THE EFFECT OF CHILD DISABILITY/DELAY STATUS AND FUNCTIONAL ABILITIES ON FATHER INVOLVEMENT: AN APPLICATION OF PROPENSITY SCORE ANALYSIS

BY

DANIEL JAMES LAXMAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Human and Community Development
with a concentration in Human Development and Family Studies
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor Brent A. McBride, Committee Chair and Director of Dissertation Research
Professor Sarah C. Mangelsdorf, Northwestern University
Associate Professor Nancy L. McElwain
Professor Joseph H. Pleck
Professor Rosa M. Santos

Abstract

Past research on father involvement with children with disabilities/delays has been limited to observational/correlational studies that do not provide estimates of causal effects. Propensity score analysis was used to estimate the causal effect of child disability/delay status on father routine caregiving, literacy, play, and responsive caregiving involvement. The association between functional abilities and father involvement was also explored. Results indicate child disability/delay status influences father involvement early on and that changes in early levels of involvement are carried forward across the first four years of the child's life. The association between functional abilities and father involvement was less clear. Implications for future research and intervention work are discussed.

Dedicated to the fathers and their children with disabilities or delays
and in acknowledgement of my family and God without whose unfailing love
this work would have never been produced.

Acknowledgements

This project would not have been possible without the support of many people. Many thanks to my advisor, Brent A. McBride, who provided encouragement, guidance, and focus to the project. Also, thanks to my committee members, Sarah C. Mangelsdorf, Joseph H. Pleck, Rosa M. Santos, and Nancy L. McElwain, who offered guidance and support. Thanks to the American Education Research Association Grants Program for awarding me a Dissertation Completion Grant, providing me with the financial means to complete this project. And finally, thanks to my wife, Kristie Love, my baby boy, Jacob, my parents, and numerous friends who endured this long process with me, always offering support and love.

This research was supported by a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation under Grant #DRL-0941014. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

Table of Contents

Chapter One: Introduction and Literature Review…….……1
Chapter Two: Methods…….....22
Chapter Three: Results……...…..89
Chapter Four: Discussion……...129
References…….....146

Chapter One: Introduction and Literature Review

Interest in the role fathers play in children's development has increased dramatically in the past few decades among researchers, policy makers, and practitioners. As a result, a growing body of research has emerged demonstrating the positive impact men can have on their children's lives when they take active roles (Pleck & Masciadrelli, 2004; Lamb, 2010). One benefit of men's involvement with their children is increased school readiness in the form of better cognitive, social/emotional, and language ability. For example, research has found that fathers influence their children in a variety of developmental domains and across developmental stages including fewer behavioral problems during preschool (e.g., Frosch, Cox, & Goldman, 2001) and elementary school years (e.g., Aldous & Mulligan, 2002; Mezulis, Hyde & Clark, 2004), improved language and cognitive development during the early childhood years (e.g., Bronte-Tinkew, Carrano, Horowitz, & Kinukawa, 2008; Roggman, Boyce, Cook, Christiansen, & Jones, 2004; Tamis-Lemonda, Shannon, Cabrera, & Lamb, 2004), and more positive peer relationships during early childhood (e.g., Petit, Brown, Mize, & Lindsey, 1998). Longer term effects of father involvement have also been demonstrated. For example, Flouri and Buchanan (2004) found that father involvement when children were at age 7 predicted their educational attainment by age 20. For a review of methodologically rigorous studies of the association between father involvement and child outcomes, see Pleck (2010).

In light of the benefits of fathers' positive involvement for preparing their children to succeed in kindergarten, it becomes important to understand why some fathers are more involved than others. Research on the antecedents of father involvement may assist practitioners and policy makers in understanding how to best encourage positive father involvement and, in turn, promote children's school readiness. However, research on the antecedents of father involvement suffers the same limitation of the general literature on father involvement: a focus on fathering within the context of children who are typically developing. This limitation is problematic given that families of children with disabilities or delays may be at greater risk for poorer adjustment, such as higher levels of stress (MacDonald & Hastings, 2010) and lower quality parenting, including maltreatment (e.g., Sullivan & Knutson, 1998, 2000). This limitation is also problematic because increased father involvement may benefit these families. For example, fathers' positive attitudes, support, and involvement have been linked to improved maternal stress and marital satisfaction in families of children with disabilities or delays (Dyson, 1997; Saloviita, Itälinna, & Leinonen, 2003; Simmerman, Blacher, & Baker, 2001). Furthermore, Bronte-Tinkew and her colleagues (Bronte-Tinkew et al., 2008) reported that father involvement was related to greater reductions in cognitive delays for infants with disabilities than for typically developing infants. Given that children with disabilities or delays and their families face additional challenges and that father involvement may be particularly beneficial for children with disabilities, it would be important for

researchers, practitioners, and policy makers to understand what factors influence fathers' involvement in these families. In searching for antecedents of father involvement with children with disabilities or delays, researchers have turned to theory for guidance

Family Systems Theory and Father Involvement with Children with Disabilities

For fathers of typically developing children, several different theories have been suggested (see Pleck, 2007, for a review). In contrast, although several researchers have suggested factors that influence father involvement with children with disabilities or delays (e.g., Lillie, 1993), little theoretical work has been done specific to fathers of children with disabilities or delays. One theory that may be useful in examining the antecedents of father involvement with children with disabilities or delays is family systems theory. Family systems theory is a subtype of general systems theory. According to family systems theory, the family system is comprised of interrelated parts, or subsystems (Cox & Paley, 1997; Whitchurch & Constantine, 1993). These subsystems include the parent-child subsystem, the marital subsystem, the coparental subsystem, as well as others. What occurs in one subsystem influences and is influenced by what is occurring in another subsystem. For example, what occurs in the coparenting subsystem might influence what occurs in a parent-child subsystem. Furthermore, individuals within a subsystem mutually influence one another. For example, a father's behavior with his child might be influenced by what his coparent is doing with their child. Indeed, one study (Stoneman, Brody, & Abbott, 1983) found that fathers of children with Down syndrome interacted less when the mother was present, a phenomenon known in the developmental psychology literature as second-order effects.

Within the parent-child subsystem, characteristics of the child may influence father involvement. This idea is shared by Belsky's (1984) process model. This model identifies three domains of determinants of parental functioning: personal psychological resources of the parents, characteristics of the child, and contextual sources of stress and support. In families of children with disabilities, characteristics of the child, such as the presence, type, or severity of a disability/delay may be an influence on father involvement. Thus, family systems theory and Belsky's process model would suggest that a starting place for understanding father involvement with children with disabilities is to look at characteristics of the child, such as the presence of a disability/delay and its severity.

An advantage of family systems theory is that family challenges are considered a systemic problem as opposed to an individual problem. Blame for family problems is not assigned to individuals, but rather is assumed to be a result of dysfunction in the family system as a whole (although, for further discussion, see Whitchurch & Constantine, 1993). This principle has particular relevance to families of children with disabilities or delays who experience the stress, growth, and rewards associated with having a child with a disability or delay not only as individuals, but as a family (Seligman & Darling, 2007). Consequently, as researchers and practitioners try to understand how and why fathers of children with

disabilities or delays are involved, they should not focus solely on the "direct" effect of a child's disability/delay status or severity on father involvement. Rather, one should additionally consider how having a child with a disability might influence the quality of the mother-child, marital, and coparenting relationships and how changes in these subsystems lead to changes in father involvement. The present study restricts its examination to looking at the direct effect of child disability status and severity on father involvement. In doing so, I recognize that this does not capture all the influences of various family systems on father involvement. However, an improved understanding of the association between disability/delay status and severity and father involvement is a necessary first step before examining the effect of other family systems. This study is part of a much larger project that will consider the influence of all family systems on father involvement with children with disabilities or delays.

Conceptualizing Father Involvement

Before further discussion of father involvement with children with disabilities/delays, it is useful to first define father involvement. The most prominent approach to conceptualizing father involvement has been that of Lamb, Pleck, Charnov, & Levine (1985). The Lamb-Pleck model consisted of engagement, access, and responsibility. Pleck (2010) recently provided a reconceptualization of this model consisting of 5 components: positive engagement activities, warmth and responsiveness, control, indirect care, and process responsibility. Positive engagement activities include interactions with the child likely to promote development. The inclusion of the qualitative parenting dimensions of warmth and responsiveness and control reflects the convergence of parenting style research with paternal involvement research in several national data sets. Indirect care includes activities that are done for the child, but not with the child, such as purchasing food for a child, making doctor's appointments, and fostering connections within the community for the child. Process responsibility refers to fathers noticing what is needed by the child for the first four components. The current study focused on the positive engagement component, though the larger project will consider additional components of father involvement. For the present study, I focused on specific types of positive engagement (i.e., literacy, play, routine caregiving, and responsive caregiving involvement) available in the data set. I examined these specific types of father engagement because they address different needs of children, such as routine physical needs through routine caregiving, physical and cognitive stimulation through play and literacy activities, and less-routine physical needs through responsive caregiving. Furthermore, specific types of father involvement may be more strongly affected by disability/delay status or disability severity and some may not be affected at all. By examining the association between disability/delay status and severity and specific types of father involvement, a nuanced understanding of the effect of disability/delay status and severity on fathers' positive engagement can be obtained. This, in turn, will provide information to practitioners and policy

makers regarding which specific types of positive engagement should be selected as interventions targets so that resources might be used most effectively.

