Developmental Psychology
1998, Vol. 34, No. 2, 276–287


0012-1649/98/$3.00

# Teacher Discipline and Child Misbehavior in Day Care: Untangling Causality With Correlational Data

David Harvey Arnold and Lorette McWilliams
University of Massachusetts at Amherst

Elizabeth Harvey Arnold
University of Connecticut

Day-care centers provide an ideal, underused setting for studying the developmental processes of child psychopathology. The influence of day-care teachers' lax and overreactive discipline on children's behavior problems was examined, as was the influence of children's behavior problems on teachers' discipline. Participants were 145 children and 16 day-care teachers from 8 classrooms in a day-care center for children from low-income families. Two techniques are presented for estimating causal relations based on correlational data gathered from day-care centers: 2-stage least squares and simultaneous structural equation modeling. Across techniques, teachers' laxness strongly influenced child misbehavior, and child misbehavior influenced both teachers' overreactivity and laxness. Teachers' overreactivity did not influence child misbehavior.

Disruptive behavior problems are characterized by high rates of noncompliance, aggression, and disruptive behavior and affect approximately 10% of grade school children (Kazdin, 1987). These problems predict later academic failure, substance abuse, violence, crime, and psychiatric disorders (Caspi, Elder, & Bem, 1987; Farrington, 1983; Loeber, 1990). Behavior problems typically emerge and become stable at an early age, and intervention has been more successful with younger children than with older ones (Dishion & Patterson, 1992; Kazdin, 1987, 1993). Consequently, leaders in this field have called for prevention programs to address these problems earlier (Hinshaw, 1992; Kazdin, 1987; Loeber, 1990).

Day-care centers are a potential vehicle for studying and addressing behavior problems in their early stages. The term "day care" is used broadly to apply to all centers that provide care to children prior to formal grade school, including programs that would typically be called "preschool." Enrollment in day care is rapidly increasing; by 1994, 61% of all 3- to 5-year-olds were enrolled in such programs, compared to 27% in 1965 (National Center for Educational Statistics, 1995). In some cases, day-care teachers may spend more time with children than do children's parents or siblings. Studies of adult–child interactions in day care can allow naturalistic study without genetic confounds, often with near random assignment of children to classrooms. Further, day-care centers provide a means for reaching children from low-income families who are at especially high risk for developing disruptive behavior disorders (Hawkins, Catalano, & Miller, 1992; Kolvin, Miller, Fleeting, & Kolvin, 1988; Offord, Alder, & Boyle, 1986; Paternite, Loney, & Langhorne, 1976; Szatmari, Boyle, & Offord, 1989). The quality of children's day-care experiences, measured broadly, appears important to children's development (Howes & Olenick, 1986; McCartney, 1984; Phillips, McCartney, & Scarr, 1987; Russell, 1990). However, very little is known about the specific processes or variables that are important (Clarke-Stewart, Allhusen, & Clements, 1995). Discipline by day-care teachers is one variable that needs to be better understood because of its potential impact on disruptive behavior.

In contrast to substantial research on parental discipline and children's behavior problems, virtually no research has examined how day-care teachers' discipline affects such problems. A handful of studies has examined the effects of teachers' reactions to behavior problems in grade school children. For example, the importance of teachers' reprimand length (Abramowitz, O'Leary, & Futtersak, 1988), timing (Abramowitz & O'Leary, 1990), firmness (Van Houten, Nau, MacKenzie-Keating, Sameoto, & Colavecchia, 1982), consistency (Acker & O'Leary, 1988), and use of consequences (Rosen, O'Leary, Joyce, Conway, & Pfiffner, 1984) has been demonstrated with older children. With respect to preschool-age children, however, only four studies even touch on the relation between discipline approaches and behavior problems, all involving small samples. Brown and Elliot (1965) observed decreased aggression after having teachers in one classroom ignore aggression and reinforce positive behaviors. Atwater and Morris (1988) examined whether teachers phrased requests to 27 children as imperatives, declaratives, or questions and found that the request form was not related to compliance. Sherburne, Utley, McConnell, and Gannon (1988) compared a modified time-out to verbal prompts in response to aggression in one classroom and found less aggression with time-out. Finally, Swiezy, Matson, and Box (1992) found increased compliance when they rewarded compliance

David Harvey Arnold and Lorette McWilliams, Department of Psychology, University of Massachusetts at Amherst; Elizabeth Harvey Arnold, Department of Psychology, University of Connecticut.

We are grateful for the assistance of the teachers and students who participated in this study. We would also like to acknowledge Jessica Griffith for her contributions and thank David Kenny and Arnold Well for their statistical help.

Correspondence concerning this article should be addressed to David Harvey Arnold, Department of Psychology, Tobin Hall, University of Massachusetts, Box 37710, Amherst, Massachusetts 01003-7710. Electronic mail may be sent to darnold@psych.umass.edu.

and ignored noncompliance in four preschoolers. These studies are the entire body of knowledge about day-care teachers' discipline and children's misbehavior. There is some basis for expecting that day-care teachers might greatly influence behavior problems, although it might also be argued that this relationship is transitory and unimportant. This question needs to be addressed empirically; until day-care interactions and their effects become a focus of study, such alternatives cannot be evaluated.

There are certainly some important dissimilarities between parent–child and teacher–child relationships, with differences, for example, in the group size and the length and frequency of interactions. Nonetheless, research on parental discipline still may provide some guidance in beginning to understand discipline and misbehavior in day care. First, parental research suggests that dysfunctional responses to early behavior problems are influential, malleable risk factors for the further development of these problems. The importance of two dimensions of discipline has been especially well-demonstrated. Excessively lax and overreactive discipline, by various names, have been associated with behavior problems in both younger and older children (e.g., Arnold, O'Leary, Wolff, & Acker, 1993; Forehand, Wells, & Griest, 1980; O'Leary, 1995; Patterson, DeBaryshe, & Ramsey, 1988; Patterson, Chamberlain, & Reid, 1982). *Laxness* refers to allowing rules to go unenforced, giving in to children's coercive behavior, and coaxing or begging children to behave. *Overreactivity* refers to displays of anger or irritation in response to misbehavior. Overreactivity and laxness are common targets in parent training (e.g., Webster-Stratton, Kolpacoff, & Hollinsworth, 1988). Although not synonymous, laxness to some extent parallels the permissive parenting style described by Baumrind as "avoiding the exercise of control" (Baumrind, 1966, p. 889), and overreactivity may share some features with the authoritarian style of favoring "punitive and forceful measures" (Baumrind, 1966, p. 890). For example, Baumrind presented evidence that such punitive, controlling parenting is associated with physical punishment, frustration, and threats (Baumrind, 1966, pp. 893–894).

Theory suggests a bidirectional relation between parental discipline and child misbehavior. Patterson (1982) postulated that a coercive cycle occurs in which children's aversive behavior causes parents to respond with overly harsh or overly lax parenting to terminate aversive child behavior. These parenting responses, in turn, contribute to greater misbehavior in the long-term through negative reinforcement and modeling, creating a vicious cycle in which parents and children reinforce dysfunctional behavior in one another. Support for parent effects comes from treatment studies in which children of parents taught to provide clear, firm, calm, appropriate, consistent discipline exhibited decreased rates of misbehavior (Kazdin, 1987). Laboratory studies have also demonstrated causal effects of parental discipline. For example, Pfiffner and O'Leary (1989) found that short, firm, immediate reprimands decreased misbehavior.

