Published in final edited form as:

Child Dev. 2012 November ; 83(6): 1960–1977. doi:10.1111/j.1467-8624.2012.01820.x.

Spanking and Child Development during the First Five Years of Life

Kathryn Maguire-Jack,
University of Wisconsin-Madison

Andrea N. Gromoske, and
University of Wisconsin-Milwaukee

Lawrence M. Berger
University of Wisconsin-Madison

Abstract

The authors used data from the Fragile Families and Child Wellbeing Study to examine whether spanking at ages 1 and 3 is adversely associated with child cognitive skills and behavior problems at ages 3 and 5. Results from cross-lagged path analyses revealed spanking at age 1 to be associated with a higher level of both spanking and externalizing behavior problems at age 3, and spanking at age 3 to be associated with a higher level of both internalizing and externalizing behavior problems at age 5. Additionally, the longer-term associations between spanking at age 1 and behavioral problems at age 5 appeared to predominantly operate through ongoing spanking at age 3. Results for cognitive skills, though less consistent, suggested no association between spanking at age 1 with poorer receptive vocabulary at age 3 or age 5.

Spanking is a common practice among parents of young children in the United States. Approximately one-third of American families report spanking children as young as 10 to 18 months of age (Regalado, Sareen, Inkelas, Wissow, & Halfon, 2004) and about 70% of families report spanking 2- to 4-year-old children (Straus & Stewart, 1999). Whether spanking is an appropriate discipline strategy or is harmful to children remains the subject of ongoing debate, and numerous studies to date have generated somewhat mixed findings. On the whole, however, whereas a minority of empirical studies suggest that spanking is an effective means of behavioral control that does not result in unintended developmental consequences or lasting damage to children (Baumrind, 1996a, 1996b, 1997; Larzelere, 1996, 2000), the preponderance of evidence suggests that it is an ineffective discipline strategy that is correlated with poor developmental outcomes (Berlin et al., 2009; Slade & Wissow, 2004; Taylor, Manganello, Lee, & Rice, 2010). Existing research has also been limited in its ability to fully account for the (most likely) reciprocal nature of the relation between spanking and child development, or to identify the causal direction of this relation. We add to the existing literature on spanking and child development by utilizing cross-lagged path models and data from the Fragile Families and Child Wellbeing Study (FFCW) to examine associations of spanking with child behavior problems and cognitive skills over the first five years of children's lives. We particularly focus on spanking at age 1, which has received less attention than spanking at older ages. As described below, our empirical strategy allows us to better account for the reciprocal nature of the relation between spanking and child outcomes than has been possible in most prior work.

Background

Conceptual framework

In modern industrial societies, parents bear primary responsibility for the socialization of children with regard to “the acceptance of values, standards, and customs of society as well as the ability to function in an adaptive way in the larger societal context” (Grusec & Davidov, 2007, p. 284). More generally, parents play a crucial role in influencing child health, development, and wellbeing and set the stage for children to become well-functioning members of society (Bornstein, 2006; Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000; Maccoby, 2000; Maccoby & Martin, 1983). In addition to the influence of parental behaviors on children, researchers have increasingly paid attention to the extent to which family relationships are reciprocal such that children’s traits, capacities, and behaviors may considerably influence their parents’ childrearing practices (Crouter & Booth, 2003). For example, the transactional effects model of parent-child relations (Sameroff, 1975) posits recurrent bidirectional influences between parent and child. A child’s functioning and behavior may elicit certain parental reactions, which then influence future child functioning and behavior. Through these transactions, children’s and parents’ behaviors enter into a system of reciprocal relations and responses (Lytton, 1990).

Parental discipline strategies are a form of socialization that may affect and be affected by child functioning and behavior. There are many types of discipline, but one of the most heavily researched and hotly debated is corporal or physical punishment. Corporal punishment involves the use of force to cause pain, but not injury, for the purpose of (behavioral) correction or control (Straus & Stewart, 1999). Current evidence suggests that there are reciprocal relations between corporal punishment and child aggression (Berlin et al., 2009; Eron, Huesmann, & Zelli, 1991; Gershoff, 2002; Sheehan & Watson, 2008).

Spanking is perhaps the most commonly used (and least extreme) form of corporal punishment (Straus & Stewart, 1999). Despite its relatively widespread use, however, there is an ongoing debate about whether spanking is an appropriate discipline technique, particularly with regard to very young children, and several bodies of developmental theory imply that it has the potential to adversely affect child development. Social learning theory (Bandura, 1978), for example, suggests that children who are spanked may become more likely to adopt aggressive behaviors because their parents have modeled aggression to them as an acceptable form of behavior. Furthermore, both psychoanalytic theory and attachment theory posit that spanking may lead to increased internalizing behavior problems for children. Psychoanalytic theory suggests that being spanked may engender feelings of hostility on the part of children, but that such feelings may be repressed given fear that, if expressed, such hostility “will lead to punishment, particularly in the form of loss of love or abandonment” (Grusec & Goodnow, 1994, p. 5). Attachment theory (Bretherton, 1985; Bretherton, 1992) proposes that “internal working models” of the parent-child relationship, which are constructed from a child’s actual experience with the parent, are used to develop expectations about self and others: ‘self’ as worthy or unworthy of care and protection and ‘others’ as available or unavailable to provide care and protection. These internal working models are used in appraising and guiding behavior in new situations (Bretherton, 1985). Thus, if a parent spanks a child when he or she is crying and in need of comfort, the child may develop an internal working model of the parent (‘other’) as rejecting and of him or her ‘self’ as unworthy of help, comfort, or love. If such a pattern continues over time, the child may become depressed or anxious, or develop a sense of low self-esteem.

Turning to cognitive skills, recent research on stress and neurobiology has shown that chronic stressors affect children’s brain processes and neurofunctioning (Anda, et al., 2006; Bremner & Vermetten, 2001). For example, studies have found that the cognitive

functioning of maltreated children differs from that of their non-maltreated peers. Although this line of research has focused on child maltreatment, similar processes may suggest that harsh parenting and spanking will result in decreased cognitive functioning. That is, to the extent that spanking causes ongoing stress for children, it may influence brain functioning in the areas of learning and memory, potentially leading to lower cognitive skills.

Finally, the age at which children are spanked may also matter. In particular, children under age 1 may be at greater risk than older children of physiological and emotional damage as a result of spanking. Compared to older children, very young children are cognitively less able to understand why they are being spanked and to alter their behavior to avoid being spanked in the future (Kopp, 1982). Furthermore, since infants largely depend on their caregivers to meet all of their needs, being spanked could inhibit the infant's ability to develop trust and attachment (Slade & Wissow, 2004; Zeanah, Boris, & Larrieu, 1997). Thus, a better understanding of whether spanking prior to age 1 affects children differently than spanking at later ages may yield important implications regarding the range of potential consequences of such discipline strategies at particular developmental stages.

Prior empirical evidence

The vast majority of empirical studies of physical discipline and child development have focused on corporal punishment in general, rather than spanking, in particular, as well as on outcomes associated with physical discipline that occurred after age 2 (Durrant, 2008; Gershoff 2002; Paolucci & Violato, 2004). A notable exception is Berlin and colleagues' (2009) recent study using Early Head Start Research and Evaluation Project (EHS) data to assess the influence of spanking by age 1 on developmental outcomes up to age 3. Although results have been somewhat mixed, such that there continues to be disagreement as to whether corporal punishment is likely to adversely influence children's development (Larzelere, 2000; Slade & Wissow 2004), the preponderance of evidence suggests that spanking and other forms of noninjury causing physical discipline are associated with at least some adverse outcomes for children, most notably in the area of externalizing behavior problems (Durrant, 2008; Gershoff 2002; Paolucci & Violato, 2004).

A multitude of studies have examined links between corporal punishment or spanking and aggressive or other externalizing behavior problems for children. As demonstrated by several recent meta-analyses and comprehensive literature reviews (Durrant, 2008; Gershoff 2002; Paolucci & Violato, 2004), the vast majority find small to moderate associations of spanking and other (minor) forms of physical discipline with these outcomes (see Larzelere [2000] for a review of the handful of studies that have found either positive or nonsignificant associations). Fewer studies have examined the links between corporal punishment or spanking and internalizing behavior problems such as depression, anxiety, and withdrawn behaviors. The studies that have considered internalizing behavior problems, however, tend to find adverse associations of roughly similar size as those for externalizing behaviors (Durrant, 2008; Gershoff, 2002; Paolucci & Violato, 2004). Finally, there is considerably less evidence linking physical discipline and cognitive skills, although, Berlin et al. (2009) found spanking at age 1 to be associated with lower cognitive skills at age 3.

As discussed above, the direction of the relation between spanking and children's outcomes is theoretically ambiguous. Most empirical studies to date have also been unable to disentangle the causal direction of any such associations. Thus, the extent to which spanking causes behavior problems or other adverse outcomes, as opposed to being driven by such problems, is unclear. Indeed, most existing studies have relied on linear regression strategies that cannot account for the reciprocal relationships between spanking and child outcomes and, therefore, offer limited insight into the causal direction of these relations (for exceptions, see, Berlin et al., 2009; Sheehan & Watson, 2008).