Past Research on Father Involvement with Children with Disabilities/Delays: The Need to Consider Disability/Delay Severity

There is mixed evidence that the quality and frequency of father-child interactions may be less optimal for fathers of children with disabilities/delays (Dyer, McBride, Santos, & Jeans, 2009; Pelham et al., 1997; Ricci & Hodapp, 2003; Roach, Orsmond, & Barratt, 1999; Sanders & Morgan, 1997). MacDonald and Hastings (2010) noted in their review of the literature on fathers of children with developmental disabilities that fathers of children with intellectual disabilities appear to engage in less caregiving activities than fathers of typically developing children. However, they suggest that the level of involvement may differ by the severity of the disability with father involvement being greater for very severe disabilities.

MacDonald and Hasting's (2010) suggestion corresponds with the efforts of other researchers who are exploring how tools assessing disability/delay severity may enhance research with children with disabilities/delays. The Individuals with Disabilities Education Act (IDEA) outlines 13 disability categories that are used for identifying children who may require special education and related services. Because reception of intervention services is in part determined by this classification system, research with children with disabilities/delays often uses disability or delay categories to study these children. Although a categorical approach is useful in studying children with disabilities/delays, a categorical approach fails to account for the significant variability in the severity of the disability/delay that may exist between children within the same category. Disability/delay severity may be defined as the degree of limitations in children's functional abilities within and across multiple developmental domains (e.g., cognitive functioning, motor skills, communication ability, etc.). For example, a child with autism spectrum disorder (ASD) may exhibit limitations within the domain of communication that may be mild, moderate, or significant. A child with ASD may also demonstrate limitations across several domains, including communication, cognitive functioning, etc. Consequently, a measure of disability/delay severity must consider the extent of limitations of functional ability both within and across developmental domains.

The use of a severity index that captures the degree of limitations within and across developmental domains may help improve research with children with disabilities/delays. For example, Virginia Buysse, Donald Bailey, and colleagues developed the ABILITIES Index (Bailey, Simeonsson, Buysse, & Smith, 1993; Buysse, Smith, Bailey, & Simeonsson, 1993), which measures the functional characteristics of children with disabilities/delays across nine domains (later expanded to 15). Chambers and colleagues (Chambers et al., 2004) found that the ABILITIES Index accounted for 40% of the

variance in total educational expenditures for special education students in a nationally representative sample of over 9000 children with disabilities. In contrast, disability status only accounted for 10% of the variance. Researchers have also explored the creation of post-hoc indexes of disability/delay severity within existing data sets. For example, Daley and colleagues (Daley, Simeonsson, & Carlson, 2009) created a post-hoc severity index for use with the Pre-Elementary Education Longitudinal Study (PEELS) data set. For all 8 cognitive, social/behavioral, and functional outcomes, the researchers found that the post-hoc severity index explained a significant amount of additional variance beyond that explained by disability/delay status alone. Thus, researchers have begun to demonstrate the utility of a severity index for use in research with children with disabilities/delays.

Researchers of father involvement with children with disabilities/delays have also begun to consider how disability severity is associated with father involvement. Two approaches have been utilized to study this association. The first approach has focused on differences in severity between groups of children with disabilities/delays. In this line of research, children with disabilities/delays are separated into groups based on their disability/delay status (e.g., ASD, Down syndrome, etc.). Researchers then examine the level of father involvement within each group and compare whether father involvement levels are lower or higher for groups of children with disabilities/delays that are considered more severe than for groups of children with disabilities/delays that are considered less severe. For example, a study by Ricci and Hodapp (2003) compared levels of father involvement with children with Down syndrome, a disability with generally less severe symptoms, with levels of father involvement with children with other intellectual disabilities and found no difference. Another study (Dyer et al., 2009) examined levels of father involvement across time in groups of children with disabilities/delays that differed in timing of diagnosis and in a group of children who were typically developing. Only one group difference emerged: one group of children with disabilities/delays did not experience the decline in fathers' functional involvement over time that the other groups experienced.

While group comparisons are a useful first step, they only provide information about how disability/delay severity at the group level is associated with father involvement, but not at the individual level. The group comparison approach fails to capture variation in severity within a disability/delay group and how such variation is associated with father involvement. Furthermore, group membership is often determined by the presence or absence of a target disability or delay. As a result, children with only that disability or delay are frequently placed in the same group with children who have been diagnosed not only with the target disability or delay, but with other disabilities or delays as well. The inclusion of children with multiple disabilities or delays in the group may bias the estimate of the effect of the target disability or delay on father involvement. There are other problems with this approach. For example, comparisons of father involvement across groups of children with disabilities or delays can be

problematic if any group has unique demographic characteristics. Such may be the case with families of children with Down syndrome. In one study, fathers of children with Down syndrome were found to exhibit greater warmth toward their child and report less depression than fathers of children with other intellectual disabilities (Stoneman, 2007). Since Down syndrome is considered to have less severe symptoms, these results would suggest that disability/delay severity is associated with a decrease in paternal warmth. However, the difference disappeared after controlling for family income, an important demographic correlate given that parents of children with Down syndrome are often older and have greater income. Demographic and other differences between groups that are not fully accounted for in the model may bias comparisons of father involvement between groups. Even when demographic or other differences are controlled for, group comparisons of disabilities/delays with differing levels of severity only provides information on the effect of disability severity at the group level. Individual variation in disability severity within each group is not addressed.

The second approach that has been used to examine the association between disability/delay severity and father involvement has explored how disability severity measured at the *individual* level is associated with father involvement. For example, initial research (McBride, Dyer, Santos, Laxman, & Jeans, 2011) using a national, longitudinal data set and a post-hoc severity index examined how disability severity at the individual level was associated with father involvement. In analyzing this association, the researchers fit models with both disability/delay status and disability/delay severity as predictors of father involvement. They found that the severity of the disability/delay was associated with lower levels of involvement and, in the case of literacy involvement, disability/delay status was associated with greater father involvement.

This second approach to examining the association between disability/delay severity measured at the individual level and father involvement is useful and an improvement over the previous approach which considered disability severity only at the group level. However, this approach is still limited because disability/delay status and disability/delay severity were treated as co-occurring characteristics of the child, but the causal relationship between disability/delay status and disability/delay severity was not specifically defined or modeled beyond acknowledging that the two were correlated. It is true that a child can be thought of as having both a disability/delay (status) and a set of limitations in functional abilities (severity) associated with that disability/delay. It is also true that the two are related in that a child with a disability/delay experiences limitations in functional abilities in one or more developmental areas. Indeed, diagnosis of a disability/delay is often made by examining a child's level of functioning within and across developmental domains, such as cognition, socioemotional development, physical development, etc. It is important to keep in mind, however, that disability/delay diagnosis is not the same thing as disability/delay status. For many disabilities/delays emerging in childhood, the disability/delay is present

long before diagnosis is possible. Furthermore, many disabilities/delays (e.g., autism spectrum disorder) are present long before any functional limitations appear because either the limitations are too difficult to accurately detect in young children or because children do not experience any functional limitations until they are failing to develop more advanced skills at an older age.

Defining the relationships between disability/delay status and disability/delay severity is challenging because there are multiple definitions of disability/delay. As noted by Simeonsson and colleagues (Simeonsson et al., 2003), "definitions of childhood disability are often characterized by overlap of health conditions, diagnoses or etiological factors" (p. 604). Furthermore, the *International Classification of Functioning, Disability, and Health* (ICF) used a biopsychosocial model of disability to define disability as "an umbrella term for impairments, activity limitations and participation restrictions" (World Health Organization, 2001, p. 213). According to the ICF, disability is an outcome of interactions between health conditions and contextual factors (external environmental factors and internal personal factors). In defining disability this way, the ICF attempts to integrate the medical and social models of disability. Controversy has surrounded the use of both models. The ICF takes the stance that:

> On their own, neither model is adequate, although both are partially valid. Disability is a complex phenomenon that is both a problem at the level of a person's body, and a complex and primarily social phenomena. Disability is always an interaction between features of the person and features of the overall context in which the person lives, but some aspects of disability are almost entirely internal to the person, while another aspect is almost entirely external. In other words, both medical and social responses are appropriate to the problems associated with disability; we cannot wholly reject either kind of intervention.
>
> A better model of disability, in short, is one that synthesizes what is true in the medical and social models, without making the mistake each makes in reducing the whole, complex notion of disability to one of its aspects (World Health Organization, 2002, p. 9).

The ICF considers interactions between all of the following in defining disability (though not all are used for classification purposes): health condition, impairments in body functions and structures, activity limitations, participation restrictions, environmental factors, and personal factors. Although the enumeration of all of these is useful in that it encourages researchers, policy makers, and practitioners to consider the multiple elements of disability, few research studies would be able to consider all of them simultaneously. The present study focuses on only a few of these elements of disability.