Understanding of child effects is more limited, but three classic studies provide some general support for them. Anderson, Lytton, and Romney (1986) had mothers interact with both a conduct-disordered and a typical boy; conduct-disordered boys elicited more negative statements. Brunk and Henggeler (1984) enlisted confederate children to play conduct-problem and anxious-withdrawn roles. They found that conduct-disordered boys elicited more commands and discipline from adults and were ignored more often. Finally, Barkley and Cunningham (1979) examined mothers interacting with their children with attention deficit hyperactivity disorder (ADHD) on stimulants and on placebos. Mothers were less directive and negative when their children took medication.

The degree to which the understanding of parent discipline generalizes to day care is unclear. Teachers might be more hesitant to administer negative consequences to students than parents would be with their own children. Alston (1982) demonstrated that day-care teachers have different concerns than parents do regarding misbehavior. Compared to grade school, day care is typically more active and less structured, which might elicit more behavior problems. Compared to home care, the high child-to-adult ratio and the possibility of contagious, modeled misbehavior could also predispose these settings to more behavior problems. Given the frequency with which children misbehave even when directly supervised by parents (e.g., Power & Chapieski, 1986), it would not be surprising if less densely supervised day-care students misbehaved more frequently. Although good normative data do not exist, day-care teachers describe disruptive behavior as their biggest challenge (Micklo, 1992). Given these factors and the potential impact of day care on future behavior, day-care teachers' discipline would seem important to study.

Understanding causal relations between discipline and misbehavior would add to theoretical knowledge in this area and have practical implications as well. On the theoretical front, such knowledge would clarify the mechanisms and contexts that are associated with coercive cycles, further understanding of child effects, and clarify the influence that teachers' discipline has on problem behaviors. Experimental studies have tended to be somewhat artificial, whereas naturalistic studies produce causally ambiguous results. Thus, a method that could supplement these approaches would add to knowledge of these relations. On the practical front, disentangling causal relations would provide guidance in designing programs to reduce problem behaviors in the classroom. Bell and Harper (1977) reviewed a wide range of approaches to disentangling child from adult effects and found no particular approach sufficient for conclusively answering such questions. Rather, convergent evidence across methods is likely to provide the best estimates of causal effects. The use of instrumental variables might provide an important piece of this convergence.

It has long been recognized in other fields of study that the identification of instrumental variables can provide for the estimation of causal effects in reciprocal relations (e.g., Baseman, 1957). Instrumental variables are exogenous variables that directly influence a variable of interest but have no direct effect on a correlated variable. In diagram form, the reciprocal relation between Variables Y1 and Y2 in Figure 1 could not be estimated if they were the only variables measured. However, the measurement of instruments, Variables X1 and X2, allow for the causal influences of Y1 and Y2 to be separated.

Day-care centers may provide such instruments because children can be observed with multiple teachers and teachers can be observed with multiple children. For example, as shown in

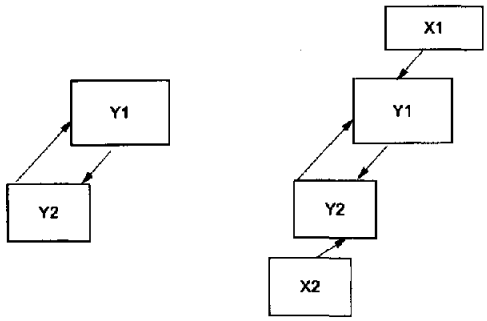


*Figure 1.* The classic causation problem. In the left portion of the figure, the extent that Y1 and Y2 cause each other cannot be determined. In the right model, Instruments X1 and X2 allow the causal effects of Y1 and Y2 to be estimated.

Figure 2, it might be assumed that a teacher's discipline of a specific child (*specific laxness* and *specific overreactivity*) is a function of the teacher's general discipline characteristics (which might also be termed a style or trait; *general laxness* and *general overreactivity*) and the specific child's behavior with the teacher (*specific misbehavior*). If a teacher's general discipline style can be measured independently of his or her interactions with that specific child, then teacher general discipline may be an instrumental, exogenous variable. Similarly, a child's misbehavior with a specific teacher (*specific misbehavior*) could be conceptualized as a function of his or her general behavioral characteristics (*general misbehavior*) and the teacher's discipline of that child (*specific laxness* and *specific overreactivity*). Again, if that child's general misbehavior can be measured independently of her or his interactions with that specific teacher, then child general misbehavior may be another instrument. One plausible way for researchers to measure a teacher's general discipline and a child's general misbehavior independently of their behavior with one another would be to measure the teacher's general discipline characteristics using his or her discipline of all other children besides that specific child and to measure the child's general misbehavior using his or her misbehavior with other teachers.

The approach of researchers' using instruments to make this estimation requires two assumptions. First, it must be assumed that a child's misbehavior with a specific teacher does not directly affect and is not directly affected by the teacher's discipline of other children. Second, it must be assumed that one teacher's discipline of a specific child does not directly affect and is not directly affected by that child's misbehavior with other teachers. These assumptions are open to question; these variables may have some effects on one another. However, if such effects exist, it is expected that they would be small and that these assumptions are generally plausible. At the conceptual level, it is expected that these teachers have established general discipline styles across their years of teaching and that their general discipline styles are relatively stable. Therefore, the misbehavior of a single child is likely to affect the teacher's discipline with that child and not the teacher's general discipline as measured by his or her discipline across all other children in the classroom. In addition, learning theory suggests that immediate, proximal consequences have the strongest influence on behaviors, suggesting that a teacher's discipline of a child would have far greater effects on that child than would her or his discipline of other children and that a child's misbehavior with a teacher would have a stronger influence on that teacher's discipline than would any knowledge the teacher might have about the child's misbehavior with other teachers. For example, if a child misbehaves a great deal with one teacher but behaves very well with another teacher, it seems unlikely that the child's good behavior with the second teacher would have a substantial impact on how overreactive or lax the first teacher is with that child. Partial tests of these assumptions can be conducted to evaluate these theoretical arguments. To the extent that the assumptions are accurate, the causal relations between discipline and misbehavior can be estimated with two-stage least squares (2SLS) and simultaneous structural equation modeling.

Two-stage least squares is powerful, conceptually straightforward, and computationally simple. Its desirable statistical properties and practical use have been well-demonstrated, particularly in economics. An EconLit search found over 300 articles using 2SLS in the last 20 years. James and Singh (1978) presented a nontechnical, readable review of 2SLS for psychology, calling for its use across a wide range of psychological research. Despite their compelling presentation, a PsychLit search revealed only 8 articles in all of psychology that have used this approach since that time, with none in developmental psychology. Simultaneous structural equation modeling (SSEM) also allows for the utilization of instruments, and is now computationally manageable, with the emergence of user-friendly versions of LISREL, EQS, and the like.

The present study examines the relation between teachers' lax and overreactive discipline and children's misbehavior in day care. Laxness and overreactivity were chosen because of their consistent relations with behavior problems in the parental literature. It was hypothesized that as with parents, lax and overreactive teacher discipline both would be related to child misbehavior. In addition, the study estimated the causal effects of teacher laxness on child misbehavior and of child misbehavior on teacher laxness and the effects of overreactivity on misbe-

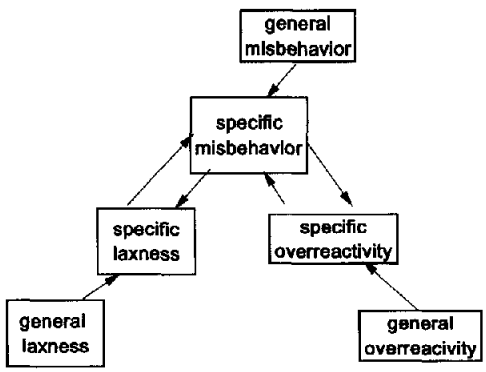


*Figure 2.* A model of how instrumental variables might separate the effects of discipline on misbehavior and the effects of misbehavior on discipline.

havior and of misbehavior on overreactivity. It was posited that both teacher and child effects would be present in each case. Models were constructed with exogenous instruments, and these models were estimated with both 2SLS and SSEM.