We investigate associations between spanking and three child outcomes – externalizing behavior problems, internalizing behavior problems, and cognitive skills – over the first five years of a child’s life utilizing FFCW data and cross-lagged path models. This analytic approach allows us to explicitly account for the reciprocal nature of relations between spanking and child outcomes by simultaneously modeling associations of spanking at ages 1, 3, and 5 with both future spanking and concurrent and future developmental outcomes, as well as associations of child behavior problems and cognitive skills with future spanking and future developmental outcomes. As such, our analyses represent a more stringent examination of the direction of the relation between spanking and child outcomes than has been possible in most prior work, although we caution that, like those of prior work, our estimates do not lend themselves to causal interpretation.

Our research explicitly builds on three recent studies: Berlin et al. (2009), Slade and Wissow (2004), and Taylor et al. (2010). In analyses that are most similar to ours, Berlin et al. (2009) investigated the bidirectional influences of spanking with regard to child externalizing behavior problems and cognitive test scores at ages 1, 2, and 3 using data from the EHS study and cross-lagged path models (Berlin et al., 2009). We extend this work by testing whether the associations between spanking and child outcomes identified at ages 1 and 3 in their work persist at age 5 among an urban birth cohort that shares many similar characteristics to the families in the EHS study (although the populations sampled, study designs, and many of the measures utilized differ across the two studies).

In a second study, Taylor et al. (2010) used FFCW data and logistic regressions to examine the relation between spanking at age 3 and a dichotomous measure of whether child externalizing behavior problems at age 5 were above or below the sample median. However, this study was unable to speak to the full effects of spanking during early childhood because the analysis began at age 3 and was therefore silent with regard to any influence of spanking that occurred prior to age 3. Furthermore, it may be important to estimate the influence of spanking throughout the distribution of behavior problems, rather than simply its influence on whether a child falls above or below the median. We analyze relations between spanking at ages 1, 3, and 5 with continuous measures of child outcomes at ages 3 and 5 in an attempt to more fully elucidate the range of these associations throughout early childhood.

In a third study that is directly related to our work, Slade and Wissow (2004) used probit regressions and data from the National Longitudinal Survey of Youth (NLSY) to analyze associations between spanking at age 2 and two dichotomous measures of behavior problems at age 6 (an indicator of whether a child was rated by his or her mother as being in the top 10% of the national distribution of behavior problems and an indicator of whether the mother had ever been summoned to school to discuss the child’s behavior problems). In addition to examining the link between spanking and behavior problems throughout the behavior problem distribution, rather than the probability of falling in the tail end of the distribution, our work also extends this study by focusing on spanking as early as age 1 and by accounting for continued spanking between ages 1 and 5 (Slade and Wissow did not adjust for ongoing spanking). Considering both whether children are spanked as early as age 1 and also whether they experience ongoing spanking is important given concerns that spanking at such an early age may, in and of itself, have particularly adverse consequences for children (Berlin et al., 2009; Kopp, 1982; Slade & Wissow, 2004). Finally, in addition to examining cognitive skills and externalizing behavior problems, we also examine whether spanking is associated with internalizing behavior problems, which was not considered in any of the aforementioned studies. On the whole, this research contributes to our understanding of the myriad of complex ways in which spanking and child development are interconnected.

Method

Participants

Our data are drawn from FFCW, a longitudinal birth cohort study of 4,898 children born between 1998 and 2000 in 20 U.S. cities. By design, FFCW over-sampled children born to unmarried parents. This process resulted in a larger number of low-income children and Black or Hispanic children than would be the case in a nationally representative sample (see Reichman, Teitler, Garfinkel, & McLanahan, 2001, for a complete description of the sample and study design). Given that sample families are relatively disadvantaged compared to the U.S. population as a whole, FFCW children are likely to also be at greater risk of both punitive parenting practices and adverse child outcomes (Conger, et al., 2002; Day, Peterson, & McCracken, 1999; Hashima & Amato, 1994). Sample families also share many of the same characteristics as those of the EHS families in Berlin and colleagues' (2009) recent study. As such, results from that study and ours are well suited for comparison: consistency or differences in results between the two studies may have implications for whether associations observed in each are likely to reflect the specific analytic sample or may more generally be characteristic of relatively disadvantaged populations in the United States.

FFCW interviewed families in person at the time of the focal child's birth and by telephone when the child was approximately 1, 3, and 5 years old. In each of these interviews, parents provided detailed information about family and household characteristics, resources, structure, and functioning; program participation; physical and mental health; and parenting behaviors. Subsequent to the age 3 and 5 interviews, families were asked to participate in an in-home assessment of multiple domains of parenting, the home environment, mother-child interactions, and child cognitive and emotional/behavioral development through both a questionnaire and a set of interviewer observed items. Mothers who refused an in-home visit were asked to complete the questionnaire portion of the in-home module by telephone. Both the spanking data and the cognitive and behavioral outcomes at ages 3 and 5 were drawn from the in-home modules. Spanking data at age 1 were drawn from the core telephone interview.

We limited our analysis sample to the 3,870 families that had completed at least one in-home interview (either in person or by phone) when the child was either age 3 or 5 in order to ensure that we had at least one non-missing observation point for one or more of the child outcomes. As with all longitudinal studies, FFCW contains missing data at each data collection point. Thus, we imputed all missing data for our full analysis sample using multiple imputation techniques (Allison, 2002; Rubin, 1987; Schafer & Graham, 2002). Multiple imputation assumes that data are missing at random (conditional on observed data) and therefore uses non-missing data to predict values for the missing data. We imputed values for all missing data in our analysis sample using the ICE program in Stata 11 (StataCorp, 2009). Specifically, we imputed 10 data sets, as recommended by Graham and colleagues (2007), using information from all variables included in our analyses.

Measures

Spanking—During each interview, the focal child's mother reported whether (1) she had spanked the focal child within the past month and, if so, how often and (2) whether her partner or spouse (if present in the household) had spanked the focal child within the past month and, if so, how often. Spanking frequency for each of these items was reported on a 5-point scale ranging from 0 to 4 for responses of "never in the past month," "once or twice," "a few times this past month," "a few times a week," or "every day or nearly every day." Our overall spanking frequency measure is based solely on the mother's report of her

own spanking for mothers who did not have a co-resident partner or spouse. For those mothers who had a co-resident partner or spouse, it reflects the most frequent spanking reported. For example, if a child was spanked every day by his or her father (or a social father) but only once or twice in the past month by his or her mother, the child was coded as being spanked every day.

Behavior problems—Child behavior problems were assessed by a subset of items from the externalizing and internalizing behavior problems subscales of the Child Behavior Checklist (CBCL) (Achenbach, 1991; 1992; Achenbach & Rescorla, 2000; McConaughy & Achenbach, 1988). The CBCL is a commonly-used measure that has been shown to be highly reliable (Greenbaum & Dedrick, 1998). It is comprised of a series of behaviors for which a child's primary caregiver (mother in our case) is asked to report whether each is not true, somewhat/sometimes true, or often/very true of the child. Each item is scored on a 0 to 2 point scale such that higher scores indicate more severe behavior problems. Items are then summed to create the various CBCL subscales. FFCW included a subset of the full CBCL. We used the aggressive subscale (15 items at age 3 and 20 items at age 5, with alphas of .86 and .84, respectively) to approximate child externalizing behavior problems and the sum of the anxious/depressed and withdrawn subscales (totaling 24 items at age 3 and 23 items at age 5, with alphas of .82 and .75) to approximate child internalizing behavior problems.

Cognitive skills—Child cognitive skills at ages 3 and 5 were assessed using the Peabody Picture Vocabulary Test- Revised (PPVT-R). The PPVT-R is a receptive vocabulary test that has been widely used to measure children's language and cognitive ability and has a reliability coefficient ranging from .54 to .90 (median=.77) (Dunn & Dunn, 1981). The PPVT-R must be administered in person. The TVIP, a Spanish version of the PPVT-R, was administered to children who spoke primarily Spanish; we include these children in our analyses.

The PPVT-R assesses a different domain of cognitive skills than does the Bayley Mental Development Index (MDI) (Bayley, 1993), the cognitive measure used in Berlin and colleagues' (2009) study. Specifically, the PPVT-R assesses a child's receptive vocabulary whereas the MDI assesses memory, problem solving, verbal communication, and perceptual acuities. Thus our results for cognitive skills will not be directly comparable to those of Berlin et al. (2009); however, an assessment of similarities and differences between our results and theirs may lend some insight into whether any observed associations are robust to diverse measures and domains of child cognitive ability. Additionally, although the PPVT-R is commonly utilized to study child cognitive skills, concerns have been raised as to whether it is fully appropriate for use with minority populations (Zigler, Abelson, & Seitz, 1973; Bracken, 1987; Stockman, 2000). Despite evidence that it exhibits adequate reliability for children across racial and ethnic groups (Halpin, Simpson, & Martin, 1990; Washington & Craig, 1999), we note that this is a potential limitation to our analyses, which are based on a largely black and Hispanic sample.

