., & Krokoff, L. J. (1990). Complex statistics are not always clearer than simple statistics: A reply to Woody and Costanzo. *Journal of Consulting and Clinical Psychology, 58,* 502–505.

Gray, M. R., & Steinberg, L. (1999). Unpacking authoritative parenting: Reassessing a multidimensional construct. *Journal of Marriage and the Family, 61,* 574–587.

Griffin, D., Murray, S., & Gonzalez, R. (1999). Difference score correlations in relationship research: A conceptual primer. *Personal Relationships, 6,* 505–518.

Hart, E. L., Lahey, B. B., Loeber, R., & Hanson, K. S. (1994). Criterion validity of informants in the diagnosis of disruptive behavior disorders in children: A preliminary study. *Journal of Consulting and Clinical Psychology, 62,* 410–414.

Holmbeck, G. N. (1996). A model of family relational transformations during the transition to adolescence: Parent-adolescent conflict and adaptation. In J. A. Graber, J. Brooks-Gunn, & A. C. Petersen (Eds.), *Transitions through adolescence: Interpersonal domains and context* (pp. 167–199). Mahwah, NJ: Lawrence Erlbaum.

Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology, 65,* 599–610.

Holmbeck, G. N. (2002). Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of Pediatric Psychology, 27,* 87–96.

Holmbeck, G. N., & O'Donnell, K. (1991). Discrepancies between perceptions of decision making and behavioral autonomy. In R. L. Paikoff (Ed.), *New directions for child development: Shared views in the family during adolescence* (no. 51, pp. 51–69). San Francisco: Jossey-Bass.

Holmbeck, G. N., Shapera, W. E., & Hommeyer, J. S. (in press). Observed and perceived parenting behaviors and psychosocial adjustment in pre-adolescents with spina bifida. In B. K. Barber (Ed.), *Intrusive parenting: How psychological control affects children and adolescents*. Washington, DC: American Psychological Association Press.

Jacob, T., & Tennenbaum, D. L. (1988). Family assessment methods. In M. Rutter, A. H. Tuma, & I. S. Lann (Eds.), *Assessment and diagnosis in child psychopathology* (pp. 196–231). New York: Guilford.

Jacob, T., & Windle, M. (1999). Family assessment: Instrument dimensionality and correspondence across family reporters. *Journal of Family Psychology, 13,* 339–354.

Kazak, A. E., Segal-Andrews, A. M., & Johnson, K. (1995). Pediatric psychology research and practice: A family/systems approach. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 84–104). New York: Guilford.

Kazdin, A. E. (1994). Informant variability in the assessment of childhood depression. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 249–271). New York: Plenum Press.

Kellam, S. G. (1990). Developmental epidemiological framework for family research on depression and aggression. In G. R. Patterson (Ed.), *Depression and aggression in family interaction*. Hillsdale, NJ: Lawrence Erlbaum.

Knight, G. P., Tein, J. Y., Shell, R., & Roosa, M. (1992). The cross-ethnic equivalence of parenting and family interaction measures among Hispanic and Anglo-American families. *Child Development, 63,* 1392–1403.

Kromrey, J. D., & Hines, C. V. (1994). Nonrandomly missing data in multiple regression: An empirical comparison of common missing-data treatments. *Educational and Psychological Measurement, 54,* 573–593.

Kronenberger, W. G., Carter, B. C., & Latta, C. (1997, April). *Situational, respondent, and questionnaire-specific*

effects on the measurement of coping behaviors of adolescents with cystic fibrosis. Paper presented at the Florida Conference on Child Health Psychology, Gainesville, FL.

La Greca, A. M., & Lemanek, K. L. (1996). Editorial: Assessment as a process in pediatric psychology. *Journal of Pediatric Psychology, 21,* 137–151.

Little, R. J. A., & Schenker, N. (1995). Missing data. In G. Arminger, C. C. Clogg, & M. E. Sobel, (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 39–75). New York: Plenum Press.