## Method

### Participants

Participants in the present study were 145 children (74 boys and 71 girls) and 16 day-care teachers (14 women and 2 men) from eight classrooms (2 teachers per classroom) in a day-care center that provides care to children from low-income families who are eligible for government subsidies. Such children are at a very high risk of developing disruptive behavior problems and are understudied. Children ranged in age from 35 to 74 months of age, with an average age of 55 months ($SD$ = 10.6 months). One-hundred and six (73.1%) of these children were African American, 34 (23.4%) were Latino, and 5 (3.4%) were Anglo American. Ten (63.5%) of the teachers were African American, 4 (25%) were Latino, and 2 were Anglo American (12.5%).

### Procedure

Each classroom was videotaped as part of a larger study. Two videotape segments, each focusing on a different teacher, were chosen from each classroom as follows. Segments of group activities were identified in which one teacher was responsible for the entire class. From these segments, a 15-min sample was randomly selected for each of two teachers in each of the eight classrooms. Fifteen-min samples were chosen for two reasons. First, previous studies and pilot data indicated that this was sufficient time to obtain stable estimates of these constructs; misbehavior occurs at high rates, which means that teachers engage in many discipline encounters in 15 min. Second, use of this relatively short period of time allowed for a much larger sample size; any cost to reliability was thought to be more than offset by the greater power achieved with the sample size, which is greater than twice the average sample used in observational studies of parenting and externalizing problems (Rothbaum & Weisz, 1994).

Fifty-five of the 145 children interacted with only one teacher because they were absent or off camera during the filming of the other teacher; thus, there was a 15-min sample of dyadic behavior for each of 235 teacher–child dyads (290 potential dyads minus the 55 missing segments). The other 90 children interacted with both teachers from their class, for a possible 180 dyads. However, 25 of these 180 child–teacher dyads allowed no ratings of laxness or overreactivity because there were no instances of misbehavior or teacher directives (see *Coding Definitions* section), leaving 155 dyads with complete data across two teachers. These 155 dyads were used in estimating causal effects, whereas all segments were used in estimating teachers' general discipline characteristics (see *Variable Formation* section).

A trained undergraduate research assistant who was unaware of the purpose of this study rated each of the videotape segments, as described below. Each child–teacher dyad was observed and coded individually, and a second coder rated one third of the dyads, randomly selected to allow for estimates of interrater reliability, assessed with intraclass coefficients, which are appropriate for such data (Bartko, 1966).

### Variable Formation

Specific laxness, specific overreactivity, and specific misbehavior were measured with the ratings and tallies described below for each teacher–child dyad. Teachers' general overreactivity and general laxness were measured by averaging their discipline ratings for all children except the child of interest. Children's general misbehavior was measured by their misbehavior with the other teacher, standardized within each videotape segment to create a measure of each child's misbehavior relative to his or her peers in the same situation with the same teacher (teachers' general discipline could not be standardized within classrooms because there were only two teachers per classroom).

### Coding Definitions

*Specific misbehavior.* Counts were made of instances of each child's misbehavior. *Misbehavior* was defined as aggressive, hostile, or noncompliant acts (e.g., hitting, pushing, verbal aggression, grabbing a toy, and ignoring direct teacher requests). An intraclass correlation was conducted to assess the reliability of the coding; the resulting reliability coefficient was .76.

*Specific laxness.* *Laxness* refers to a teacher's not enforcing rules, not following through on requests or directives, and coaxing or begging a child to behave rather than using firm, clear directives. Ratings were made on the basis of how frequently the teacher exhibited lax behavior with a specific child versus how frequently he or she firmly handled that child's misbehavior. The severity of lax behavior displayed by the teacher was also considered in making the rating. For example, a teacher who ignored an aggressive child would be rated as more lax than a teacher who waited a bit before getting a child to clean up. Global ratings were made on the basis of these observations, with 7 indicating *high* levels of laxness and 1 indicating *little or no* laxness. The intraclass correlation coefficient indicated that the interrater reliability was .77 for laxness.

*Specific overreactivity.* *Overreactivity* refers to responding to misbehavior with anger, irritation, frustration, or annoyance rather than being calm and businesslike. Global ratings were made on the basis of these observations, with 7 indicating *high* levels of overreactivity and 1 indicating *no* overreactivity. The intraclass correlation coefficient for overreactivity was .90.

For both laxness and overreactivity, coders took into account instances when teachers responded to misbehavior without overreactivity or laxness as well as times in which overreactivity or laxness were noted. That is, a teacher who responded with overreactivity to 2 of 20 misbehaviors would receive a much lower score than would a teacher who was overreactive on 2 of 3 occasions.

### General Analytic Procedures

*2SLS.* Two-stage least squares is one means of estimating causality with instruments. Suppose one is trying to examine the effect of Y1 on Y2 (see Figure 1). In addition to the usual assumptions of regression analyses, it must be assumed that X1 is not a direct cause of Y2 and is not caused by either Y1 or Y2. In the first stage of 2SLS, ordinary regression is used to predict the causal variable (Y1) from the exogenous instruments (X1 and X2). This regression equation is termed a *reduced-form equation.* If the assumptions above are met, the predicted values of Y1 provide proxy estimates of Y1 that have been purged of the influence of Y2 and allow for an unbiased estimate of its causal effects. In the second stage, a regression equation is estimated in which this proxy variable predicts the dependent variable (Y2). This second stage provides an estimate of the causal effect of the variable of interest.

*SSEM.* Simultaneous structural equation modeling provides another method of using exogenous, instrumental variables to estimate causal effects, allowing the model depicted in Figure 2 to be estimated. As opposed to the separate steps used in 2SLS, SSEM simultaneously estimates a system of equations specified by a model using maximum likelihood estimation procedures.

### Comparison of 2SLS and SSEM

Both 2SLS and SSEM can use exogenous instrumental variables to decompose reciprocal relations. Simultaneous structural equation model-

ing provides more flexibility than 2SLS by allowing correlated error terms. Further, because SSEM simultaneously considers all information, it provides more efficient parameter estimates than does 2SLS. However, SSEM becomes unbiased only with large samples and is perhaps more sensitive to violations of assumptions than is 2SLS. In essence, because SSEM simultaneously estimates a web of relations, errors in one strand of the web may have unknown effects throughout the system. Schmitt and Bedeian (1982) have thus argued that 2SLS is preferable. Given the efficiency, flexibility, and popularity of SSEM, however, it remains a viable alternative for using instrumental variables.

Each approach requires the assumption of a static model. That is, it must be assumed that the causal relations between the variables remain relatively stable across time (Kenny, 1996). Both approaches allow the advantage of addressing an error in independent variables. In traditional regression, an error in independent variables biases results. Two-stage least squares avoids this bias with its two-stage procedure, and SSEM can explicitly model error. Thus, both techniques can provide estimates unbiased by error in independent variables.

### *Specific Analyses*

Under both approaches, each teacher–child dyad was entered as a separate case. Dyads were chosen as the unit of analysis because that was the interaction of interest.