Child emotionality—As discussed above, a considerable concern in analyses such as ours is how best to account for reciprocal relationships between spanking and child functioning. Because cognitive skills and behavior problems in FFCW were measured at only ages 3 and 5, we cannot account for these factors at age 1. However, the study did measure child emotionality via mother report at age 1. Emotionality relates to the tendency to become aroused easily and intensely (Mathieson & Tambs, 1999). It is likely to be associated with later cognitive skills and behavior problems as well as with maternal spanking. Indeed, empirical evidence suggests it is associated with increased parental stress (Gelfand, Teti, & Fox, 1992) which, in turn, is associated with harsh discipline (Pinderhughes, Dodge, Zelli, Bates, & Pettit, 2000). FFCW assessed child emotionality at age 1 using a subset of 3 of the

5 items ($\alpha = .60$) from the Emotionality, Activity, and Sociability (EAS) Temperament Survey for Children (Mathieson & Tambs, 1999). For each of the items, mothers were asked to report on a 5-point scale whether a given behavior is “not at all like my child” to “very much like my child.” The items were then summed to create a total child emotionality score ranging from 3 to 15.

Covariates—We controlled for a variety of time-invariant and time-varying factors that prior studies have found to be associated with spanking (Belsky, 1984; Berlin et al., 2009; Xu, Tung, & Dunaway, 2000) as well as child cognitive skills and behavior problems (Berlin et al., 2009; Chase-Lansdale, Brooks-Gunn, & Paikoff, 1991). Time-invariant factors included maternal age at the time of the child’s birth, maternal race/ethnicity (Black, non-Hispanic; Hispanic; and another non-Hispanic race or ethnicity; with white, non-Hispanic as the reference group), maternal PPVT-R score at child age 3 (or at age 5 if not assessed at age 3), and child sex. Time-varying factors (measured at child ages 1, 3, and 5) included maternal depression, household income, whether the mother was single, whether the child participated in Early Head Start at age 1 and Head Start at ages 3 or 5, whether the mother worked in the past week, the number of adults living in the household, and the number of children living in the household. We include Early Head Start and Head Start participation as covariates both because involvement in such programs may influence parental spanking behaviors (Love et al., 2005) and children’s developmental outcomes (Forness et al., 1998; Belsky et al., 2007), and also for consistency with Berlin and colleagues’ (2009) analyses using EHS study data.

Analytic Strategy

We estimated a series of cross-lagged path models using a maximum likelihood structural equation modeling (SEM) approach (Kline, 2011). This strategy allowed us to simultaneously estimate the effects of spanking on children’s future cognitive skills and behavior problems and the effects of current cognitive skills and behavior problems on future spanking (Finkel, 1995). A simplified version of our conceptual model is presented in Figure 1. Our goal was to simultaneously estimate: (1) associations of spanking at ages 1, 3, and 5 with cognitive skills or behavior problems at ages 3 and 5; (2) associations of the developmental outcome at age 3 with both spanking and the outcome at age 5; and (3) associations of child emotionality at age 1 with both spanking and cognitive skills or behavior problems at ages 3 and 5. In addition, though omitted from the figure for ease of presentation, we control for the full set of time invariant covariates at age 1 and the full set of time-period relevant time-varying covariates at ages 1, 3, and 5. The coefficients of interest are those shown in the figure: β_1 is the effect of age 1 spanking on the age 5 outcome; β_2 is the effect of age 1 spanking on the age 3 outcome; β_3 is the effect of the age 3 outcome on the age 5 outcome; β_4 is the effect of age 1 spanking on age 3 spanking; and β_5 is the effect of age 3 spanking on the age 5 outcome.

In order to gain insight into the influence of both ongoing spanking (at age 3) and the age 3 developmental outcome on the association between age 1 spanking and the age 5 outcome, for each outcome we engaged in a model building exercise in which we first estimated three simplified versions of the model before estimating the full model depicted in Figure 1. We began our model building with a simple model (Model A) in which we estimated only the effects of age 1 spanking on age 5 spanking and the age 5 outcome, constraining all other paths to be equal to zero. This model allowed us to assess the ‘full’ effect of age 1 spanking on the developmental outcome at age 5 without accounting for the portion of that effect that can be explained by spanking at age 3 or by the influence of age 1 spanking on the age 3 outcome. In Model B, we added the age 3 outcome to assess the extent to which the direct effect of age 1 spanking on the age 5 outcome is explained by the effect of age 1 spanking

on the outcome at age 3, constraining all other paths to be zero. In Model C, we included the age 3 spanking measure, but excluded the age 3 outcome, constraining all other paths to be zero. This model enabled us to assess the extent to which ongoing spanking may explain the association between spanking at age 1 and cognitive skills or behavior problems at age 5. Model D is the full model, which provides simultaneous estimates of all of these direct and indirect paths, as well as within-wave correlations between spanking and the developmental outcome. For this model, we calculated all of the indirect pathways through which spanking at age 1 is associated with the relevant outcome at age 5 in order to gain insight into the potential mediating roles of age 3 spanking and the age 3 developmental outcome. In all of our primary models, the spanking measures are defined in terms of the frequency with which the child was spanked at a given time point. All time invariant covariates (including age 1 emotionality) and all time-period relevant time-varying covariates were included in all models.

We examined several fit statistics for each model in order to determine how well the model fit the data. Specifically, we present the chi-square statistic, for which a p-value of greater than .05 is generally considered to indicate good absolute fit (although this statistic is heavily influenced by sample size), the comparative fit index (CFI) and the Tucker-Lewis Index (TLI), for which values greater than .90 are generally considered to indicate good fit, the root-mean-square error of approximation (RMSEA), for which a value of less than .05 is generally considered to indicate good fit, and the standardized root-mean-square residual (SRMR), for which a value of .08 or less is generally considered to indicate good fit (Kline, 2011). Although there is ongoing debate regarding which of these fit statistics should be preferred and in what particular circumstances, as well as which particular values thereof should be considered to indicate good, adequate, and poor fit (see, e.g., issue 42 of *Personality and Individual Differences*, 2007), we followed the advice of Kaplan (2009) and MacKinnon (2008) and relied most heavily on the RMSEA. Furthermore, we compared the fit of our more parsimonious and more complex models using the Bayesian Information Criteria (BIC) in order to determine if the more complex models fit the data better. A decrease in BIC from a more parsimonious to a more complex model indicates improved model fit (Kline, 2011).

In addition to our primary models, we also engaged in supplemental analyses in which we estimated 4 alternative specifications of Model D (the full model) in order to investigate the robustness of our primary results. First, we estimated a version using dichotomous indicators of whether the child was spanked at each time point in lieu of the spanking frequency measures. Second, we estimated a version in which we examined the timing and consistency with which spanking occurred at ages 1 and 3 through a set of 3 mutually exclusive indicators (1 = yes) of whether the child was spanked at age 1 only, ages 1 and 3, or age 3 only, with the reference category being that the child was not spanked at either age 1 or age 3. Third, we estimated a version in which we coded the spanking variables to indicate the age at which spanking was initiated (onset), comparing children who began being spanked at age 1, age 3, and age 5 to children who were never spanked. Finally, in order to determine whether associations between spanking and child outcomes were driven by extremely frequent or 'excessive' spanking, we re-estimated the model excluding the 183 children (5% of our analysis sample) who had been spanked "every day or nearly every day" at age 1, 3, or 5. All models were estimated using Mplus 6.0 software (Mplus, 2010), which averaged all results over the 10 imputed datasets.

Results

Descriptive Statistics

Table 1 presents descriptive statistics for our full sample and by spanking status at ages 1, 3, and 5. In our sample, spanking was least common at age 1, at which time 30% of children were reported to have been spanked, peaked at 56% at age 3, and decreased slightly to 51% at age 5. In addition, we see that 6% of children were spanked at age 1 but not at age 3, 23% at age 1 and 3, and 33% at age 3 but not at age 1. Turning to the frequency with which children were spanked at age 1 (not shown in the table), 16% of mothers reported that they or their spouse or partner had spanked their child once or twice during the past month, 6% spanked the child a few times in the past month, 6% a few times a week, and about 2% daily. At ages 3 and 5, these figures were 28% and 32% for once or twice in the past month, 15% and 13% for a few times in the past month, 10% and 5% for a few times a week, and about 3% and 1% for daily, respectively.