The definition of disability/delay severity used in this study (the degree of limitations in functional abilities within and across developmental domains) overlaps with the following elements identified by the ICF: impairments in body structure and function (physiological and psychological), activity limitations, and participation restrictions. Personal factors are considered in part in this study in

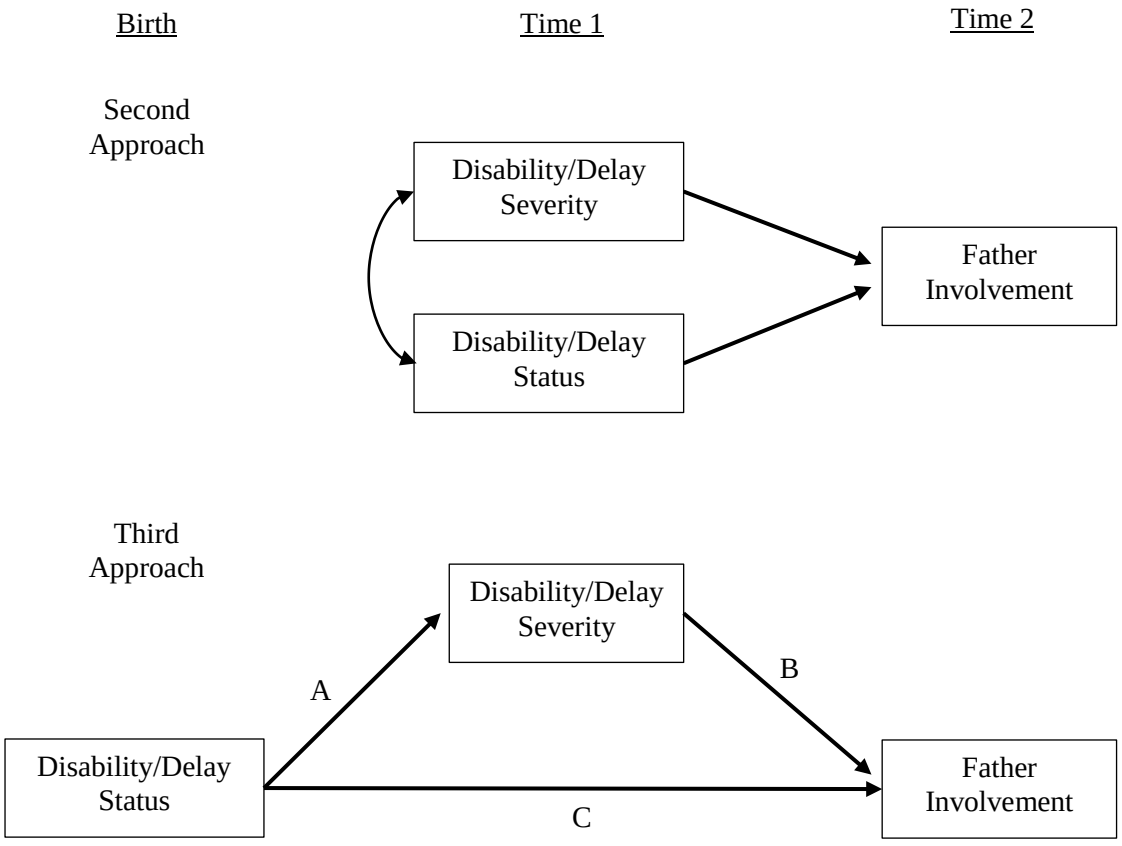
that analyses control for some personal factors (e.g., child's sex, race/ethnicity, etc.) in examining the association between disability status and severity and father involvement. An examination of how personal factors affect the association between disability status and severity and father involvement is beyond the scope of this study. Likewise, environmental factors are not directly considered in this study, though it should be noted that father involvement is itself an environmental factor.

"Disability/delay status," as used in this study, best maps on to the ICFs concept of "health condition," which includes diseases, disorders, injuries, trauma, etc. As noted before, diagnosis of a child's health condition (disability/delay status) is often made by examining a child's level of functioning within and across developmental domains. Often, the diagnosis of a health condition requires that a certain degree of impairment in body structure and function, limitation in activity, and or restrictions in participation be present. This would seem to imply that disability/delay status and disability/delay severity are the same thing. However, this is not the case because, as noted before, a child can have a disability/delay long before any limitations appear. Disability or delay status can be thought of as a characteristic of the child—a health condition—that is present from birth or is present following an incident of injury or illness after birth. In many cases, the disability or delay does not change (it is often time-invariant) meaning that the child continues to have the underlying health condition even though symptoms may change over time. In contrast, disability/delay severity almost always changes over time (it is time-variant) as a child fails to meet developmental milestones, loses previous obtained abilities, or gains new abilities as a result of intervention services or other environmental influences. For many disabilities/delays, interventions and other environmental factors may address and ameliorate limitations in functional abilities, but the underlying health condition is rarely considered cured. Thus, disability/delay severity is a distinct construct from disability/delay status, though both are elements of the umbrella concept of disability/delay.

With this background, it may be appropriate to think of disability/delay severity (limitations in functional abilities) as the manifestation of disability/delay status, or underlying health condition. Thus, one might say that disability/delay severity is partially the result of disability/delay status (environmental factors may be one of the other causes). A model that treats disability/delay status as a cause of disability/delay severity may be more appropriate than a model that treats them simply as co-occurring. If researchers were to use this third approach to study the association between disability/delay status and father involvement, disability/delay severity would be treated as a partial or full mediator of the association between disability/delay status and father involvement. A comparison of the second and third approaches is shown in Figure 1.

Figure 1.

Comparison of the Second and Third Approaches.



Disability/Delay Status, Disability/Delay Severity, Father Involvement, and Causal Inference

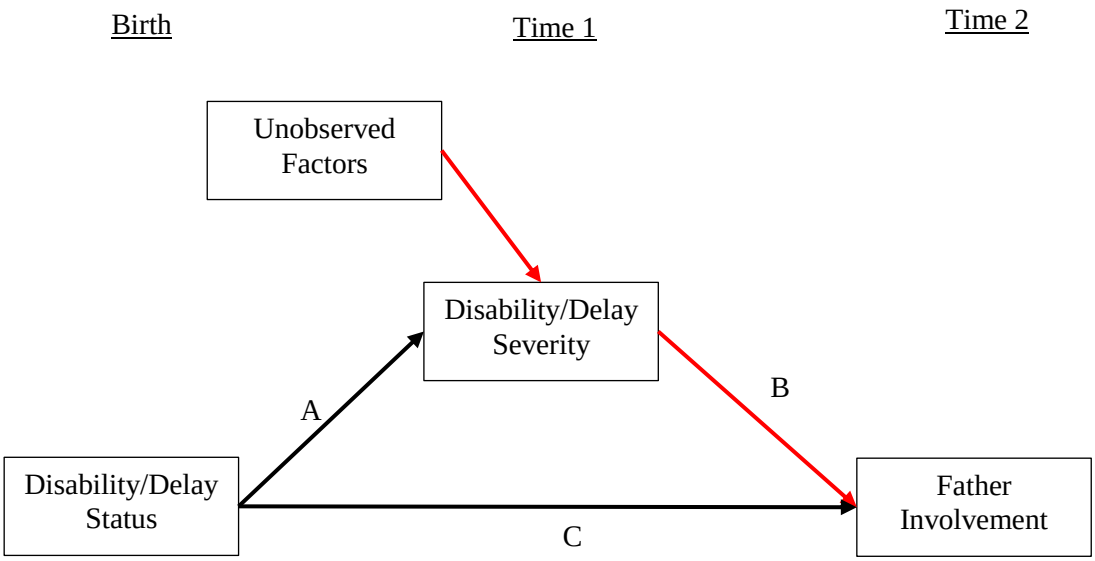
Although the third approach improves upon the second approach by treating disability/delay status as a cause of disability/delay severity, it still faces one limitation that is shared by the second and first approaches as well. All three approaches are limited in that they do not account for the effect of other factors beyond disability/delay status and severity that influence father involvement. While disability/delay status and severity are important predictors of father involvement, they are not the only predictors. Similarly, disability/delay status is not the only predictor of disability/delay severity. Other factors are associated with father involvement and may also be associated with disability/delay status or severity. These other factors need to be controlled for before any causal inference can be made about the effect of disability/delay status or severity on father involvement or the effect of disability/delay status on disability/delay severity. The presence of these other factors may bias the association between status, severity, and father involvement if these other factors are related to any combination of the three. For example, family socioeconomic status may be associated with the presence of a disability/delay, the severity of the disability/delay, and father involvement. Controlling for socioeconomic status in the analyses may help account for this bias. However, in order to demonstrate causal effects, one must control for all variables that are associated with any combination of disability/delay status, disability/delay severity, and father involvement. The addition of all "third variables" to the regression model leaves an increasingly smaller amount of variance left to be explained by the predictors of interest. Each additional variable reduces the power of the model to detect significant associations resulting in an increase in Type II errors. Furthermore, causal inference also requires the functional form of the association between each predictor variable and the outcome be correctly specified (i.e., linear, quadratic, cubic, etc.). The correct functional form of the association between a predictor and father involvement is often not known for many of the "third variables."

One approach that addresses these limitations is propensity score analysis (Guo & Fraser, 2010; Murnane & Willett, 2010; Rosenbaum & Rubin, 1983). This approach is a quasi-experimental design that essentially attempts to simulate an experimental design. In experimental designs, random assignments to treatment status allows for estimates of the causal effect of the treatment on the outcome because random assignment should nullify the effects of any third variables. Propensity score analysis artificially creates a control group for the treatment group. It does this by taking each member of the "treatment" group (i.e., children with disabilities/delays) and matching him or her to one or more members of the "control" group (i.e., children who are typically developing) who have a similar propensity for being in the treatment group (i.e., having a disability/delay), but do not. The two matched groups can then be compared to identify the causal effect of the treatment (i.e., disability/delay status) on the outcome (i.e., disability/delay severity or father involvement). Further details are given in the methods section.