For the 2SLS analyses, the first stage involved estimating reduced-form equations predicting the endogenous variables from the instruments. Teachers' specific laxness and overreactivity and children's specific misbehavior were each predicted from all three exogenous instruments: general laxness, overreactivity, and misbehavior. The effects of teachers' discipline on child misbehavior were tested in the second stage, which involved entering the predicted discipline values together with children's general misbehavior in a regression equation to predict children's specific misbehavior. This equation allows for a causal estimate of teachers' specific discipline on children's specific misbehavior. The effects of child's misbehavior on teacher's discipline were tested in the second stage, as two regression equations were formed in which the predicted child's misbehavior values and teachers' general overreactivity and laxness scores were used to predict teachers' specific overreactivity and laxness. Because these procedures were conducted with a standard statistical computer package, the standard errors of this two-stage process needed to be and were corrected (for a description of the correction procedure, see Erlanger & Winsborough, 1976; Hout, 1977).

For the SSEM analyses, the Structural Equations Modeling module (SEPATH) of Statistica (StatSoft, 1995) was used to estimate the model presented in Figure 2. Error terms for teacher and child ratings were allowed to be correlated, as were laxness and overreactivity error terms, in case the coder's ratings were biased across dyads. Multiple indicators would allow for the constructs of laxness and overreactivity to be represented as latent variables, which may make more conceptual sense than representing these variables as directly measured. Multiple indicators of these variables could have been artificially created by, for example, considering each teacher's interactions with each child as a measure of general laxness or dividing observations into shorter time periods. However, doing so would have required a larger sample and would tend to take advantage of sample-specific characteristics, so the direct use of the data with its accompanying limitations was preferred.

Under both approaches, unstandardized coefficients are presented. Although the scales used to measure discipline are somewhat arbitrary, they nonetheless have more substantive meaning than do standardized coefficients, and this same scale can and has been used across studies. Furthermore, unstandardized coefficients from various methods or samples can be compared, whereas standardized coefficients cannot be compared (King, 1986; Namboodiri, Carter, & Blalock, 1975).

Analyses were first conducted for the entire sample and then separately on the basis of gender, ethnicity, and age. There were no theoretically guided predictions of how the relationship between discipline and misbehavior might differ across these demographic variables. However, these variables have been shown to be related to discipline and misbehavior (e.g., Condry & Ross, 1985; Cornbleth & Korth, 1980; Fagot, 1984), suggesting the need to explore possible differences. Unfortunately, the numbers of Latino and Anglo American children were not sufficiently large to provide stable estimates with 2SLS and SSEM. Therefore, only simple analyses were examined separately by ethnicity.

## Results

### *Descriptive Information*

Table 1 presents the means and standard deviations for the central variables of the study, both for the entire sample and by gender, ethnicity, and age. On average, approximately 49 instances of misbehavior were observed per class in each 15-min observation period. Although it is difficult to place this rate of

Table 1
*Means and Standard Deviations for the Central Variables*

| | Misbehavior | | Laxness | | Overreactivity | |
|---|---|---|---|---|---|---|
| Category and group | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| All children | 5.0 | 3.9 | 3.7 | 2.0 | 1.5 | 1.0 |
| Gender | | | | | | |
| Girls | 3.3 | 2.6 | 3.2 | 2.0 | 1.7 | 1.1 |
| Boys | 6.5 | 4.1 | 4.0 | 1.9 | 1.4 | 0.9 |
| Ethnicity | | | | | | |
| African American | 4.9 | 3.8 | 3.7 | 2.0 | 1.4 | 0.9 |
| Latino | 5.6 | 4.1 | 3.8 | 2.0 | 1.9 | 1.4 |
| Age[a] | | | | | | |
| Younger | 5.2 | 4.3 | 3.3 | 1.8 | 1.2 | 0.6 |
| Older | 4.7 | 3.6 | 3.8 | 2.1 | 1.6 | 1.1 |

*Note.* Misbehavior refers to the number of instances per child. Laxness and overreactivity are measured on a 7-point scale, with higher scores indicating more laxness and overreactivity.
[a] Older is defined as 55 months or greater, younger as less than 55 months.

misbehavior in context because of a lack of normative data in this area, it seems consistent with day-care teachers' reports that disruptive behavior is their biggest challenge in the classroom (Micklo, 1992) and is consistent with these children's high risk status. The overreactivity ratings were positively skewed, with a large number of 1 ratings. This distribution may have attenuated the relation of overreactivity with other variables. Boys misbehaved significantly more frequently than did girls ($p < .001$). Teachers were more lax toward boys than toward girls ($p < .01$), were more overreactive toward Latino than toward African American children ($p < .05$), and more overreactive toward older than toward younger children ($p < .05$).

### *Simple Relations Between Discipline and Misbehavior*

The correlations between the variables measured in this study are presented in Table 2 by gender, ethnicity, and age. Comparisons between subgroups must be made cautiously given the difficulties in comparing correlation coefficients with the smaller number of participants in subgroups. Nonetheless, examination of these relationships may generate hypotheses that would help explain the divergent trajectories of these groups. In contrast to the differences in means described above, the relationships between variables followed fairly similar patterns across the groups. One possible point of divergence is that overreactivity was significantly correlated with misbehavior in boys but not in girls, and in African American but not in Latino children. For the entire group, as predicted, specific laxness and specific overreactivity correlated significantly with specific child misbehavior. Teachers' specific laxness and specific overreactivity were, in contrast to the parental literature, not significantly correlated. General laxness and general overreactivity were slightly negatively correlated.

### *Initial Analyses*

For the 2SLS analyses, the reduced-form equations are presented in Table 3. With respect to the SSEM analyses, the parameter estimates of the impact of the general characteristics on the specific variables were as follows for the large group: general

Table 2
*Simple Correlations Among Specific and General Misbehavior, Laxness, and Overreactivity*

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **All children** | | | | | | |
| 1. Specific misbehavior | — | | | | | |
| 2. General misbehavior | .34*** | — | | | | |
| 3. Specific laxness | .67*** | .13 | — | | | |
| 4. General laxness | .41*** | −.03 | .68*** | — | | |
| 5. Specific overreactivity | .25** | .27** | −.03 | −.09 | — | |
| 6. General overreactivity | −.09 | −.07 | −.14 | −.18* | .39*** | — |
| **Girls and boys[a]** | | | | | | |
| 1. Specific misbehavior | — | .17 | .62*** | .39** | .04 | −.08 |
| 2. General misbehavior | .25* | — | .08 | −.04 | .21 | .05 |
| 3. Specific laxness | .68*** | .04 | — | .69*** | −.20 | −.15 |
| 4. General laxness | .44*** | −.07 | .68*** | — | −.29* | −.29* |
| 5. Specific overreactivity | .28** | .25* | .01 | .00 | — | .46*** |
| 6. General overreactivity | −.10 | −.14 | −.14 | −.10 | .37*** | — |
| **African American and Latino children[b]** | | | | | | |
| 1. Specific misbehavior | — | .33*** | .65*** | .43*** | .32*** | −.01 |
| 2. General misbehavior | .28 | — | .11 | −.02 | .30** | −.09 |
| 3. Specific laxness | .71*** | .14 | — | .73*** | .04 | −.05 |
| 4. General laxness | .32 | −.13 | .49** | — | −.05 | −.16 |
| 5. Specific overreactivity | .04 | .15 | −.26 | −.20 | — | .22* |
| 6. General overreactivity | −.42* | −.14 | −.51** | −.30 | .65*** | — |
| **Older and younger children[c]** | | | | | | |
| 1. Specific misbehavior | — | .40*** | .71*** | .44** | .24* | −.27* |
| 2. General misbehavior | .30* | — | .10 | −.09 | .31** | −.08 |
| 3. Specific laxness | .64*** | .18 | — | .69*** | −.03 | −.31** |
| 4. General laxness | .38** | .04 | .66*** | — | −.22 | −.40*** |
| 5. Specific overreactivity | .29* | .22 | −.01 | .12 | — | .46*** |
| 6. General overreactivity | .08 | −.06 | .06 | .11 | .29* | — |