In addition, the descriptive statistics in Table 1 clearly reveal differences in child behavior problems and child cognitive skills at ages 3 and 5 according to whether or not a child was spanked at ages 1, 3, and 5. On the whole, children who were spanked at age 1 exhibited higher levels of externalizing and internalizing behavior problems at ages 3 and 5 and poorer cognitive skills at age 3 relative to those who were not spanked at age 1. Children who were spanked at age 3 exhibited more externalizing behavior problems at ages 3 and 5 and more internalizing behavior problems at age 5, but also slightly higher cognitive skills at age 5. Finally, children spanked at age 5 exhibited higher levels of externalizing behavior problems at ages 3 and 5 and more internalizing behavior problems at age 5; there were no differences in cognitive skills at either age.

The raw data also revealed considerable differences between children who were and were not spanked in terms of child emotionality (at age 1) as well as a host of background characteristics. Children with higher emotionality scores at age 1 were more likely to be spanked at each subsequent age. Furthermore, children who were spanked tended to have younger mothers who had higher depression scores than those of children who were not spanked. Black children and boys were particularly likely to have been spanked, whereas Hispanic children were less likely to experience spanking. In addition, children who were spanked at age 1 were more likely to live with a single mother at ages 1, 3, and 5, as well as to attend Head Start at ages 3 and 5, than those who were not spanked at age 1. Finally, children who were spanked at age 1 experienced lower household income and more maternal work at ages 1, 3, and 5 compared to those who were not spanked at age 1. Conversely, however, children who were spanked at age 3 experienced slightly higher income than those who were not spanked at age 3; there were no differences in spanking by household income at age 5. We next turn to results from cross-lagged path analyses that adjust for these factors.

Associations between Spanking and Child Behavior Problems

Externalizing behavior problems—Results from our 4 primary path models for externalizing behavior problems are presented in Figures 2 through 5. Figure 2 presents results from the most basic model (Model A). The model fit statistics suggest that this simple model did not fit the data very well. However, when we added age 3 externalizing behavior problems to the model (Model B; Figure 3) the model fit was improved considerably (as indicated by the decrease in the BIC statistic) and the RMSEA indicated adequate model fit. In Model C (Figure 4), we included a measure of age 3 spanking frequency in place of the age 3 externalizing behavior problems measure. Again, this model demonstrated an improvement in model fit relative to Model A, and the RMSEA suggested

adequate model fit (we did not compare Models B and C to one another because they are alternatives to each other; Model C does not build upon Model B). Although we caution that direct interpretation of the coefficients and changes therein from Models A through C should be avoided given that (particularly) Model A fit the data relatively poorly, a general pattern emerged that suggests that there is no direct association between spanking at age 1 and externalizing behavior at age 5 once either the indirect effect of age 1 spanking through age 3 spanking or age 3 externalizing behavior problems is taken into account. As discussed below, this pattern was supported by results from Model D, which fit the data quite well.

Results from our full cross-lagged path model (Model D) are presented in Figure 5. This model exhibits an improvement in model fit compared to both Model B and Model C, and has fit statistics that indicate it fits the data very well. Consistent with our findings from Models B and C, the direct effect of age 1 spanking on age 5 externalizing behavior problems in Model D was nonsignificant. However, each of the indirect paths—that through age 3 spanking and that through age 3 externalizing behavior problems—linking age 1 spanking to age 5 externalizing behavior problems was statistically significant. This finding suggests that the association of age 1 spanking with age 5 externalizing behavior problems reflects both ongoing spanking at age 3 and the more proximate influence of spanking at age 1 on externalizing behavior problems at age 3.

Table 2 presents estimates of the indirect effects. These results confirm that each of the indirect pathways from age 1 spanking to age 5 externalizing behavior problems is significant and reveal that the regression coefficients for the indirect paths from age 1 spanking to age 5 externalizing behavior problems tend to be larger than those from age 1 emotionality to age 5 spanking. This may suggest that spanking has a larger effect on future externalizing behavior problems than externalizing behavior problems have on future spanking. At the same time, results from this model indicate that there are reciprocal (transactional) influences between parent and child behaviors. Furthermore, we see that the estimate of the indirect effect of age 1 spanking on age 5 externalizing behavior problems through age 3 spanking is smaller than the estimate of the indirect effect through age 3 externalizing behavior problems. This may imply that the initial increase in (age 3) externalizing behavior problems that is associated with age 1 spanking is the predominant path through which early spanking influences later externalizing behavior problems.

Turning to our supplemental analyses, we first estimated a variant of Model D in which we included at each time point dichotomous indicators that the child was spanked in lieu of the spanking frequency measures. Results (not shown) were consistent with those of our primary model. Second, we examined patterns with regard to the timing and consistency with which spanking occurred through a set of mutually exclusive indicators of whether the child was spanked at age 1 only, at ages 1 and 3, or at age 3 only. These results (not shown) revealed no association between spanking at age 1 and externalizing behavior problems at either age 3 or age 5. Furthermore, despite that there may have been limited statistical power to detect significant effects of age 1 spanking only, given that only 6% of children were spanked at age 1 but not age 3, it is notable that the coefficients for spanking at age 1 only were extremely small (close to zero) and that the direction of the coefficient with regard to age 5 externalizing behavior problems was negative. This may suggest that early spanking (by age 1), if discontinued thereafter, may not have a long-term influence on externalizing behavior problems, at least at age 5. On the other hand, spanking at both age 1 and 3 was associated with greater externalizing behavior problems at both age 3 and 5. This finding provides further support that observed associations between spanking at age 1 and externalizing behavior problems at age 5 are likely explained by ongoing spanking. Third, we examined patterns with regard to the age at which spanking was initiated (1, 3, or 5) and found (results not shown) that the direct association between spanking initiated at age 1 and

externalizing behavior problems at age 5 was smaller than that between spanking initiated at age 3 and externalizing behavior problems at age 5. Finally, to assess whether adverse associations between spanking and externalizing behavior problems were primarily driven by the minority of children who were exposed to high levels of spanking, we re-estimated Model D excluding those children who were spanked every day or nearly every day (at any age). Overall, these findings (not shown) were consistent with those of our primary models suggesting that our results were not unduly reflective of children who experienced excessive spanking.

Internalizing behavior problems—We repeated Models A through D to estimate associations of spanking with internalizing behavior problems, which have received considerably less attention in prior work (and were not included in the Berlin et al. [2009], Slade and Wissow [2004], or Taylor et al. [2010] studies). Model fit statistics from Model A (not shown) indicated rather poor fit; however, moving from Model A to either Model B or Model C (not shown) improved model fit and led to RMSEAs that indicated adequate fit. The general pattern of results from these models suggests that after accounting for ongoing spanking or ongoing internalizing behavior problems there is no direct association between spanking at age 1 and internalizing behavior problems at age 5. This finding was supported by results of the full cross-lagged path model (Model D; Figure 6), which revealed no significant direct association between age 1 spanking and age 5 internalizing behavior problems when the indirect effects through age 3 spanking and age 3 internalizing behavior problems were included. However, results for the indirect effects (Table 2) for this model are not fully consistent with those for the externalizing behavior problems model: Only 2 of the 6 estimates for the potential indirect pathways were significant. The magnitudes of these estimates suggest that the indirect association of age 1 spanking with age 5 internalizing behavior problems operates through ongoing spanking at age 3 to a greater extent than through increased internalizing behavior problems at age 3.

Results from our supplemental analyses were generally consistent with those from our primary models. The estimation using dichotomous spanking variables revealed no direct association of age 1 or age 3 spanking with age 5 internalizing behavior problems. The timing model indicated that neither being spanked at age 1 only, both age 1 and age 3, nor only at age 3 was associated with internalizing behavior problems at age 5. The spanking onset model indicated that the initiation of spanking at age 1 was not related to internalizing behavior problems at age 3 or age 5, and that spanking initiation at age 3 was not associated with greater internalizing behavior problems at age 5. Finally, when we excluded children exposed to very frequent spanking, we found an association between age 3 spanking and higher age 5 internalizing behavior problems, but no direct association between age 1 spanking and age 5 internalizing behavior problems.

Associations between Spanking and Child Cognitive Skills

Consistent with the prior models, the simplest model (Model A, not shown) for cognitive skills had relatively poor fit. However, the model fit improved considerably and was, overall, quite good with regard to models B, C, and D. Models A and B (not shown) suggested an association between spanking at age 1 and greater cognitive skills at age 5. But this association was nonsignificant in models C (not shown) and D (Figure 7). Results from the full model (Model D; Figure 7) revealed no direct association of age 1 spanking with age 3 or age 5 cognitive skills, and no association between age 3 spanking and age 5 cognitive skills. As was the case with internalizing behavior problems, 5 of the 6 indirect pathways were nonsignificant (Table 2).