Propensity score analysis works best when there are only two groups (a treatment and a control group). Propensity score analysis can be done when there are more than two groups, but it is considerably more difficult to successfully do so. For continuous measures, such as measures of children's disability severity, propensity score analysis is not a viable option. Consequently, the causal effect of disability/delay severity cannot be estimated (Path B in Figure 2). However, propensity score analysis could be used to improve the quality of the estimates of the effect of disability/delay status on father involvement through disability/delay severity because it would remove some bias from the estimates of Paths A and Paths C in Figure 2. However, as is the case with a true experimental design, it cannot provide *causal* estimates because the mediator occurs post-treatment (i.e., after a child is born with a disability/delay) and, consequently, may be influenced by factors other than treatment status (Jo, Stuart, MacKinnon, & Vinokur, 2011). These other factors may influence disability/delay severity for children with disabilities/delays differently than for children who are typically developing and may thus bias the estimates in the model. As can be seen in Figure 2, the presence of unobserved factors that influence disability/delay severity differently for children with disabilities/delays than for children who are typically developing biases the estimate of the effect of disability/delay severity on father involvement. Bias in this estimate in turn biases the estimates of the indirect effect of disability/delay status on father involvement through disability/delay severity. Consequently, estimates of the causal effects of disability/delay severity on father involvement or disability/delay status on father involvement *through* disability/delay severity cannot be obtained. This is a problem for both experimental and quasi-experimental methods exploring a mediator of the effect of a treatment on an outcome. Approaches to estimating the causal effects of a mediator such as disability/delay severity are currently being developed (e.g., Jo et al., 2011). However, these approaches are intended for use in randomized control trials and have not yet been applied to studies where treatment status was not randomly assigned. Because treatment status (disability/delay status) cannot be randomly assigned, this limitation cannot be overcome using current methodologies. However, this study's use of propensity score matching is justified in that it allowed the causal effect of disability/delay status on disability/delay severity and the causal effect of disability/delay status on father involvement to be estimated, even though estimates of other causal effects, including causal mediation, cannot be obtained.

Figure 2.

Mediation Model with Unobserved Factors.



An advantage of using propensity score analysis to examine the effect of disability/delay status on disability/delay severity and on father involvement is that there is no longer the need to control for all "third variables" because propensity score matching creates two groups as if they were randomly assigned. The "third variables" were used to create the propensity score on which children were matched. The current study used propensity score matching to identify the causal effects of disability/delay status on disability/delay severity and father involvement.

The logic of using propensity score analysis to estimate the causal effect of disability/delay status on father involvement requires that the study's disability/delay group only include disabilities or delays that existed at birth and thus prior to father involvement as conceptualized in this study. This does not mean that the symptoms are necessarily present or detectable at birth, but that the "foundation" of the disability or delay—the underlying health condition—is present at birth. If this is true, then prenatal biological factors (e.g., genetic inheritance and in utero development) and environmental factors would be sufficient to estimate each child's propensity for having a disability/delay. If the disability or delay is not present at birth, then it may be partially or fully the result of environmental factors that occur post-birth, including father involvement. Consequently, the estimate of the effect of disability/delay status on father involvement could include the effect of earlier father involvement on later father involvement through disability/delay status. The solution is to include only disabilities/delays in this study's disability/delay group that can be assumed to be present at birth as a result of genetic inheritance or in utero development, though they may not be diagnosable until much later. For example, because Down syndrome is clearly a result of a prenatal biological cause, children with Down syndrome should be included in the study sample. In contrast, if failure to thrive were caused by environmental factors, such as abuse or neglect, that were correlated with or include father involvement, then children diagnosed only with failure to thrive should not be included in the study sample. Many of the disabilities that would meet the criteria are often referred to as developmental or congenital disabilities/delays whereas disabilities not meeting the criteria might be referred to as acquired disabilities/delays (see DSM-V, American Psychiatric Association, 2013; Dalby et al., 2009). Further details about the criteria used to identify disabilities that are exclusively or primarily caused by prenatal biological factors are given in the methods section.

For this study, disability/delay status is assumed to be an underlying health condition that is present from birth and does not change, meaning that the child continues to have the disability/delay. While intervention services and other environmental factors may reduce the severity of children's disabilities/delays, the underlying health condition is assumed to continue to exist. Obviously, this assumption is not true of all disabilities/delays. However, the sample of children with disabilities/delays in this study is limited to children with disabilities/delays that are exclusively or primarily caused by prenatal biological factors. Given this sample restriction, it is safe to assume that the underlying health

condition of the disabilities/delays in this study is present from birth and continue to be present. In other words, if the environment could not have caused the disability or delay (the underlying health condition), then it likely could not remove it either. For this sample, severity of the disability/delay may change, but disability/delay status does not. Finally, given the sample restrictions, limitations in the representativeness of this study should be noted. Specifically, this study examined the effect of having a child with a disability/delay *that is exclusively or primarily caused by prenatal biological factors* on father involvement.

Disability/Delay Severity vs. Functional Abilities

Up to this point, I have referred to the extent of limitations in children's functional abilities within and across multiple developmental domains as disability/delay severity. However, "disability/delay severity" was calculated for children who do not have a known disability or developmental delay. Consequently, the term is somewhat misleading. Furthermore, research in special education in recent years has focused on examining the strengths and abilities of children with disabilities/delays, rather than solely focusing on the limitations associated with their disability/delay. Indeed, the ICF focuses on the levels of health of all people rather than on the disabilities, impairments, and limitations of those with a diagnosed health condition (World Health Organization, 2001). In consideration of these points, I use the term "functional abilities" to refer to the extent of children's functional abilities within and across multiple developmental domains (e.g., cognitive ability, general health, social skills, hearing).

Research Questions and Hypothesized Associations

The conceptual model for this study is presented in Figure 3.The current study explored four research questions:

1. How does the presence of a child's disability or developmental delay influence father involvement across the first four years of life?
2. How are children's functional abilities associated with father involvement?
3. Do children's functional abilities mediate the association between disability/delay status and father involvement?
4. How is father involvement associated with children's functional abilities and does father involvement mediate the effect of disability/delay status on functional abilities?

The first three questions draw from family systems theory and Belsky's (1984) process model and explore how characteristics of the child (i.e., disability/delay status and functional abilities) influence father involvement. The third research question combines the first two questions into a mediation analysis. Specifically, Question 1 examines Paths D-FI_1 and D-FI_2 in the model shown in Figure 3. (For clarity, the relevant paths for Questions 1 – 4 are highlighted in Figures 4 – 7, respectively.) Question 2 examines the Path FA. Question 3 adds Path D-FA_1 to the model to examine whether the effect of

disability/delay status on father involvement (Path D-FI_2) is mediated by children's functional abilities (Paths D-FA_1 and FA). Finally, Question 4 builds on Questions 1 – 3 and takes into account that, according to family systems theory, individuals within a subsystem *mutually* influence one another. Thus, in addition to children's functional abilities influencing father involvement, father involvement may also influence children's functional abilities (Path FI). The effect of father involvement on disability/delay status was not examined because disability/delay status is thought to be present from birth and thus cannot be influenced by father involvement which occurs after birth. (The sample for this study was restricted to children with disabilities/delays that would be present from birth). However, as was explored in Question 1, disability/delay status may influence father involvement. Father involvement, then, may be a mediator of the effect of disability/delay status on children's functional abilities. This possibility is suggested by the findings of Bronte-Tinkew and colleagues (Bronte-Tinkew et al., 2008) that father involvement was associated with reductions in cognitive delays for infants with disabilities. For Question 4, I examined whether the effect of disability/delay status on children's functional abilities (Path D-FA_2) was mediated by father involvement (Paths D-FI_1 and FI).

As shown in Figures 3 – 7, these questions were explored at multiple time points (birth, 9 months, 2 years, and 4 years). For Question 1, the effect of disability/delay status on father involvement was explored for father involvement at all three time points. For Question 2, the effect of children's functional abilities on father's subsequent involvement was examined from 9 months to 2 years and from 2 years to 4 years. The mediation components of Questions 3 and 4 were examined longitudinally from birth to 9 months to 2 years, from birth to 9 months to 4 years, and from birth to 2 years to 4 years. Finally, for Question 4, the effect of father involvement on children's functional abilities was examined from 9 months to 2 years, from 9 months to 4 years, and from 2 years to 4 years. Additional details are included in the methods section.

Past research has provided weak and inconsistent evidence that fathers of children with disabilities/delays are less optimally involved than fathers of typically developing children, or that fathers are less involved with children who are more limited in their functional abilities (Dyer et al., 2009; MacDonald & Hastings, 2010; Pelham et al., 1997; Ricci & Hodapp, 2003; Roach et al., 1999; Sanders & Morgan, 1997). Given these mixed findings and the observational/correlational nature of the studies, little guidance is available for hypothesizing the direction of the association between child disability/delay status and functional abilities and father involvement. Consequently, the analyses for Questions 1 and 2 were exploratory. Question 3 explored whether children's functional abilities mediated the effect of child disability/delay status on later father involvement. Given the mixed findings of past research, I did not hypothesize the direction of the indirect effect of disability/delay status on father involvement through children's functional abilities. However, I did hypothesize that child disability/delay status would have

some effect (positive or negative) on children's functional abilities as limitations in functional abilities appeared or developed over time as a result of the disability/delay. Children's functional abilities would, in turn, influence later father involvement. Question 4 examined the role of father involvement in contributing to children's later functional abilities and in mediating the effect of disability/delay status on children's functional abilities. Although the direction of the mediating effect cannot be specified based on our current knowledge, there is a substantial body of research suggesting that father involvement is beneficial for children in general (e.g., Lamb, 2010) and perhaps even particularly beneficial for children with disabilities/delays (e.g., Bronte-Tinkew et al., 2008). Consequently, I hypothesized that greater father involvement would be associated with higher levels of children's functional abilities.