*Note.* Values are unstandardized coefficients. Older = 55 months or greater; younger = 55 months.
[a] Girls above diagonal; boys below diagonal. [b] African American children above diagonal; Latino children below diagonal. [c] Older children about diagonal; younger children below diagonal.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3
*Reduced-Form Regression Equations From the First Stage of the Two-Stage Least Squares Analyses*

| | Independent variable | | | | | | |
|---|---|---|---|---|---|---|---|
| | General laxness | | General overreactivity | | General misbehavior | | |
| Dependent variable | $\beta$ | *SE* | $\beta$ | *SE* | $\beta$ | *SE* | $R^2$ |
| **All children** | | | | | | | |
| Specific laxness | .89 | .08*** | −.05 | .24 | .33 | .12** | .49*** |
| Specific overreactivity | −.01 | .05 | .84 | .15*** | .32 | .08*** | .24*** |
| Specific misbehavior | .70 | .12*** | .05 | .35 | .93 | .18*** | .29*** |
| **Boys** | | | | | | | |
| Specific laxness | .84 | .10*** | −.24 | .30 | .15 | .15 | .48*** |
| Specific overreactivity | .05 | .07 | .94 | .22*** | .35 | .11** | .23*** |
| Specific misbehavior | .79 | .16*** | −.09 | .51 | .76 | .25** | .28*** |
| **Girls** | | | | | | | |
| Specific laxness | .96 | .13*** | .19 | .37 | .31 | .26 | .49*** |
| Specific overreactivity | −.10 | .07 | .68 | .19*** | .23 | .14 | .27*** |
| Specific misbehavior | .50 | .15** | .09 | .43 | .47 | .30 | .19** |
| **Younger children** | | | | | | | |
| Specific laxness | .86 | .12*** | .08 | .33 | .33 | .19** | .41*** |
| Specific overreactivity | .02 | .07 | .77 | .18*** | .25 | .10* | .23*** |
| Specific misbehavior | .69 | .21** | .29 | .54 | .81 | .32* | .20*** |
| **Older children** | | | | | | | |
| Specific laxness | .90 | .11*** | −.21 | .37 | .32 | .16 | .56*** |
| Specific overreactivity | −.01 | .08 | .93 | .26*** | .39 | .11** | .26*** |
| Specific misbehavior | .65 | .13*** | −.32 | .44 | 1.0 | .19*** | .46*** |

* $p < .05$. ** $p < .01$. *** $p < .001$.

misbehavior on specific misbehavior, .67; general laxness on specific laxness, .59; and general overreactivity on specific overreactivity, .86. All of these coefficients were significant ($p < .001$).

### *Causal Effects of Discipline*

Parameter estimates of teacher effects are presented for both methods in Table 4. As predicted, under both estimation approaches, laxness had strong effects on child misbehavior. With 2SLS, each point increase in laxness was associated with an increase of .79 misbehaviors, $t(151) = 4.9$, $p < .001$. With SSEM, each point increase in laxness was associated with an increase of .78 misbehaviors ($p < .001$). The entire SSEM model is presented in Figure 3, by gender in Figure 4, and by age in Figure 5. Contrary to prediction, overreactivity was not estimated to causally affect misbehavior for either 2SLS or SSEM.

No clear differences were observed between subgroups of children. The one exception is that it appears that the effect of laxness on misbehavior might be stronger for boys than for girls. The lack of a significant effect of overreactivity on misbehavior was observed for both younger and older children. The direction of the effect was opposite for older and younger children; for older children, overreactivity was associated with less misbehavior, and for younger children, overreactivity was associated with more misbehavior. However, because neither of these effects are significant and this pattern was not observed in the simple correlations, this pattern is likely a chance occurrence.

### *Causal Effects of Misbehavior*

Parameter estimates of child effects are presented in Table 4. As predicted, under both approaches, child misbehavior was estimated to significantly affect teacher laxness. With 2SLS, it was estimated that each additional misbehavior caused a .35-point increase in teacher laxness, $t(151) = 2.2$, $p = .03$, whereas under SSEM, it was estimated that each additional misbehavior was associated with a .40-point increase in teacher laxness ($p < .001$). Support for child effects on overreactivity was also found. Two-stage least squares estimates suggest that each additional misbehavior led to a .35-point increase in teacher overreactivity, $t(151) = 5.0$, $p < .001$, whereas the SSEM parameter estimate was .15 ($p < .01$).

Table 4
*A Comparison of Methods of Estimating the Effects of Teacher Discipline on Child Misbehavior and the Effects of Child Misbehavior on Teacher Discipline for All Children, by Gender, and by Age*

| Method of estimation | Teacher effects | | | | Child effects | | | |
|---|---|---|---|---|---|---|---|---|
| | Lax on child | | Over on child | | Child on lax | | Child on over | |
| | $\beta$ | *SE* | $\beta$ | *SE* | $\beta$ | *SE* | $\beta$ | *SE* |
| **All children** | | | | | | | | |
| 2SLS | .79 | .16*** | .10 | .80 | 0.35 | .16* | .35 | .07*** |
| SSEM | .78 | .11*** | .03 | .28 | 0.40 | .10*** | .15 | .05** |
| **Boys** | | | | | | | | |
| 2SLS | .93 | .28*** | −.16 | .78 | 0.20 | .28 | .38 | .09*** |
| SSEM | .95 | .16*** | −.03 | .41 | 0.29 | .15* | .17 | .08* |
| **Girls** | | | | | | | | |
| 2SLS | .56 | .24* | .14 | .71 | 0.33 | .39 | .25 | .13 |
| SSEM | .55 | .12*** | .07 | .12 | 1.27 | .13*** | −.14 | .15 |
| **Younger children** | | | | | | | | |
| 2SLS | .78 | .35* | .39 | .97 | 0.36 | .30 | .27 | .08** |
| SSEM | .76 | .22*** | .24 | .47 | 0.48 | .17** | .15 | .08 |
| **Older children** | | | | | | | | |
| 2SLS | .73 | .18*** | −.34 | .64 | .34 | .26 | .41 | .14** |
| SSEM | .77 | .11*** | −.19 | .32 | .33 | .13* | .19 | .08* |

*Note.* Values are unstandardized coefficients. Older = 55 months or greater; younger = less than 55 months. Lax = laxness; over = overreactivity. 2SLS = two-stage least squares; SSEM = simultaneous structural equation modeling.
* $p < .05$. ** $p < .01$. *** $p < .001$.

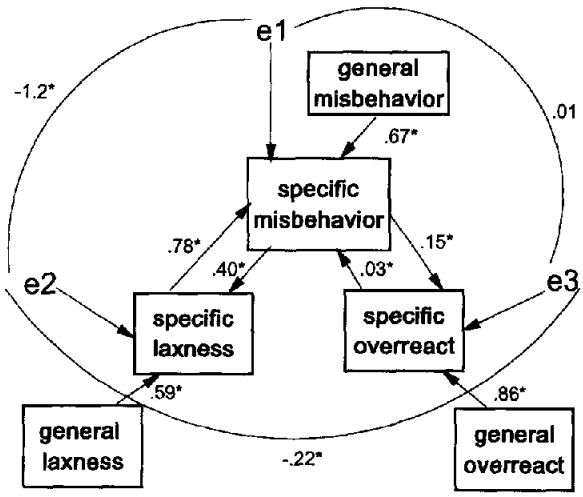


*Figure 3.* The model of reciprocal teacher and child effects estimated with simultaneous structural equation modeling. Error terms are shown only in the cases in which they were allowed to correlate with each other. *$p < .05$.