Turning to the supplemental analyses, consistent with our primary results, the model using dichotomous spanking measures revealed that age 1 spanking was not directly associated with age 5 cognitive skills; however, the model revealed a significant association between age 3 spanking and higher age 5 cognitive skills. With regard to the timing at which spanking occurred, we found no association of being spanked at only age 1 with age 5 cognitive skills, whereas both being spanked only at age 3 and being spanked at both age 1 and age 3 were associated with higher cognitive skills. The spanking onset model revealed that both spanking initiated at age 1 and spanking initiated at age 3 were associated with higher PPVT-R scores at age 5. Finally, results from the model which excluded children who had been spanked very frequently revealed no direct association between spanking at age 1 and cognitive skills at age 5; spanking at age 3 was not associated with cognitive skills at age 5. Overall, these models are suggestive of tenuous and non-intuitive associations between spanking (particularly at age 3) and higher levels of cognitive skills at age 5. However, such associations were nonsignificant in our preferred primary model (Model D).

Discussion

This study examined associations of spanking with children's externalizing behavior problems, internalizing behavior problems, and cognitive skills over the first five years of life using longitudinal data from FFCW and cross-lagged path models. Our analyses explicitly accounted for the reciprocal nature of relations between spanking and child outcomes by simultaneously modeling associations of spanking at ages 1, 3, and 5 with both future spanking and concurrent and future developmental outcomes, as well as associations of current behavior problems and cognitive skills with future spanking and future developmental outcomes. We hypothesized that spanking at age 1 would be directly associated with higher levels of externalizing and internalizing behavior problems and lower cognitive skills at age 5, even after adjusting for ongoing spanking at age 3 and explicitly modeling the reciprocal relationship between spanking and these outcomes.

Overall, we found that the association between spanking at age 1 and greater externalizing behavior problems at age 5 operated fully through two indirect pathways: spanking at age 3 and behavior problems at age 3. That is, most children who were spanked at age 1 continued to be spanked as they aged, and spanking at age 1 was associated with greater externalizing behavior problems at age 3. In turn, both spanking at age 3 and greater externalizing behavior problems at age 3 were associated with greater externalizing behavior problems at age 5. In addition, spanking at age 1 was indirectly linked to greater internalizing behavior problems at age 5 in that most children who were spanked at age 1 continued to be spanked as they aged (we found no significant association between spanking at age 1 and internalizing behavior problems at age 3). These findings, which were consistent across our primary and supplemental models, fail to support our hypothesis that spanking by age 1 would have direct effects on age 5 behavior problems over and above those of spanking at age 3. That the direct association of age 1 spanking with both externalizing and internalizing behavior problems at age 5 is nonsignificant when the indirect path through age 3 spanking is included in our models may suggest that any influence of early spanking on later behavior problems operates fully through later (ongoing) spanking; it may also reflect that early spanking is confounded with later spanking.

At first glance, our finding that spanking at age 1 is associated with increased externalizing behavior problems at age 3 may seem contrary to those of Berlin and colleagues (2009), who found that spanking at age 1 was related to increased externalizing behavior problems at age 2, but not at age 3. However, we do not have a behavior problems measure at age 2. As such, it is possible that the association we find between age 1 spanking and age 3 externalizing behavior problems is driven by the unmeasured relationship between age 1

spanking and externalizing behavior problems at age 2. Indeed, this would parallel our finding that spanking by age 1 has an influence on our most proximal measure of externalizing behavior problems (age 3), and appears to influence later behavior problems, in part, through this initial association. Berlin et al. (2009) also found a significant association between externalizing behavior problems at age 2 and externalizing behavior problems at age 3, providing some additional support for this hypothesis. Furthermore, our externalizing behavior problems results are consistent with those of Taylor et al. (2010) who, using logistic regressions and FFCW data, found an association between spanking at age 3 and externalizing behavior problems at age 5. Because we utilized cross-lagged path models, however, we are able to better account for the reciprocal relation between spanking and child behavior. That both approaches produce similar results demonstrates that these associations are robust.

Associations between spanking and internalizing behavior problems have received much less attention in prior work. Our results suggest that spanking may have an impact on internalizing behavior problems as children begin to reach middle childhood, but not during early childhood. That is, we find that spanking at age 1 is not associated with age 3 internalizing behavior problems, but that spanking at age 3 is associated with greater internalizing behavior problems at age 5; furthermore, age 1 spanking is associated with age 5 internalizing behavior problems only indirectly through ongoing spanking at age 3. That internalizing behavior problems associated with spanking do not appear until age 5 may reflect the tendency for depressive symptoms, loneliness, and anxiety to manifest during middle childhood (or later) rather than during early childhood or toddlerhood (Zahn-Waxler, Klimes-Dougan, & Slattery, 2000).

Our cognitive skills analyses produced somewhat mixed and inconsistent results. The full cross-lagged path model (Model D) revealed no direct or indirect associations of either age 1 or age 3 spanking with cognitive skills at age 5. However, our supplemental analyses indicated that there may be a significant association between age 3 spanking and greater cognitive skills at age 5. This result is both puzzling and inconsistent with the findings of Berlin and colleagues (2009), who analyzed data on similarly disadvantaged children to those in the FFCW sample. One possibility for divergent findings between our study and theirs may be the use of different measures of cognitive skills. Berlin and colleagues utilized the MDI, whereas we used the PPVT-R. The two tests measure different aspects of children's cognitive development. The PPVT-R measures a child's receptive vocabulary by allowing him or her to choose a picture that is best associated with a verbally-presented vocabulary word (Dunn & Dunn, 1997); the MDI more holistically measures cognitive development by assessing memory, problem solving, verbal communication, and perceptual acuities (Bayley, 1993). It may be that spanking adversely influences aspects of cognitive development other than receptive vocabulary. In addition, Berlin et al. (2009) followed children only until age 3. It is also possible that any initial or early adverse influences of spanking on cognitive skills fades as children age (although we find no evidence of initial effects). In any case, our finding here, though robust to several specifications in our supplemental analyses, is not present in our primary specification, does not lend itself easily to interpretation, and runs counter to what we had hypothesized. As such, we find this result puzzling and cannot speculate as to why early spanking may be associated with higher future cognitive skills, in general, or why the results of our primary and supplemental specifications are contradictory in this regard.

Finally, our findings highlight that there do indeed appear to be reciprocal influences of spanking and child behavior in that we find links between both child emotionality and earlier measures of child behavior problems with later spanking. These reciprocal relations suggest a cyclical pattern of negative parent-child interactions and imply that spanking may

not be an effective way to discourage problematic behaviors. At the same time, the strength of these reciprocal effects differs with regard to internalizing and externalizing behavior problems such that externalizing behavior problems are much more strongly linked to future spanking than are internalizing behavior problems. This makes sense to the extent that the former are more likely to frustrate or anger parents than the latter. In addition, the influence of spanking on behavior problems appears to be stronger than that of behavior problems on spanking, as illustrated by larger standardized coefficients linking spanking to child behavior than child behavior to spanking.

This study has several limitations that should be considered when interpreting our results. First, FFCW did not provide a definition of spanking to mothers. Therefore, mothers in the study answered the questions regarding spanking based on their own conceptualization of spanking, which is likely not uniform across all participants. Second, because FFCW includes a substantial overrepresentation of single-parent and low-income families, our results will be most applicable to these family types and may not generalize to more advantaged families. Third, we rely on maternal reports for both parental spanking and child behavior problems. Mothers' reports of their own and their partners' spanking may be biased such that they underestimate the true level of spanking in the household. In addition, literature on child maltreatment has indicated that mothers who use harsher forms of discipline tend to misperceive their child's behavior and exhibit negative attributional biases (Bradley & Peters, 1991; Chilamkurti & Milner, 1993; Dadds, Mullens, McAllister, & Atkinson, 2003). To the extent that this is also true of mothers who spank their children, it could lead to biased measurement of child behavior and may have thereby influenced our results. Furthermore, because the key spanking and behavior problems measures are all reported by the mother, it is possible that our results may overestimate the true associations among them (Bank, Dishion, Skinner & Patterson, 1990). Fourth, our spanking measure does not differentiate which caregiver is engaging in the spanking behaviors at each time point, or whether the child is being spanked by more than one caregiver. These factors may matter and should be the subject of future work. Fifth, our cognitive measure is limited to receptive vocabulary and does not include the wider range of cognitive skills that have been assessed in other studies. Thus, differences between our results in this area and those of prior work may simply reflect that our measure of cognitive skills is more limited. Finally, as is the case with all observational studies, it is possible that our estimates are biased by omitted variables that are associated with both spanking and the outcomes of interest. Such variables may include omitted background characteristics as well as other aspects of harsh or aggressive parenting or family violence, such as more extreme forms of corporal punishment, intimate partner violence, and parenting stress. Our study is silent in this regard. As such, our estimates do not lend themselves to causal interpretation. Future research should seek to further understand links between spanking, other forms of harsh parenting and family violence, and children's development and wellbeing.