Figure 3.

Conceptual Model.

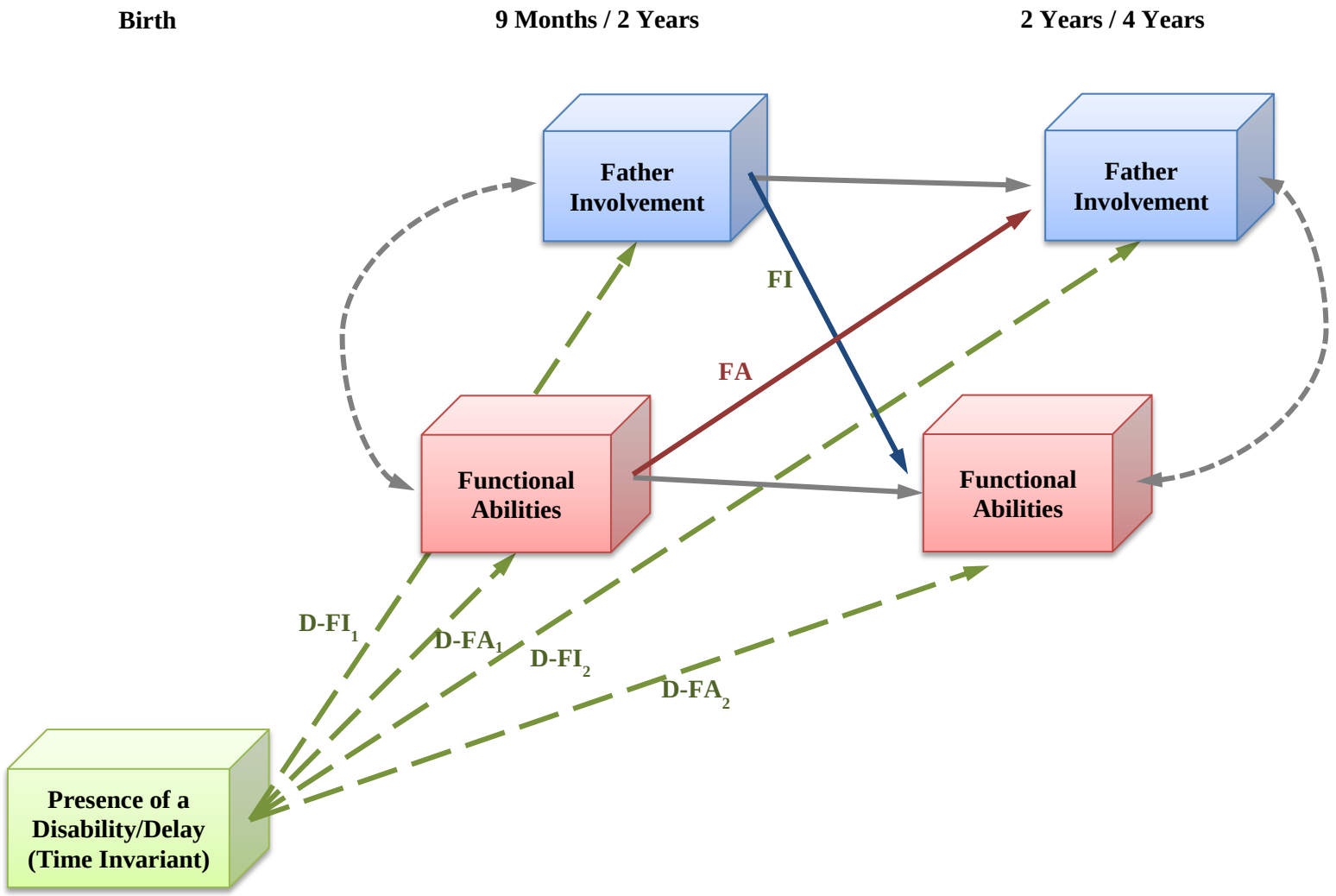


Figure 4.

Conceptual Model for Question 1: Effect of Child Disability/Delay Status on Father Involvement: Paths D-FI$_1$ and D-FI$_2$.

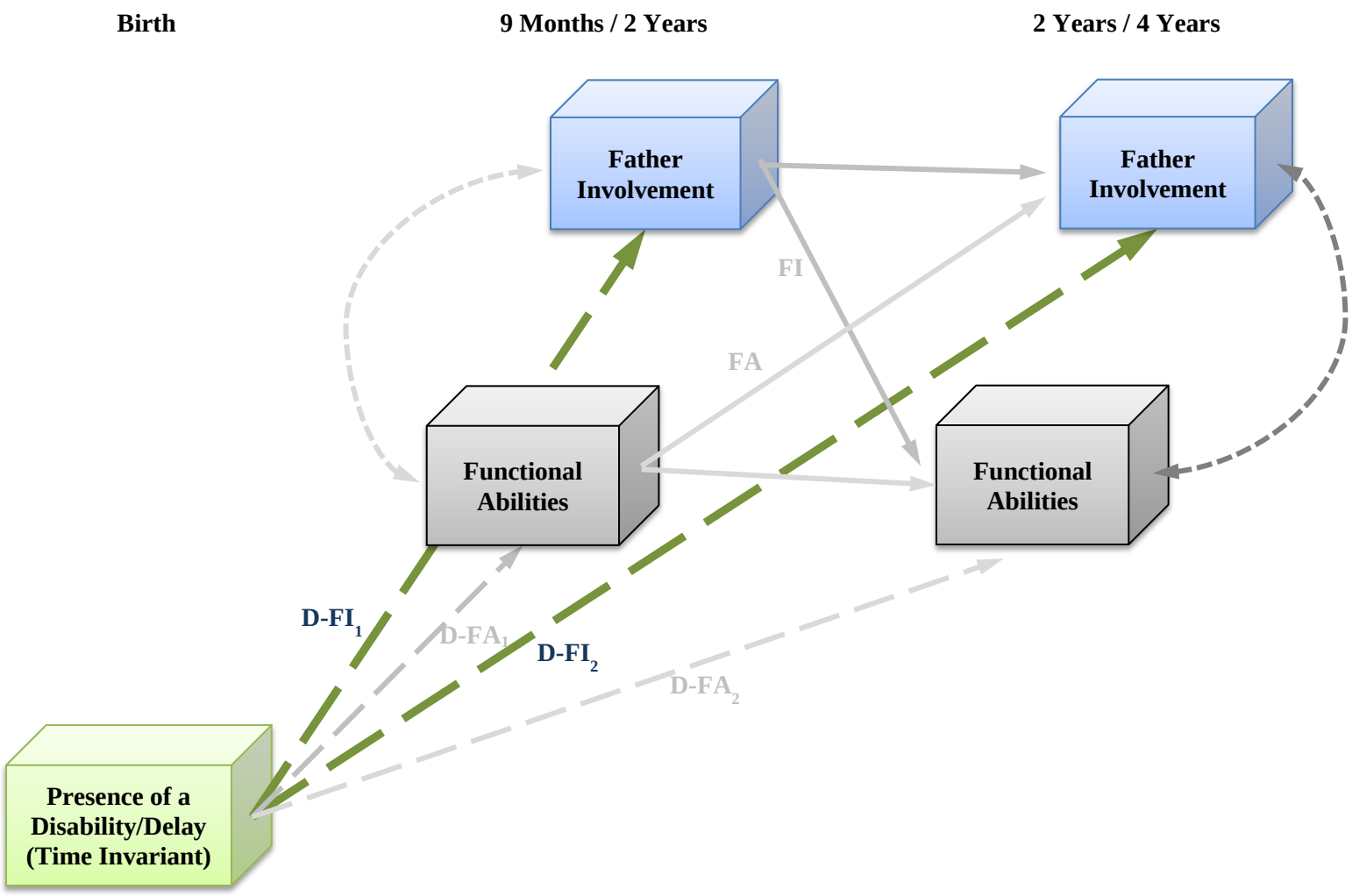


Figure 5.

Conceptual Model for Question 2: Longitudinal Association between Children's Functional Abilities and Father Involvement: Path FA.