Child effects were similar across different age groups. The effect of child misbehavior on overreactivity was significant for boys but not for girls. This gender difference could be caused by having a much greater range of misbehavior in boys than in girls, which would result in greater power. Examination of the regression coefficients for 2SLS supports this hypothesis; the magnitude of the effect for girls is only somewhat smaller than the effect for boys. However, the magnitudes of the coefficients from SSEM appear to be quite different for boys than for girls. Simultaneous structural equation modeling analyses are especially sensitive to small sample sizes, and so these differences should be interpreted tentatively.

### *Tests of Assumption Consistency*

As presented by James and Singh (1978), statistical tests can be performed that provide some evidence regarding the appropriateness of the assumptions made regarding the instruments' relations with other variables. Specifically, one endogenous variable (e.g., Y1) can be predicted from the other with the second-stage parameter estimate, and the residuals can be calculated. These residuals are then considered the dependent variable, with the instruments being the independent variables. If the appropriate instruments are essentially unrelated to these

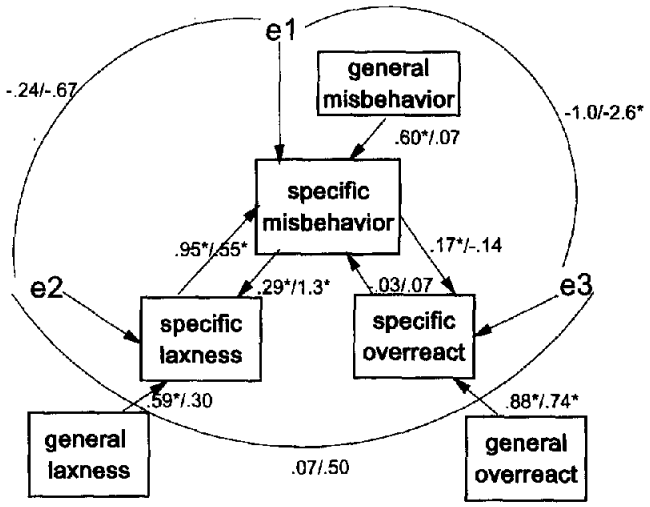


*Figure 4.* The model of reciprocal teacher and child effects estimated with simultaneous structural equation modeling by gender. Coefficients for boys are presented before the slash marks. Coefficients for girls are presented after the slash marks. Error terms are shown only in the cases in which they were allowed to correlate with each other. $*p < .05$.

residuals, then support for the assumptions is provided. For example, specific laxness was estimated from specific misbehavior with the parameter estimate obtained in the second stage. The residuals from this estimation were then entered into a regression equation, with the instruments as the independent variables. If the regression weight for general misbehavior is close to zero, support is provided for the assumption that general misbehavior has no direct effect on specific laxness. As presented in Table 5, the consistency tests indicated that all such coefficients were near zero, consistent with the assumptions made. It should be noted that these tests cannot conclusively establish the necessary assumptions but can provide some evidence for their reasonableness.

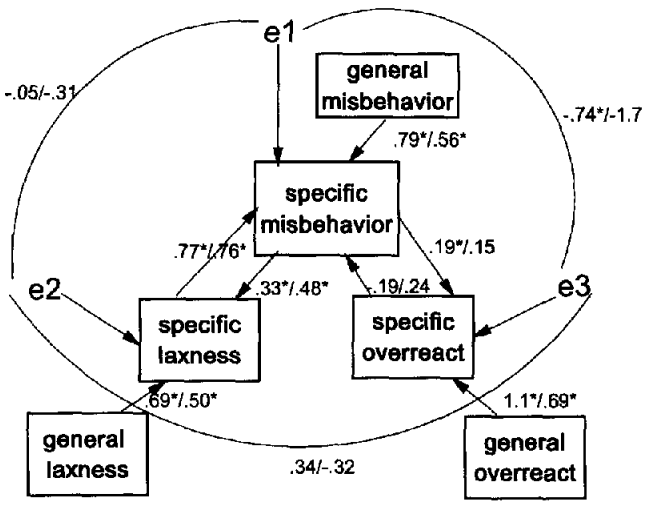


*Figure 5.* The model of reciprocal teacher and child effects estimated with simultaneous structural equation modeling for older and younger children. Older is defined as 55 months or greater; younger as less than 55 months. Coefficients for older children are presented before the slash marks. Coefficients for younger children are presented after the slash marks. Error terms are shown only in the cases in which they were allowed to correlate with each other. $*p < .05$.

## Discussion

The bidirectional effects of teachers' discipline and children's misbehavior were examined with 2SLS and SSEM. The results from the two estimation methods converged in suggesting that teachers' laxness strongly influenced child misbehavior and that child misbehavior influenced both teacher overreactivity and laxness. Teachers' overreactivity did not appear to affect children's misbehavior. These results highlight the importance of considering teachers in developmental models.

The influence of lax discipline on misbehavior is consistent with theory and with empirical studies of classrooms of older children and parents with preschool-age children (e.g., Pfiffner & O'Leary, 1989; Van Houten et al., 1982). The causal importance of laxness in day-care settings suggests that teachers who do not set and enforce clear, firm, consistent, and appropriate classroom rules are likely to face higher levels of misbehavior, which may trigger coercive cycles. Thus, the results suggest that helping teachers learn to set and enforce such rules may be critical to programs aimed at preventing disruptive behavior disorders.

In this study, teacher overreactivity was not found to causally influence misbehavior. It is possible that the skewed overreactivity ratings obtained in this sample might have obscured an effect. In many dyads, no overreactivity was exhibited, perhaps because of a general social prohibition against teachers' overreactivity or, perhaps, less overreactivity than usual was exhibited because teachers were being videotaped. However, child effects on overreactivity suggest enough variability in overreactivity to detect effects, and so other possibilities should be further evaluated. It may be generally true that overreactivity is a reaction to rather than a cause of misbehavior. Although modeling theory would suggest overreactivity effects, the empirical literature provides little evidence that overreactivity has causal

Table 5
*Tests of Logical Consistency for Residuals of Variables Not Expected To Be Directly Related to Specific Child Misbehavior and Teacher Laxness and Overreactivity*

| Dependent variable | Independent variable | Parameter value | $p$ |
|---|---|---|---|
| Specific misbehavior | General laxness | .003 | .98 |
| | General overreactivity | −.001 | .99 |
| Specific laxness | General misbehavior | −.002 | .99 |
| Specific overreactivity | General misbehavior | .003 | .98 |

*Note.* Values represent the relationship between the independent variables and the dependent variables, with the appropriate direct and indirect effects controlled (see James & Singh, 1978). If the assumptions of two-stage least squares are met, these values will approximately equal zero.

effects, and it may not have them. For example, Maccoby (1980) described children of authoritarian parents as being quiet and obedient (p. 385).

Alternatively, overreactive discipline may have stronger causal effects for parents than for teachers. Perhaps the intense, long-term parent–child relationship causes overreactive parenting to be modeled or to provide a child's only means of attention in some cases. This effect might be buffered in day care by the group context, in which modeling and attention from a large group of peers and other teachers might diffuse effects. Or perhaps the escalation of the coercive cycle that overreactivity is thought to cause in children might not be found in day care; for example, a child might be inhibited from escalating because of the public nature of the setting, because of differences in the teacher–child relationship, or because of pressure from other children or teachers. Thus, coercive cycles found in parent–child relationships might not occur or might operate differently in day care. Some support for the notion that coercive cycles may be different in the day care is found in the lack of correlation between overreactivity and laxness and the relatively weak relation between misbehavior and overreactivity.