Despite these limitations, this study provides new evidence regarding links between spanking during early childhood and later adverse developmental outcomes for children. Most notably, we find that any direct effects of spanking by age 1 on child outcomes at age 5 are fully explained by ongoing spanking between ages 1 and 5 and/or the initial, proximal influence of spanking on the outcome (at age 3). Our finding that spanking is associated with higher levels of future externalizing and internalizing behavior problems, net of earlier levels of behavior problems, suggests that parents' use of spanking as a discipline strategy may adversely influence children's behavioral development. Our results also suggest that the reciprocal and transactional nature of spanking and (particularly externalizing) behavior problems may potentially result in escalated problem behaviors, rather than increased ongoing behavioral control on the part of the parent. On the whole, then, our behavior problems results suggest that interventions that encourage parents who spank their children

at young ages to discontinue this practice may help to diminish the likelihood that children will develop or continue to exhibit problem behaviors. By providing parents with alternative strategies for disciplining their children, practitioners may have the potential to better help parents control the problem behaviors that are likely to elicit spanking in the first place. In addition to face-to-face interventions in this area, universal efforts such as social marketing campaigns aimed at decreasing the use of spanking may be beneficial (Taylor et al., 2010). Finally, it will be important for future studies to investigate how associations between spanking and child outcomes may vary by the context in which spanking occurs and, in particular, the extent to which spanking is normative or nonnormative in a given context (see, e.g., Gershoff et al., 2010; Simons et al., 2002), as such factors may mediate or moderate the associations revealed by our models.

Acknowledgments

The Fragile Families and Child Wellbeing Study is funded by NICHD grant numbers R01HD36916, R01HD39135, and R01HD40421, as well as a consortium of private foundations and other government agencies. This research was supported by NICHD grant number K01HD054421 (to Berger). We are grateful to David Kaplan, Dan Meyer, Sarah Bruch, Mike Brondino, and Sarah Beal for their excellent advice and assistance.

References

- Achenbach, T. Integrative Guide to the 1991 CBCL/4-18, YSR, and TRF Profiles. Burlington, VT: University of Vermont, Department of Psychology; 1991.
- Achenbach, T. Manual for the Child Behavior Checklist/2-3 and 1992 Profile. Burlington, VT: University of Vermont Department of Psychiatry; 1992.
- Achenbach, T.; Rescorla, L. Manual for the ASEBA Preschool Forms and Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2000.
- Allison, PD. Missing data. Thousand Oaks, CA: Sage; 2002.
- Anda RF, Felitti VJ, Bremner D, Walker JD, Whitfield C, Perry BD, et al. The enduring effects of abuse and related adverse experiences in childhood: A convergence of evidence from neurobiology and epidemiology. *European Archives of Psychiatry and Clinical Neuroscience*. 2006; 256:174–186. [PubMed: 16311898]
- Bandura A. Social learning theory of aggression. *Journal of Communication*. 1978; 28(3):12–29. [PubMed: 690254]
- Bank, L.; Dishion, T.; Skinner, M.; Patterson, GR. Method Variance in Structural Equation Modeling: Living with “Glop.” In: Patterson, GR., editor. *Depression and Aggression in Family Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.; 1990.
- Baumrind D. A blanket injunction against disciplinary use of spanking is not warranted by the data. *Pediatrics*. 1996a; 98(4, Pt. 2):828–831. [PubMed: 8885981]
- Baumrind D. The discipline controversy revisited. *Family Relations*. 1996b; 45:405–415.
- Baumrind D. Necessary distinctions. *Psychological Inquiry*. 1997; 8:176–182.
- Bayley, N. Bayley Scales of Infant Development Second Edition. San Antonio: The Psychological Corporation; 1993.
- Belsky J. The determinants of parenting: A process model. *Child Development*. 1984; 55(1):83–96. [PubMed: 6705636]
- Belsky J, Lowe Vandell D, Burchinal M, Clarke-Stewart KA, McCartney K, Tresch Owen M. Are there long-term effects of early child care? *Child Development*. 2007; 78(2):681–701. [PubMed: 17381797]
- Berlin L, Malone P, Ayoub C, Ispa J, Fine M, Brooks-Gunn J, et al. Correlates and consequences of spanking and verbal punishment for low-income White, African American, and Mexican American toddlers. *Child Development*. 2009; 80(5):1403–1420. [PubMed: 19765008]
- Bornstein, MH. Parenting: Science and practice. In: Sigel, IE.; Renninger, KA., editors. *Handbook of Child Psychology, Vol. 4: Child Psychology and Practice*. New York: Wiley; 2006. p. 892-949.

- Bradley EJ, Peters RD. Physically abusive and nonabusive mothers' perceptions of parenting and child behavior. *American Journal of Orthopsychiatry*. 1991; 61(3):455–460. [PubMed: 1951653]
- Bracken BA. Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*. 1987; 5(4):313–326.
- Bremner JD, Vermetten E. Stress and development: Behavioral and biological consequences. *Development and Psychopathology*. 2001; 13:473–489. [PubMed: 11523844]
- Bretherton I. Attachment theory: Retrospect and prospect. *Monographs of the Society for Research in Child Development*. 1985; 50(1–2):3–35.
- Bretherton I. The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology*. 1992; 28:759–779.
- Chase-Lansdale PL, Brooks-Gunn J, Paikoff RL. Research and programs for adolescent mothers: Missing links and future promises. *Family Relations*. 1991; 40(4):396–403.
- Chilamkurti C, Milner JS. Perceptions and evaluations of child transgressions and disciplinary techniques in high- and low-risk mothers and their children. *Child Development*. 1993; 64:1801–1814. [PubMed: 8112120]
- Collins WA, Maccoby E, Steinberg L, Hetherington EM, Bornstein MH. Contemporary research on parenting: The case for nature and nurture. *American Psychologist*. 2000; 55:218–232. [PubMed: 10717969]
- Conger RD, Wallace LE, Sun Y, Simons RL, McLoyd VC, Brody GH. Economic pressure in African American families: A replication and extension of the family-stress model. *Developmental Psychology*. 2002; 38(2):179–193. [PubMed: 11881755]
- Crouter, AC.; Booth, A. *Children's Influence on Family Dynamics: The Neglected Side of Family Relationships*. Mahwah, NJ: Erlbaum; 2003.
- Dadds MR, Mullins MJ, McAllister RA, Atkinson E. Attributions, affect, and behavior in abuse-risk mothers: A laboratory study. *Child Abuse & Neglect*. 2003; 27:21–45. [PubMed: 12510029]
- Day RD, Peterson GW, McCracken C. Predicting spanking of younger and older children by mothers and fathers. *Journal of Marriage and Family*. 1999; 60(1):79–94.
- Dunn, LM.; Dunn, LM. *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service; 1981.
- Dunn, LM.; Dunn, LM. *Examiner's Manual for the PPVT-III*. Circle Pines, MN: American Guidance Service; 1997.
- Durrant J. Physical punishment, culture, and rights: Current issues for professionals. *Journal of Developmental and Behavioral Pediatrics*. 2008; 29(1):55–66. [PubMed: 18300726]
- Eron, LD.; Huesmann, LR.; Zelli, A. The role of parental variables in the learning of aggression. In: Pepler, NDJ.; Rubin, KH., editors. *The development and treatment of childhood aggression*. Hillsdale, NJ: Erlbaum; 1991. p. 169–188.
- Finkel, SE. *Causal Analysis with Panel Data*. Thousand Oaks, CA: Sage Publications, Inc.; 1995.
- Forness SR, Cluett SE, Ramey CT, Ramey SL, Zima BT, Hsu C, Kavale K, MacMillan DL. Special education identification of Head Start children with emotional and behavioral disorders in second grade. *Journal of Emotional and Behavioral Disorders*. 1998; 65:194–204.
- Gelfand M, Teti D, Fox C. Sources of parenting stress for depressed and nondepressed mothers of infants. *Journal of Clinical Child Psychology*. 1992; 21(3):262–272.
- Gershoff E. Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin*. 2002; 128(4):539–579. [PubMed: 12081081]
- Gershoff ET, Grogan-Kaylor A, Lansford JE, Chang L, Zelli A, Deater-Deckard K, Dodge KA. Parent discipline practices in an international sample: Associations with child behaviors and moderation by perceived normativeness. *Child Development*. 2010; 81:487–502. [PubMed: 20438455]
- Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications on multiple imputation theory. *Prevention Science*. 2007; 8:206–213. [PubMed: 17549635]
- Greenbaum PE, Dedrick RF. Hierarchical confirmatory factor analysis of the Child Behavior Checklist/4–18. *Psychological Assessment*. 1998; 10(2):149–155.