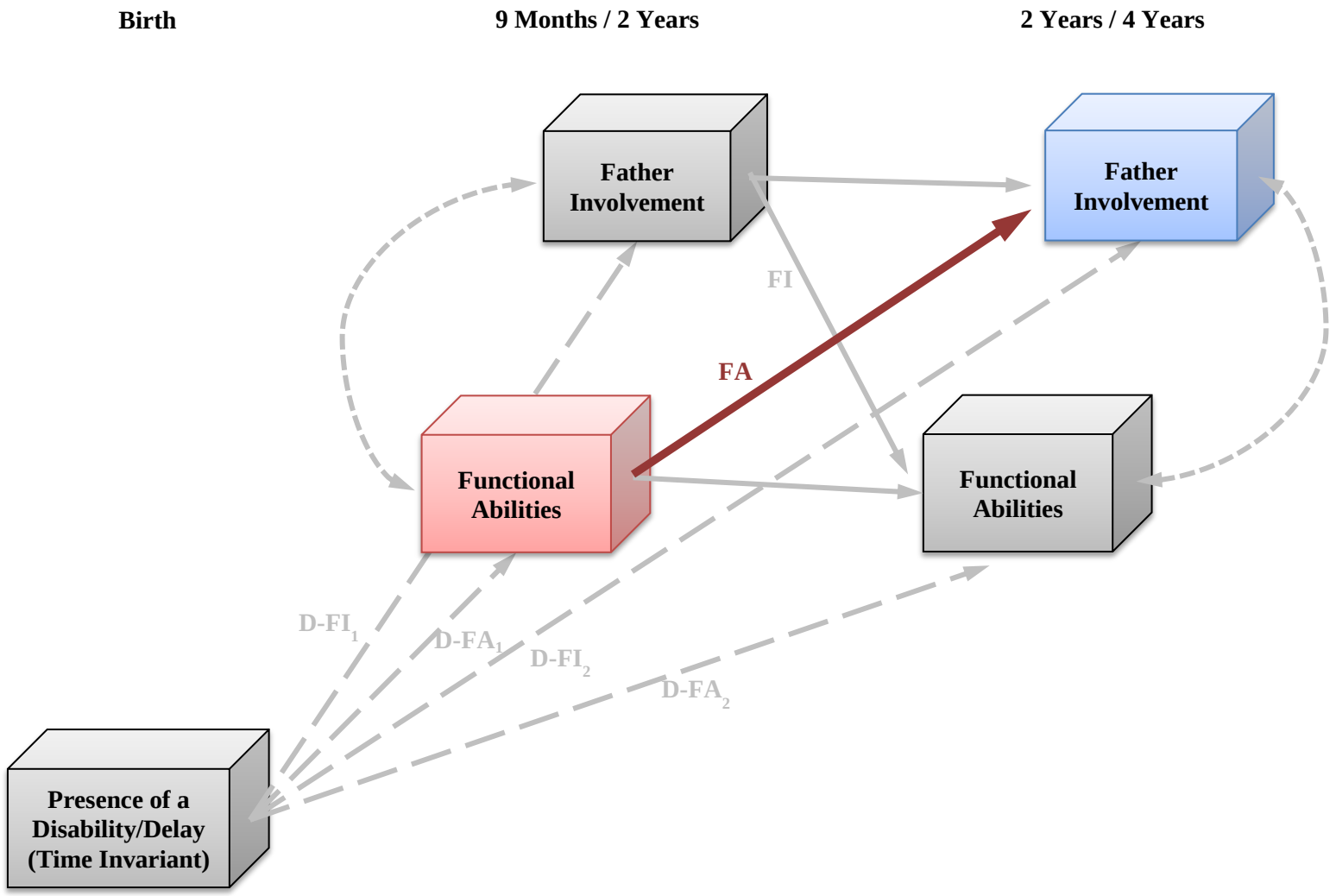


Figure 6.

Conceptual Model for Question 3: Mediation of the Effect of Child Disability/Delay Status on Father Involvement by Children's Functional Abilities: Path D-FI$_2$ by Paths D-FAI$_1$ and FA.

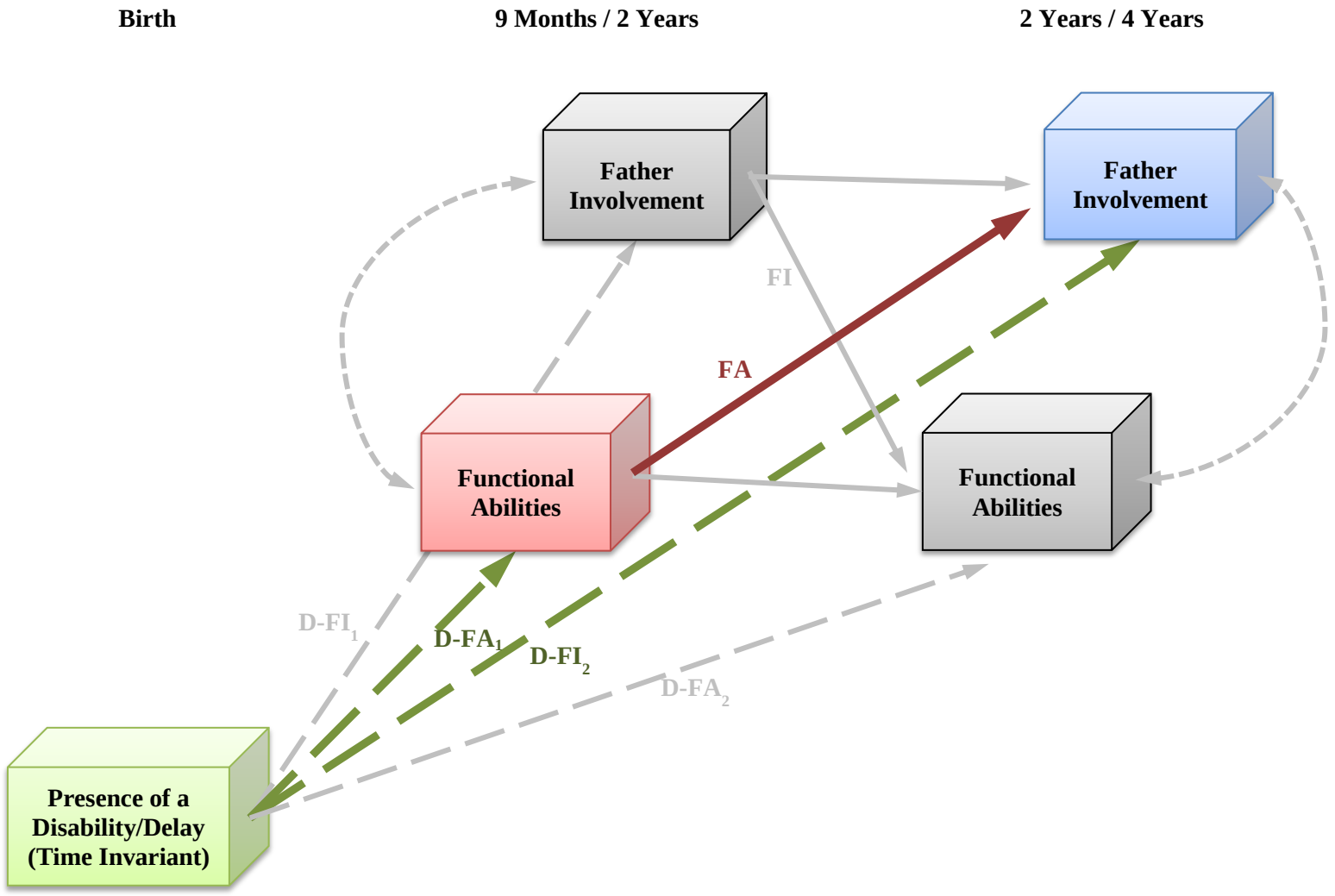
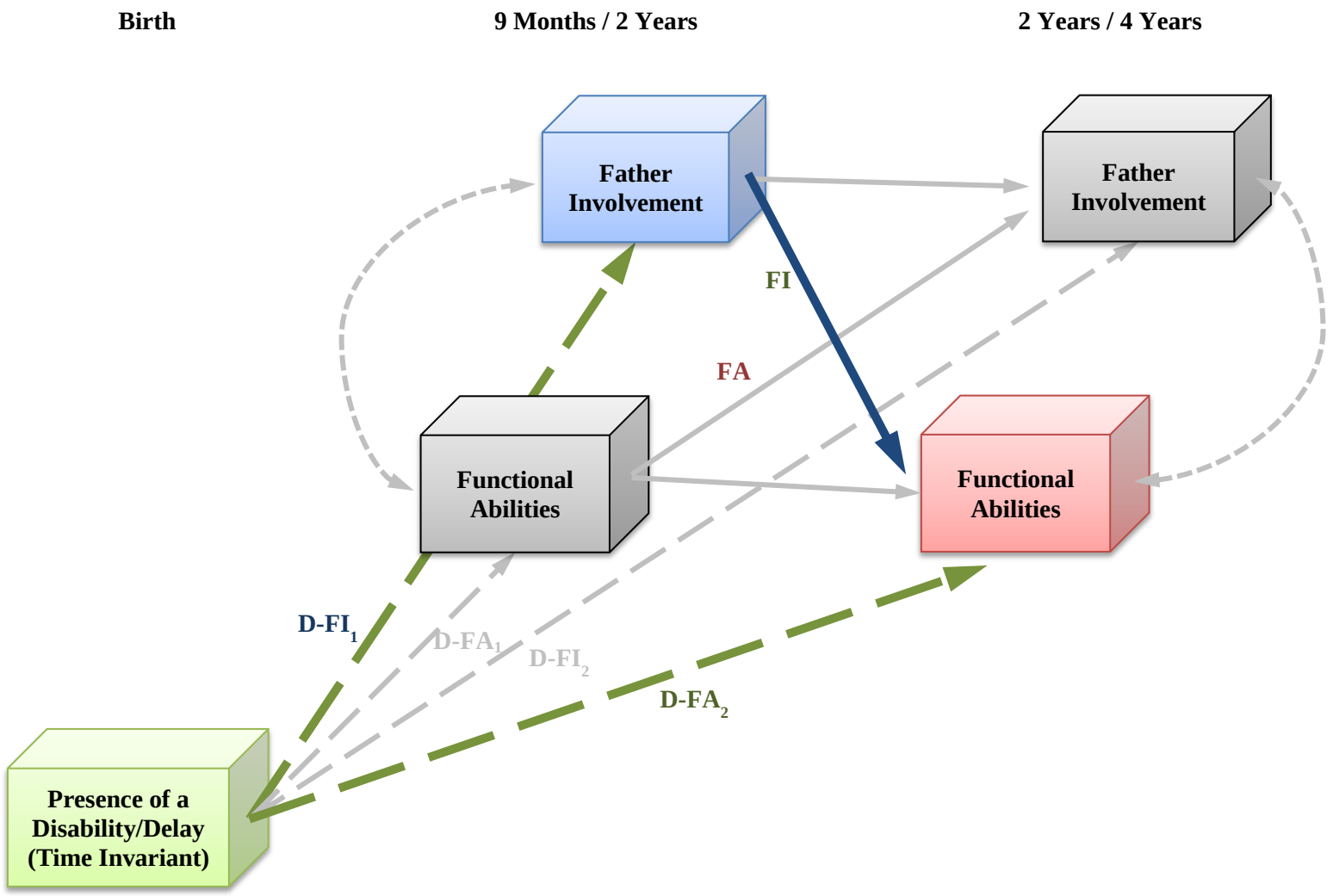


Figure 7.