Finally, the construct of overreactivity may need to be redefined for non-Anglo American cultures. For example, recent evidence suggests that, in contrast to Anglo American families, the use of moderate physical punishment was not related to behavior problems in African American families (Deater-Deckhard, Dodge, Bates, & Pettit, 1996), suggesting that cultural diversity in discipline approaches needs to be better understood. These various possibilities point to the importance of future studies aimed at better understanding coercive cycles in day care. Of course, even if overreactivity in day-care teachers does not have causal effects on externalizing problems, it might well affect internalizing or other problems.

With respect to child effects, support was found for the commonly held but little tested idea that children affect teachers. Specifically, children's misbehavior affected both laxness and overreactivity. These results support including child effects in models of development. Research is needed to examine the process by which misbehavior affects day-care teachers to help teachers minimize such effects. For example, child behavior may affect a caregiver's mood, attributions, or expectations that appear to affect behavior toward children (e.g., Jouriles, Murphy, & O'Leary, 1989; Smith & O'Leary, 1995).

The preschool period is marked by a dramatic gender divergence in both the forms and the amounts of aggression (Crick & Grotpeter, 1995; Eagly & Steffan, 1986; Zahn-Waxler, 1993). This divergence might be better understood, in part, by examinations of different socialization and discipline practices. Consistent with prior studies, boys exhibited more misbehavior than did girls (for a review, see Eagly & Steffan, 1986), and teachers reacted differently to their problem behaviors (e.g., Fagot, 1984; Fagot & Hagan, 1985). Consistent with coercion theory, this difference in misbehavior was paralleled by teachers' greater laxness toward boys. Because laxness was coded to adjust for the amount of misbehavior exhibited, this represents greater laxness *per misbehavior.* Furthermore, the causal analyses suggest that laxness may have a greater effect on misbehavior in boys than in girls. That is, they may be more likely to "get away with what they can." These findings in combination suggest a possible pattern in which coercive cycles might escalate faster in boys than in girls. This might, in part, explain the divergent trajectories of boys and girls, in which boys are more likely to develop externalizing problems. These hypotheses should be evaluated in future studies.

Surprisingly, teachers were not more overreactive toward boys than toward girls. This might be accounted for by teachers' differential expectations of boys and girls. Prior studies have found that similar behaviors are interpreted as more aggressive in girls than in boys (e.g., Condry & Ross, 1985), and so teachers might be more likely to interpret girls' misbehavior as inappropriate and react with overreactivity. Such hypotheses need further evaluation.

Potential threats to the validity of the present study should be noted. The fact that the same coder coded both teacher discipline and child misbehavior could have potentially biased and inflated results. However, evidence against bias includes the following: First, the error terms of these variables were allowed to correlate in the SSEM, and the results were very similar to 2SLS. Second, the ratings across coders were reliable. Although identical biases are possible, the good reliability estimates make bias less likely. Third, the tallies of child misbehavior were based on specific, operationalized definitions that should minimize potential biases. In addition, even if separate coders had been used, their ratings might still have been influenced by the presence of the other dyad member. This issue might eventually be solved with technologies that allow for only the teacher or the child to be observed, but, for now, this issue should be noted as a potential caution in interpreting the results. As mentioned above, other possible biasing factors could be that teachers influenced children's behavior with other teachers or that children's behavior with other teachers influenced a teacher. If assumptions about the absence of such effects were incorrect, they may have biased the results of this study. Although evaluation of the assumptions provided some support for them, future studies should further examine them directly. Alternatively, measuring children's and teachers' general characteristics 1 year and their specific interactions the next year would help ensure the assumptions. The relatively short length of each dyadic observation may also threaten the validity of these results. However, the convergent validity of the variables of interest suggest that the short time period was sufficient to obtain reliable and valid indicators of the variables.

Finally, relations may be different in different groups. In particular, relations among these variables may differ as a function of the age and the gender of children. Future studies should examine larger groups of boys and girls separately in narrow age groups. Similarly, studies should examine different types of day care, and children and teachers from different socioeconomic status and ethnic groups.

If the results of the present study are supported by future research, the implications for day-care centers should be considered. For example, it might be that laxness generally stems from inappropriately low teacher–student ratios. Or perhaps teachers must increase the priority they give to teaching and to managing social behaviors. Or perhaps the training that day-care providers

receive should emphasize strategies for consistently enforcing classroom rules.

Continued study will be critical to understanding the relation between discipline and misbehavior in day care. Future studies might also examine the extent to which teacher interactions with children influence their development in other contexts. In addition, future studies might expand the techniques of the present study to other questions in such areas as peer relationships or, perhaps, families with multiple children. In addition, the similarities and differences in teacher and parent discipline and their effects would be interesting to examine. The methods used in this article are unique to this area of study, and more widespread use of these methods will provide convergent evidence with respect to their use. In addition to increased use, Monte Carlo simulation studies of these methods might provide useful information about their robustness to violations of assumptions. Finally, direct empirical study of the underlying assumptions would help evaluate the methods as well. Ultimately, a comparison of results across a range of approaches, including true experiments and microanalyses of interaction patterns, will provide the most convincing understanding. Such studies should be conducted across a wide range of situations, types of centers, child ages, and ethnic and socioeconomic status groups. Such studies should consider functional classifications of behaviors, focusing on quality of responses to children rather than on simple counts of heterogeneous behaviors.

## References


Abramowitz, A. J., & O'Leary, S. G. (1990). Effectiveness of delayed punishment in an applied setting. *Behavior Therapy, 21,* 231–239.

Abramowitz, A. J., O'Leary, S. G., & Futtersak, M. W. (1988). The relative impact of long and short reprimands on children's off-task behavior in the classroom. *Behavior Therapy, 19,* 243–247.

Acker, M. M., & O'Leary, S. G. (1988). Effects of consistent and inconsistent feedback on inappropriate child behavior. *Behavior Therapy, 19,* 619–624.

Alston, L. (1982). Parent versus teacher reactions to problematic behavior of Head Start children. *Journal of Applied Developmental Psychology, 3,* 155–165.

Anderson, K. E., Lytton, H., & Romney, D. M. (1986). Mothers' interactions with normal and conduct-disordered boys: Who affects whom? *Developmental Psychology, 22,* 604–609.

Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment, 5,* 137–144.

Atwater, J. B., & Morris, E. K. (1988). Teachers' instruction and children's compliance in preschool classrooms: A descriptive analysis. *Journal of Applied Behavior Analysis, 21,* 157–167.

Barkley, R. A., & Cunningham, C. E. (1979). The effects of methylphenidate on the mother–child interactions of hyperactive children. *Archives of General Psychiatry, 36,* 201–208.

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports, 19,* 3–11.

Baseman, R. L. (1957). A generalized classical model of linear estimation of coefficients in a structural equation. *Econometrica, 25,* 77–83.

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Development, 37,* 887–907.

Bell, R. Q., & Harper, L. V. (1977). *Child effects on adults.* Hillsdale, NJ: Erlbaum.

Brown, P., & Elliot, R. (1965). Control of aggression in a nursery school class. *Journal of Experimental Child Psychology, 2,* 103–107.

Brunk, M. A., & Henggeler, S. W. (1984). Child influences on adult controls: An experimental investigation. *Developmental Psychology, 20,* 1074–1081.

Caspi, A., Elder, G. H., & Bem, D. J. (1987). Moving against the world: Life course patterns of explosive children. *Developmental Psychology, 23,* 308–313.