- Grusec, JE.; Davidov, M. Socialization in the family: The role of parents. In: Grusec, JE.; Hastings, PD., editors. *Handbook of Socialization: Theory and Research*. New York: The Guilford Press; 2007. p. 284-308.
- Grusec JE, Goodnow JJ. Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. *Developmental Psychology*. 1994; 30(1):4-19.
- Halpin G, Simpson RG, Martin SL. An investigation of racial bias in the Peabody Picture Vocabulary Test-Revised. *Educational and Psychological Measurement*. 1990; 50(1):183-189.
- Hashima PY, Amato PR. Poverty, social support, and parental behavior. *Child Development*. 1994; 65:394-403. [PubMed: 8013229]
- Kaplan, D. *Structural equation modeling: Foundations and extensions*. Thousand Oaks, CA: Sage Publications, Inc.; 2009.
- Kline, RB. *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press; 2011.
- Kopp CB. Antecedents of self-regulation: A developmental perspective. *Developmental Psychology*. 1982; 18(2):199-214.
- Larzelere RE. A review of the outcomes of parental use of nonabusive or customary physical punishment. *Pediatrics*. 1996; 98(4, Pt. 2):824-828. [PubMed: 8885980]
- Larzelere RE. Child outcomes of non-abusive and customary physical punishment by parents: An updated literature review. *Clinical Child and Family Psychology Review*. 2000; 3(4):199-221. [PubMed: 11225737]
- Love JM, Kisker EE, Ross C, Raikes H, Constantine J, Boller K, et al. The effectiveness of Early Head Start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology*. 2005; 41(6):885-901. [PubMed: 16351335]
- Lytton H. Child and parent effects in boys' conduct disorder: A reinterpretation. *Developmental Psychology*. 1990; 26(5):683-697.
- Maccoby EE. Parenting and its effects on children: On reading and misreading behavior genetics. *Annual Review of Psychology*. 2000; 51:1-27.
- Maccoby, EE.; Martin, JA. Socialization in the context of the family: Parent-child interaction. In: Hetherington, EM., editor. *Handbook of Child Psychology: Vol. 4*. New York: Wiley; 1983. p. 1-101.
- Mathieson KS, Tambs K. The EAS temperament questionnaire – Factor structure, age trends, reliability, and stability in a Norwegian sample. *Journal of Child Psychology and Psychiatry*. 1999; 40:431-439. [PubMed: 10190344]
- McConaughy, S.; Achenbach, T. *Practical Guide for the Child Behavior Checklist and Related Materials*. Burlington: University of VT, Department of Psychiatry; 1988.
- Mplus (Version 6) [Computer software]. Los Angeles, CA: Statmodel;
- Paolucci E, Violato C. A meta-analysis of the published research on the affective, cognitive, and behavioral effects of corporal punishment. *Journal of Psychology*. 2004; 138(3):197-222. [PubMed: 15264439]
- Pinderhughes EE, Dodge KA, Zelli A, Bates JE, Pettit GS. Discipline responses: Influences of parents' socioeconomic status, ethnicity, beliefs about parenting, stress, and cognitive-emotional processes. *Journal of Family Psychology*. 2000; 14(3):380-400. [PubMed: 11025931]
- Regalado M, Sareen H, Inkelas M, Wissow LS, Halfon N. Parents' discipline of young children: Results from the National Survey of Early Childhood Health. *Pediatrics*. 2004; 113:1952-1958. [PubMed: 15173466]
- Reichman NE, Teitler JO, Garfinkel I, McLanahan SS. *Fragile Families: Sample and Design*. Children and Youth Services Review. 2001; 23(4-5):303-326.
- Rubin, DB. *Multiple imputation for nonresponse in surveys*. New York: Wiley; 1987.
- Sameroff AJ. Transactional models in early social relations. *Human Development*. 1975; 18:65-79.
- Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychological Methods*. 2002; 7(2):147-177. [PubMed: 12090408]

- Sheehan MJ, Watson MW. Reciprocal influences between maternal discipline techniques and aggression in children and adolescents. *Aggressive Behavior*. 2008; 34:245–255. [PubMed: 18085532]
- Simons RL, Lin K-H, Gordon LC, Brody GH, Murry V, Conger RD. Community differences in the association between parenting practices and child conduct problems. *Journal of Marriage and Family*. 2002; 64:331–345.
- Slade E, Wissow L. Spanking in early childhood and later behavior problems: A prospective study of infants and toddlers. *Pediatrics*. 2004; 113(5):1321–1330. [PubMed: 15121948]
- Stata (Version 11) [Computer Software]. College Station, TX: StataCorp;
- Straus MA, Stewart JH. Corporal punishment by American parents: National data on prevalence, chronicity, severity, and duration, in relation to child and family characteristics. *Clinical Child and Family Psychology Review*. 1999; 2(2):55–70. [PubMed: 11225932]
- Stockman IJ. The new Peabody Picture Vocabulary Test-III: An illusion of unbiased assessment? *Language, Speech, and Hearing Services in Schools*. 2000; 31:340–353.
- Taylor CA, Manganello JA, Lee SJ, Rice JC. Mother's spanking of 3-year-old children and subsequent risk of children's aggressive behavior. *Pediatrics*. 2010; 125(5):e1057–e1065. [PubMed: 20385647]
- Washington JA, Craig HK. Performances of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, and Hearing Services in Schools*. 1999; 30:75–82.
- Xu X, Tung Y, Dunaway RG. Cultural, human, and social capital as determinants of corporal punishment: Toward an integrated theoretical model. *Journal of Interpersonal Violence*. 2000; 15:603–630.
- Zahn-Waxler C, Klimes-Dougan B, Slattery MJ. Internalizing problems of childhood and adolescence: Prospects, pitfalls, and progress in understanding the development of anxiety and depression. *Development and Psychopathology*. 2000; 12(3):443–466. [PubMed: 11014747]
- Zeanah C, Boris N, Larrieu J. Infant development and development risk: A review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1997; 36(3):165–178. [PubMed: 9031569]
- Zigler E, Abelson WD, Seitz V. Motivational factors in the performance of economically disadvantaged children on the Peabody Picture Vocabulary Test. *Child Development*. 1973; 44(2): 294–303.

Figure 1.
Conceptual model

Fit statistics
 BIC = 101144.35
 $\chi^2 = 1668.35$
 df = 89
 $p = .000$
 RMSEA = .068
 $p = .000$
 CFI = .401
 TLI = .091
 SRMR = .055

Figure 2. Model A standardized coefficients for cross-lagged path analysis of spanking at ages 1 and 5 and externalizing behavior problems at ages 1 and 5. * $p < .05$, ** $p < .01$, *** $p < .001$.

Fit statistics
 BIC = 99686.5
 $\chi^2 = 729.15$
 df = 120
 $p = .000$
 RMSEA = .036
 $p = 1.00$
 CFI = .713
 TLI = .577
 SRMR = .034

Figure 3. Model B standardized coefficients for cross-lagged path analysis of spanking at ages 1 and 5 and externalizing behavior problems at ages 1, 3, and 5. * $p < .05$, ** $p < .01$, *** $p < .001$.

Fit statistics
 BIC = 99675.3
 $\chi^2 = 724.11$
 df = 120
 $p = .000$
 RMSEA = .036
 $p = 1.00$
 CFI = .716
 TLI = .581
 SRMR = .035

Figure 4. Model C standardized coefficients for cross-lagged path analysis of spanking at ages 1, 3, and 5 and externalizing behavior problems at age 5. * $p < .05$, ** $p < .01$, *** $p < .001$.

Fit statistics
 BIC = 98206.2
 $\chi^2 = 96.702$
 df = 108
 $p = .774$
 RMSEA = .00
 $p = 1.00$
 CFI = 1.00
 TLI = 1.00
 SRMR = .014

Figure 5.

Model D standardized coefficients for the full cross-lagged path analysis of spanking at ages 1, 3, and 5 and externalizing behavior problems at ages 3 and 5. * $p < .05$, ** $p < .01$, *** $p < .001$.

Fit statistics
 BIC = 94110.8
 $\chi^2 = 218.35$
 df = 108
 $p = .000$
 RMSEA = .016
 $p = 1.00$
 CFI = .941
 TLI = .904
 SRMR = .019

Figure 6.

Model D standardized coefficients for the full cross-lagged path analysis of spanking at ages 1, 3, and 5 and internalizing behavior problems at ages 3 and 5. * $p < .05$, ** $p < .01$, *** $p < .001$. For comparison purposes, the BIC statistics for models A, B, and C (results not shown) were 96488.0, 95553.9 and 95037.0.