Conceptual Model for Question 4: Longitudinal Association between Father Involvement and Children's Functional Abilities and Mediation of the Effect of Child Disability/Delay Status on Children's Functional Abilities by Father Involvement: Path FI and Path D-FA$_2$ by Paths D-FI$_1$ and FI.



Chapter Two: Methods

Data Source

The data for this project come from the National Center for Education Statistics (NCES) Early Childhood Longitudinal Study-Birth Cohort (ECLS-B) data set. When weights are applied, the data set is a nationally representative sample of children born in 2001 to mothers age 15 or older. Children who died or were adopted before the 9-month assessment were not represented in the data. Furthermore, over time, the sample is not nationally representative of children who died or moved abroad. The sample was drawn from US birth certificates and oversampled specific populations: low-birth-weight and very-low-birth weight children, twins, American Indians/Native Alaskans, Asian/Pacific Islanders, and Chinese children. Information on children and parents was collected when children were 9 months, 2 years, and 4 years of age using parent and teacher questionnaires, video recorded observations of parents and children, and direct assessments of children.

Complex survey design. Because the ECLS-B employed a complex survey design with a multi-stage cluster sample and oversampled certain subgroups, standard errors may be underestimated resulting in Type I errors and unweighted data would not be nationally representative. Therefore, I followed guidelines outlined by NCES (Snow et al., 2009) for adjusting standard errors using the jackknife replication method and the Taylor series method and applied appropriate weights so that findings can be generalized to the U.S. population of children born in 2001. Included with the data are multiple sets of weights generated by NCES to account for differential sampling probabilities and attrition rates. The particular set of weights to be used in an analysis is dependent on the type of data used (e.g., direct assessments, father report, etc.) and the time of assessment. I examined the associations between child disability/delay status, functional abilities, and father involvement at multiple time points (9 months, 2 years, and 4 years), which indicates the use of at least three different sets of weights (one for each time point). However, I examined the effect of child disability/delay status and functional abilities on father involvement at each time point *using the same sample*. Consequently, I used the weights for the 4-year time point for analyses examining parenting at all three time points (W3D0). There are fewer children with these weights at the 4-year time point because only children that still participated in the 4-year time point can have a weight, and several have dropped out over time. (It should be noted that the 4-year weights account for bias introduced by participant attrition). As a result of using the 4-year weights for all analyses, the sample size was smaller for analyses at 9 months and 2 years than if 9-month and 2-year weights had been used. Although the decreased sample size for analyses at earlier time points may result in less precision in estimates, the alternative of running analyses using a different set of weights may lead to problems with comparing the effect of disability/delay status and functional abilities on father involvement at one time point to the effect at another time point. This is because slightly different

samples would be used at each time point and the resulting estimates may not be fully comparable. However, the use of the 4-year weights still produce estimates based on a nationally representative sample of children born in 2001, so the only resulting loss from using the 4-year weights is that the 2-year and 4-year estimates may be less precise. Given the size of the sample, this is not a significant concern.

Sample

The ECLS-B contains data on 10,700 children. Note: per NCES requirements when using ECLS-B data, all N's in this paper were rounded to the nearest 50. Weights appropriate for the proposed analyses (i.e., 4-year weights for analyses examining father involvement, W3D0) were available for 3650 children and were constructed so that the 3650 children were representative of children born in the US in 2001.

Some restrictions on the sample were necessary. First, to ensure that the same individual completed the father questionnaires at each wave, it was necessary to restrict the sample to children living with their biological father through at least 4 years of age. This restriction was also necessary to examine the causal effect of disability/delay status on father involvement—I could only assume that I was estimating the effect of disability/delay status on father involvement if the person reporting father involvement at a later time point was the same person that has experienced and been influenced by the child's disability/delay since his or her birth. Second, because some of the prenatal information used to generate each child's propensity score was collected retrospectively from mothers at 9 months, it was necessary to restrict the sample to families in which the child still lived with his or her biological mother at 9 months. Third, the sample was restricted to children who were singletons. The reason for this restriction is that in the ECLS-B, fathers of twins, triplets, etc. reported their level of involvement on one questionnaire only. Consequently, each member of a set of twins, triplets, etc. has the exact same scores for father involvement. Inclusion of these cases would bias the analyses not only because twin's father involvement scores are not independent, but also because a child with a disability/delay and his or her twin without a disability/delay would have the exact same father involvement scores. Furthermore, there are likely qualitative differences in father involvement for families of twin children with disabilities/delays than for families of singleton children with disabilities/delays. Restricting the sample to singleton children avoids these potential sources of bias and provides for clearer interpretation of results. After implementing these restrictions, the sample consisted of 2700 children who lived with their resident biological father through at least 4 years of age. All children had a resident biological mother through 2 years of age and 99.89% of children had a resident biological mother through 4 years of age.

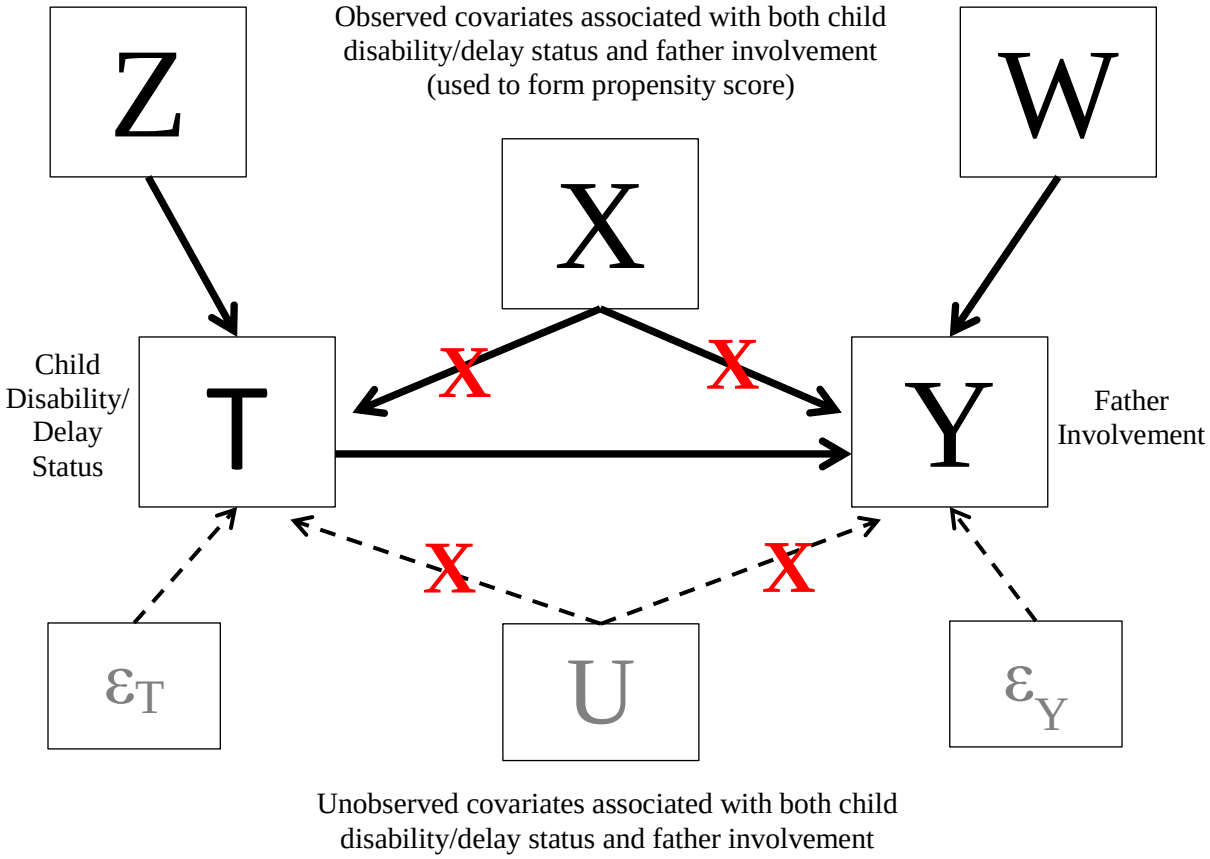
Analytic Plan Part 1: Overview of Propensity Score Analysis