Clarke-Stewart, K. A., Allhusen, V. D., & Clements, D. C. (1995). Nonparental caregiving. In M. H. Bornstein (Ed.) *Handbook of parenting: Volume 3* (pp. 151–176). Mahwah, NJ: Erlbaum.

Condry, J. C., & Ross, D. F. (1985). Sex and aggression: The influence of gender label on the perception of aggression in children. *Child Development, 56,* 225–233.

Cornbleth, C., & Korth, W. (1980). Teacher perceptions and teacher–student interaction in integrated classrooms. *Journal of Experimental Education, 48,* 259–263.

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development, 66,* 710–722.

Deater-Deckhard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1996). Physical discipline among African American and European American mothers: Links to children's externalizing behaviors. *Developmental Psychology, 32,* 1065–1072.

Dishion, T. J., & Patterson, G. R. (1992). Age effects in parent training outcome. *Behavior Therapy, 23,* 719–729.

Eagly, A. H., & Steffan, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin, 100,* 309–330.

Erlanger, H. S., & Winsborough, H. H. (1976). The subculture of violence thesis: An example of a simultaneous equation model in sociology. *Sociological Methods and Research, 5,* 231–246.

Fagot, B. I. (1984). The consequents of problem behavior in toddler children. *Journal of Abnormal Child Psychology, 12,* 385–396.

Fagot, B. I., & Hagan, R. (1985). Aggression in toddlers: Responses to the assertive acts of boys and girls. *Sex Roles, 12,* 341–351.

Farrington, D. P. (1983). Offending from 10 to 25 years of age. In K. T. Van Deusen & S. A. Mednick (Eds.), *Prospective studies of crime and delinquency* (pp. 17–38). Boston: Kluwer-Nijhoff.

Forehand, R., Wells, K., & Griest, D. (1980). An examination of the social validity of a parent-training program. *Behavior Therapy, 11,* 121–123.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112,* 64–105.

Hinshaw, S. P. (1992). Academic underachievement, attention deficits, and aggression: Comorbidity and implications for intervention. *Journal of Consulting and Clinical Psychology, 60,* 893–903.

Hout, M. (1977). A cautionary note on the use of two-stage least squares. *Sociological Methods and Research, 5,* 335–346.

Howes, C., & Olenick, M. (1986). Family and child care influences on toddler compliance. *Child Development, 57,* 202–216.

James, L. R., & Singh, B. K. (1978). An introduction to the logic, assumptions, and basic analytic procedures of two-stage least squares. *Psychological Bulletin, 85,* 1104–1122.

Jouriles, E. H., Murphy, C. M., & O'Leary, K. D. (1989). Effects of maternal mood on mother–son interaction patterns. *Journal of Abnormal Child Psychology, 17,* 513–525.

Kazdin, A. E. (1987). *Conduct disorder in childhood and adolescence.* Newbury Park, CA: Sage.

Kazdin, A. E. (1993). Treatment of conduct disorder: Progress and di-

rections in psychotherapy research. *Development and Psychopathology, 5,* 277–310.

Kenny, D. A. (1996). The design and analysis of social-interaction research. *Annual Review of Psychology, 47,* 59–86.

King, G. (1986). How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science, 30,* 666–687.

Kolvin, F. J., Miller, J. W., Fleeting, M., & Kolvin, P. A. (1988). Social and parenting factors affecting criminal-offence rates: Findings from the Newcastle thousand family study (1947–1980). *British Journal of Psychiatry, 152,* 80–90.

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review, 10,* 1–41.

Maccoby, E. E. (1980). *Social development: Psychological growth and the parent–child relationship.* New York: Harcourt Brace Jovanovich.

McCartney, K. (1984). Effect of quality of day care environment of children's language development. *Developmental Psychology, 20,* 244–260.

Micklo, S. J. (1992). *Perceived problems of public school prekindergarten teachers in Florida related to demographic variables.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Namboodiri, N. K., Carter, L. R., & Blalock, H. M. (1975). *Applied multivariate analyses and experimental designs.* New York: McGraw-Hill.

National Center for Educational Statistics (1995). *Digest of education statistics.* Washington, DC: U.S. Government Printing Office.

Offord, D. R., Alder, R. J., & Boyle, M. B. (1986). Prevalence and sociodemographic correlates of conduct disorder. *American Journal of Social Psychiatry, 6,* 272–278.

O'Leary, S. G. (1995). Parental discipline mistakes. *Current Directions in Psychological Science, 4,* 11–13.

Paternite, C. E., Loney, J., & Langhorne, J. E. (1976). Relationships between symptomatology and SES-related factors in hyperkinetic/mbd boys. *American Journal of Orthopsychiatry, 46,* 291–301.

Patterson, G. R. (1982). *Coercive family process.* Eugene, OR: Castalia Publishing.

Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of a parent training program. *Behavior Therapy, 13,* 638–650.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1988). A developmental perspective on antisocial behavior. *American Psychologist, 44,* 329–335.

Pfiffner, L. J., & O'Leary, S. G. (1989). Effects of maternal discipline and nurturance on toddler's behavior and affect. *Journal of Abnormal Child Psychology, 17,* 527–540.

Phillips, D., McCartney, K., & Scarr, S. (1987). Child-care quality and children's social development. *Developmental Psychology, 23,* 537–543.

Power, T. G., & Chapieski, M. L. (1986). Child rearing and impulse control in toddlers: A naturalistic investigation. *Child Development, 22,* 272–275.

Rosen, L. A., O'Leary, S. G., Joyce, S. A., Conway, G., & Pfiffner, L. J. (1984). The importance of prudent negative consequences for maintaining the appropriate behavior of hyperactive students. *Journal of Abnormal Child Psychology, 12,* 581–604.

Rothbaum, F., & Weisz, J. R. (1994). Parental caregiving and child externalizing behavior in nonclinical samples: A meta-analysis. *Psychological Bulletin, 116,* 55–74.

Russell, A. (1990). The effects of child–staff ratio on staff and child behavior in preschools: An experimental study. *Journal of Research in Childhood Education, 4,* 77–88.

Schmitt, N., & Bedeian, A. G. (1982). A comparison of LISREL and two-stage least squares analysis of a hypothesized life–job satisfaction reciprocal relationship. *Journal of Applied Psychology, 67,* 806–817.

Sherburne, S., Utley, B., McConnell, S., & Gannon, J. (1988). Decreasing violent or aggressive theme play among preschool children with behavior disorders. *Exceptional Children, 55,* 166–172.

Smith, A. M., & O'Leary, S. G. (1995). Attributions and arousal as predictors of maternal discipline. *Cognitive Therapy and Research, 19,* 459–471.

StatSoft (1995). Statistics for windows [Computer software]. Tulsa, OK: Author.

Swiezy, N. B., Matson, J. M., & Box, P. (1992). The good behavior game: A token reinforcement system for preschoolers. *Child and Family Behavior Therapy, 14,* 21–32.

Szatmari, P., Boyle, M. H., & Offord, D. R. (1989). Ontario Child Health Study: Prevalence of attention deficit disorder with hyperactivity. *Journal of Child Psychology and Psychiatry, 30,* 219–230.

Van Houten, R. V., Nau, P. A., MacKenzie-Keating, S. E., Sameoto, D., & Colavecchia, B. (1982). An analysis of some variables influencing the effectiveness of reprimands. *Journal of Applied Behavior Analysis, 15,* 65–83.

Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting and Clinical Psychology, 56,* 558–566.

Zahn-Waxler, C. (1993). Warriors and worriers: Gender and psychopathology. *Development and Psychopathology, 5,* 79–89.

Received July 25, 1996
Revision received August 26, 1997
Accepted August 28, 1997 ■