Fit statistics
 BIC = 115087.7
 $\chi^2 = 392.78$
 df = 108
 $p = .000$
 RMSEA = .026
 $p = 1.00$
 CFI = .858
 TLI = .767
 SRMR = .024

Figure 7. Model D standardized coefficients for cross-lagged path analysis of spanking at ages 1, 3, and 5 and cognitive skills at ages 3 and 5. * $p < .05$, ** $p < .01$, *** $p < .001$. For comparison purposes, the BIC statistics for models A, B, and C (results not shown) were 117398.5, 116541.1 and 115942.6

Table 1

Descriptive Statistics

	Age 1				Age 3		Age 5	
	Full sample (N=3,870)	Not spanked (N=2,727)	Spanked (N=1,143)		Not spanked (N=1,705)	Spanked (N=2,165)	Not spanked (N=1,904)	Spanked (N=1,966)
<i>Any spanking</i>								
Age 1	.30	--	1.00		.15	.41***	.19	.40***
Age 3	.56	.46	.78***		--	1.00	.35	.77***
Age 5	.51	.43	.69***		.27	.70***	--	1.00
<i>Spanking timing</i>								
Age 1, not 3	.06	--	.22		.15	--	.08	.05**
Age 3, not 1	.33	.47	--		--	.59	.24	.42***
Age 1 and 3	.23	--	.78		--	.41	.11	.35***
<i>Spanking frequency (0–4 point scale)</i>								
Age 1	.52 (.95)	--	1.75 (.95)		.22 (.63)	.75 (1.09)***	.29 (.70)	.74 (1.10)***
Age 3	.99 (1.1)	.77 (.97)	1.56 (1.19)***		--	1.77 (.90)	.50 (.80)	1.46 (1.16)***
Age 5	.77 (.94)	.61 (.84)	1.14 (1.04)***		.35 (.65)	1.10 (.99)***	--	1.52 (.77)
<i>Externalizing behavior problems</i>								
Age 3	9.87 (5.71)	9.44 (5.57)	10.89 (5.91)***		8.74 (5.51)	10.77 (5.71)***	8.83 (5.47)	10.89 (5.76)***
Age 5	11.91 (6.03)	11.57 (5.97)	12.71 (6.10)***		10.84 (5.84)	12.74 (6.04)***	10.69 (5.76)*	13.08 (6.06)***
<i>Internalizing behavior problems</i>								
Age 3	5.59 (4.03)	5.42 (3.95)	5.97 (4.19)**		5.55 (4.08)	5.61 (4.00)	5.52 (4.02)	5.65 (4.03)
Age 5	5.59 (4.53)	5.44 (4.47)	5.95 (4.66)**		5.40 (4.59)	5.75 (4.48)*	5.23 (4.02)	5.94 (4.59)***
<i>PPVT-R</i>								
Age 3	86.80 (16.98)	87.40 (17.10)	85.38 (16.62)**		87.28 (16.94)	86.43 (16.95)	87.25 (17.34)	86.37 (16.61)
Age 5	93.12(16.27)	93.13 (16.77)	93.10 (15.03)		92.02 (16.94)	93.98 (15.67)**	92.77 (16.88)	93.45 (15.65)
<i>Child emotionality</i>								
Age 1	8.50 (3.18)	8.21 (3.11)	9.20 (3.22)***		8.120 (3.21)	8.74 (3.14)***	8.31 (3.16)	8.69 (3.19)***

Time-invariant covariates (measured at age 1 unless otherwise noted)

	Age 1				Age 3				Age 5			
	Full sample (N=3,870)	Not spanked (N=2,727)	Spanked (N=1,143)		Not spanked (N=1,705)	Spanked (N=2,165)			Not spanked (N=1,904)	Spanked (N=1,966)		
Maternal age at FC birth	25.11 (6.04)	25.62 (6.13)	23.91 (5.66)		25.63 (6.12)	24.70 (5.96)			25.70 (6.28)	24.54 (5.76)		***
Maternal PPVT at age 3	90.15 (13.31)	90.52 (13.75)	89.26 (12.14)	*	89.45 (13.93)	90.70 (12.76)	*		89.64 (13.75)	90.63 (12.84)	*	
Child is female	.48	.49	.44	**	.51	.45	***		.51	.45	***	
Mother is Black	.49	.43	.66	***	.45	.53	***		.44	.55	***	
Mother is Hispanic	.26	.30	.18	***	.32	.22	***		.31	.21	***	
Mother of other race	.03	.04	.02		.03	.03			.04	.03	*	
<i>Time-varying covariates</i>												
<i>Household income</i>												
Age 1	9.68 (1.52)	9.73 (1.52)	9.58 (1.54)	**	9.61 (1.61)	9.74 (1.45)	*		9.69 (1.56)	9.68 (1.49)		
Age 3	9.82 (1.40)	9.89 (1.42)	9.68 (1.33)	***	9.73 (1.54)	9.90 (1.27)	***		9.81 (1.45)	9.84 (1.34)		
Age 5	9.89 (1.46)	9.95 (1.48)	9.75 (1.40)	***	9.83 (1.51)	9.94 (1.41)	*		9.88 (1.48)	9.90 (1.44)		
<i>Single mother</i>												
Age 1	.41	.40	.46	***	.41	.42			.42	.41		
Age 3	.42	.39	.47	***	.42	.42			.42	.42		
Age 5	.44	.41	.49	***	.43	.44			.43	.44		
<i>Early Head Start/Head Start participation</i>												
Age 1 (Early)	.05	.04	.05		.04	.05			.05	.05		
Age 3	.09	.08	.12	***	.09	.09			.08	.10	*	
Age 5	.16	.15	.18	**	.15	.16			.15	.16		
<i>Maternal depression</i>												
Age 1	.93 (2.12)	.83 (2.02)	1.17 (2.33)	***	.78 (1.97)	1.04 (2.23)	***		.82 (2.00)	1.03 (2.23)	**	
Age 3	1.26 (2.44)	1.14 (2.36)	1.55 (2.61)	***	1.01 (2.24)	1.46 (2.58)	***		1.09 (2.32)	1.42 (2.55)	***	
Age 5	1.06 (2.28)	.99 (2.22)	1.22 (2.40)	**	.88 (2.10)	1.21 (2.40)	***		.89 (2.10)	1.23 (2.42)	***	
<i>Worked during the past week</i>												
Age 1	.53	.51	.58	***	.49	.57	***		.50	.56	**	
Age 3	.47	.44	.55	***	.45	.49			.45	.50	*	

	Age 1				Age 3		Age 5	
	Full sample (N=3,870)	Not spanked (N=2,727)	Spanked (N=1,143)		Not spanked (N=1,705)	Spanked (N=2,165)	Not spanked (N=1,904)	Spanked (N=1,966)
Age 5	.44	.40	.54 ^{***}		.42	.45	.41	.46 [*]
<i>Number of adults</i>								
Age 1	2.17 (.97)	2.18 (.97)	2.15 (.96)		2.15 (.97)	2.18 (.97)	2.15 (.95)	2.19 (.99)
Age 3	2.03 (.91)	2.05 (.91)	1.99 (.93)		2.00 (.90)	2.06 (.92)	2.03 (.91)	2.03 (.91)
Age 5	1.94 (.91)	1.96 (.86)	1.91 (.85)		1.95 (.86)	1.94 (.85)	1.94 (.87)	1.95 (.85)
<i>Number of children</i>								
Age 1	2.32 (1.32)	2.36 (1.32)	2.22 (1.32) ^{**}		2.41 (1.37)	2.25 (1.28) ^{***}	2.37 (1.37)	2.27 (1.28) [*]
Age 3	2.33 (1.34)	2.38 (1.35)	2.21 (1.32) ^{**}		2.40 (1.41)	2.28 (1.29) [*]	2.37 (1.38)	2.29 (1.31)
Age 5	2.51 (1.35)	2.54 (1.36)	2.42 (1.34) [*]		2.60 (1.40)	2.43 (1.31) ^{***}	2.55 (1.39)	2.47 (1.32)

Note. Descriptive information averaged over 10 multiply imputed datasets. Means (and standard errors) or proportions (for dichotomous variables) presented.

* p<.05,

** p<.01,

*** p<.001.

\$watermark-text

\$watermark-text

\$watermark-text

Table 2

Indirect and Total Effects for Model D

	Externalizing behavior		Internalizing behavior		Cognitive scores	
	Unst. b (SE)	Stand. β	Unst. b (SE)	Stand. β	Unst. b (SE)	Stand. β
Indirect Effect						
S1→S3→O5	0.172 (.051)	0.027**	0.064 (.036)	0.013*	0.241 (.149)	0.014
S1→O3→O5	0.236 (.052)	0.037***	0.016 (.032)	0.003	-0.116 (.111)	-0.007
S1→O3→S5	0.008 (.002)	0.008***	0.000 (.001)	0.000	0.000 (.000)	0.000
E1→S3→S5	0.005 (.002)	0.018*	0.006 (.002)	0.019*	0.006 (.002)	0.019*
E1→O3→S5	0.008 (.001)	0.026***	0.001 (.001)	0.004	0.000 (.000)	0.000
E1→S3→O5	0.006 (.003)	0.003*	0.002 (.002)	0.001	0.008 (.006)	0.002
Total Effect						
S1→O5	0.349 (.051)	0.055**	0.079 (.084)	0.017	0.698 (.315)	0.041*
E1→S5	0.008 (.005)	0.029	0.008 (.005)	0.029	0.008 (.005)	0.029

Note. S1 is age 1 spanking, S3 is age 3 spanking, O3 is age 3 developmental outcome, O5 is age 5 developmental outcome, E1 is age 1 emotionality.

* $p < .05$.

** $p < .01$.

*** $p < .001$.