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

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.

Mason, M. J. (1999). Methods, plainly speaking: A review of procedural and statistical methods for handling attrition and missing data in clinical research. *Measurement and Evaluation in Counseling and Development, 32,* 111–118.

Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin, 113,* 181–190.

Melby, J. N., Conger, R. D., Ge, X., & Warner, T. D. (1995). The use of structural equation modeling in assessing the quality of marital observations. *Journal of Family Psychology*, 9, 280–293.

O'Donnell, D., Biederman, J., Jones, J., Wilens, T. E., Milberger, S., Mick, E., & Faraone, S. V. (1998). Informativeness of child and parent reports on substance use disorders in a sample of ADHD probands, control probands, and their siblings. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 752–758.

Ohannessian, C. M., Lerner, R. M., Lerner, J. V., & von Eye, A. (1995). Discrepancies in adolescents' and parents' perceptions of family functioning and adolescent emotional adjustment. *Journal of Early Adolescence, 15,* 490–516.

Paikoff, R. L. (Ed.). (1991). *New directions for child development: Shared views in the family during adolescence* (no. 51). San Francisco: Jossey-Bass.

Patterson, J. M. (1990). Family and health research in the 1980s: A family scientist's perspective. *Family Systems Medicine, 8,* 422–434.

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

Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology, 20,* 51–63.

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

Quittner, A. L., & DiGirolamo, A. M. (1998). Family adaptation to childhood disability and illness. In R. T. Ammerman & J. V. Campo (Eds.), *Handbook of pediatric psychology and psychiatry* (vol. 2, pp. 70–102). Boston, MA: Allyn & Bacon.

Quittner, A. L., & Opipari, L. C. (1994). Differential treatment of siblings: Interview and diary analyses comparing two family contexts. *Child Development, 65,* 800–814.

Raaijmakers, Q. A. W. (1999). Effectiveness of different missing data treatments in surveys with Likert-type data: Introducing the relative mean substitution approach. *Educational and Psychological Measurement, 59,* 725–748.

Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement, 20,* 335–343.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology, 47,* 537–560.

Rovine, M. J. (1994). Estimating nonshared environment using sibling discrepancy scores. In E. M. Hetherington, D. Reiss, & R. Plomin (Eds.), *Separate social worlds of siblings: The impact of nonshared environment on development* (pp. 33–61). Hillsdale, NJ: Lawrence Erlbaum.

Schafer, J. (1997). *Analysis of incomplete multivariate data*. London: Chapman and Hall.

Schwarz, J. C., Barton-Henry, M. L., & Pruzinsky, T. (1985). Assessing child rearing behaviors: A comparison of ratings made by mother, father, child and sibling on the CRPBI. *Child Development, 56,* 462–479.

Sessa, F. M., Avenevoli, S., Steinberg, L., & Morris, A. S. (2001). Correspondence among informants on parenting: Preschool children, mothers, and observers. *Journal of Family Psychology, 15,* 53–68.

Shaffer, D., Lucas, C. P., & Richters, J. E. (1999). *Diagnostic assessment in child and adolescent psychopathology.* New York: Guilford.

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

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York: Harper Collins.

Verhulst, F. C., Althaus, M., & Berden, M. G. (1987). The child assessment schedule: Parent-child agreement and validity measures. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 28,* 455–466.

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

Welsh, D. P., Galliher, R. V., & Powers, S. I. (1998). Divergent realities and perceived inequalities: Adolescents', mothers', and observers' perceptions of family interactions and adolescent psychological functioning. *Journal of Adolescent Research, 13,* 377–402.

Youngstrom, E., Izard, C., & Ackerman, B. (1999). Dysphoria-related bias in maternal ratings of children. *Journal of Consulting and Clinical Psychology, 67,* 905–916.

Youngstrom, E., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology, 68,* 1038–1050.