In the explanation of propensity score analysis that follows, I discuss only father involvement as the outcome to simplify the presentation. The same logic applies to predicting children's functional

abilities from disability/delay status. As previously noted, comparisons of fathers of children with disabilities/delays and fathers of children who are typically developing have revealed differences between the two groups, although findings have been mixed. However, those differences or lack of differences may be a result of factors other than disability/delay status such as any of a number of risk factors for child disability/delay status that may also be associated with father involvement (e.g., parents' age, poor prenatal care, SES, etc.). These risk factors are represented by X in Figure 8. Father involvement is represented by Y. Child disability/delay status is represented by T, or "treatment" status. If not controlled for in analyses, the presence of "third variables" (X) biases the estimates of the effect of child disability/delay status leading to an incorrect estimate of the causal effect of disability/delay status on father involvement. Other third variables that are not observed (i.e., are not included in the model because they were not measured) also lead to biased estimates and are labeled U in Figure 8. The use of a true experimental design with random assignment to a treatment or condition (e.g., disability/delay vs. typically developing) would, in effect, remove all Us and Xs from the model. Random assignment to treatment status results in their being no significant association between U and T or X and T. Furthermore, there would be no significant association between U and Y and X and Y because randomization would balance the effect of U and X on Y. The use of a true experimental design to examine the effect of disability/delay status would be impossible (and unethical if it were possible), but an approach that approximates an experiment is the use of propensity score matching (Dehejia & Wahba, 2002; Guo & Fraser, 2010; Murnane & Willett, 2010; Rosenbaum & Rubin, 1983).

Figure 8.

Statistical Model.



Before explaining the logic of propensity score analysis, it is helpful to review further how this study would be done using an experimental design. In an experimental design, children would be randomly assigned to one of two "treatment" conditions: control = typically developing, treatment = having a disability or delay. In theory, the best way to examine the effect of disability/delay status on father involvement would be to take each child and father and observe father involvement with the child when the child had a disability/delay, rewind the clock, and observe father involvement with the child again, but this time when the child did *not* have a disability/delay. Then we could compare the difference in father involvement with the same child between the two different conditions to estimate the effect of disability/delay status. This cannot be done even in experimental studies because the outcome is never observed under its counterfactual condition (Holland, 1986; Rubin, 1974) (i.e., a father observed raising a child with a disability/delay cannot be observed raising the same child without a disability/delay and vice versa). However, random assignment to one or the other experimental conditions results (when done correctly and with a large enough sample) in two groups that are equivalent to each other in every way except treatment condition. This is not to say that every member of a group is the exact same as every member of the other group. Rather, every group member has unique characteristics, but these characteristics are spread randomly between the two groups so that neither group has a significantly higher proportion of any characteristics. Since the two groups are identical (at the group level) in all ways except treatment status, it can be assumed that how the control group performs on the outcome is a good estimate of how the treatment group would have performed in their counterfactual condition (i.e., in the control condition). Thus, a comparison of the two groups could yield an estimate of the effect of the treatment for those who received the treatment.

If this study were done as a randomized experiment, the process of randomly assigning children to the disability/delay condition or the typically developing condition could be thought of as matching the group of children with disabilities/delays to the group of typically developing children. The two groups would be a good match because they would be identical as a group in every way except treatment status. There are limitations to this approach because a researcher is not matching at the individual level, but at the group level. As a result, each child does not have a perfect match, but group-level matching through random assignment done correctly and with a large enough sample addresses these limitations. In a fictitious world, the best approach would be to observe each child and father under both conditions. Thus, each father-child pair would be its own perfect match and be matched at the individual-level (or, in this case, pair-level). Again, this is not possible, but the idea of matching a father-and-child-with-a-disability/delay pair to a father-and-child-who-is-typically-developing pair to estimate the effect of disability/delay is the foundation of propensity score analysis.

Propensity score matching consists of matching each treatment case (i.e., child with a disability/delay and his or her father) to one or more control cases (i.e., typically developing children and their fathers) that are similar on a vector of pre-treatment (i.e., pre-birth) covariates **X**. The control case(s) are used to estimate the counterfactual for the treatment case, or, in other words, the outcome we would observe if the treatment case could be observed under the control condition (i.e., if we could observe father involvement with a child with a disability if the child had instead been born *without* a disability/delay and was typically developing). The purpose of this study is to estimate the effect of child disability/delay status on father involvement for children with a disability (i.e., the average treatment effect on the treated, ATT):

$$ATT = \bar{\delta}_T = \bar{Y}_T^t - \bar{Y}_T^c \quad (1)$$

In Equation 1, the average treatment effect on the treated $\bar{\delta}_T$ is the difference between $\bar{Y}_T^t$, or the outcome for the treatment cases under the treatment condition (i.e., father involvement with children with disabilities under the condition of the children having a disability/delay) and $\bar{Y}_T^c$, or the outcome for the treatment cases under the control condition (i.e., father involvement with children with disabilities under the condition of the children *not* having a disability/delay). Although $\bar{Y}_T^t$ can be observed and estimated, $\bar{Y}_T^c$ cannot. However, we can observe $\bar{Y}_C^c$, or the outcome for the control cases under the control condition (i.e., father involvement with children who are typically developing under the condition of the children being typically developing). Implementing the conditional independence assumption, we assume that under the control condition, the outcomes are independent of treatment assignment, conditional on **X.** In other words, the conditional independence assumption holds that, conditional on **X,** *the level of involvement of fathers with children with disabilities/delays under the typically developing condition* is equivalent to and can thus be estimated by *the level of involvement of fathers with typically developing children under the typically developing condition*:

$$(\bar{\boldsymbol{Y}}_T^c \mid \boldsymbol{X}) = (\bar{\boldsymbol{Y}}_C^c \mid \boldsymbol{X}) \quad (2)$$

Another way of saying this is that conditional on **X,** "assignment" to the disability/delay or the typically developing condition is as if by random. Because we are now conditioning on **X**, Equation 1 becomes:

$$\bar{\delta}_T = \sum_X \frac{n_{Tx}}{n_T} [(\bar{\boldsymbol{Y}}_T^t \mid \mathbf{X}) - (\bar{\boldsymbol{Y}}_T^c \mid \boldsymbol{X})] \quad (3)$$

The term $\sum_X \frac{n_{Tx}}{n_T}$ is used for weighting in subsequent analyses and is discussed later. Based on the conditional independence assumption, we can substitute $(\bar{\boldsymbol{Y}}_C^c \mid \boldsymbol{X})$ for $(\bar{\boldsymbol{Y}}_T^c \mid \boldsymbol{X})$ in Equation 3 yielding a formula for the average treatment effect on the treated (ATT) that can be observed:

$$\bar{\delta}_T = \sum_X \frac{n_{Tx}}{n_T} \left[\left(\bar{\boldsymbol{Y}}_T^t \,\middle|\, \mathbf{X}\right) - (\bar{\boldsymbol{Y}}_C^c \mid \boldsymbol{X})\right] \quad (4)$$

The conditional independence assumption holds only if **X** includes all relevant pretreatment variables that are associated with both disability status (T) and father involvement (Y). In other words, there are no

unobserved variables (U) that affect both disability/delay status and father involvement that were not included in **X** (see Figure 8, noting that the paths from U to T and Y have been crossed out—the paths from X to T and Y have also been crossed out because they are addressed through matching as discussed below). Thus, the covariates included in **X** should be exhaustive. Otherwise, estimates of the ATT may be biased.

If all relevant pretreatment covariates are included in **X**, then each father-child-with-a-disability/delay pair can be matched to one or more father-child-who-is-typically-developing pairs who have similar values for all of the covariates in **X** to estimate the effect of disability/delay status on father involvement**.** If there are multiple covariates included in **X**, it would be extremely difficult or even impossible to match on all covariates. For example, if **X,** included only to dichotomous covariates, then there would only be 4 configurations to match on. However, the addition of 3 more dichotomous covariates increases the number of configurations to 32, and the addition of 5 more increases the number to 128, and so on. Fortunately, Rosenbaum and Rubin (1983) demonstrated that matching on the propensity score is equivalent to matching on all the covariates in **X**. The propensity score represents the probability ("propensity") that a child has a disability based on **X.** A child's propensity score can be estimated by fitting a logistic regression model in which disability/delay status is predicted from all of the covariates in **X.**

Once the propensity score is formed, cases are ordered by their propensity score. Then, each individual in the treatment condition (child with a disability/delay) is matched to one or more individuals from the control condition (children who are typically developing) with a similar propensity score. Control cases can be used as a match for more than one treatment case ("matching with replacement"). A caliper can be used to restrict the range of possible matches so that matched control cases are sufficiently similar in their propensity score to treated cases. Rossenbaum and Rubin (1985) suggested that a caliper of 0.25 standard deviations may provide an effective range. For example, assume the standard deviation of the propensity score was 0.80 and a specific child with a disability/delay had a propensity score of 0.22. Only children who are typically developing and have a propensity score in the range of 0.22 – 0.25*0.80 = 0.02 and 0.22 + 0.25*0.80 = 0.42 are eligible matches.

When children with disabilities/delays are matched to children who are typically developing, the matching is reflected in the dataset through weighting. By using the weights in the analyses, estimates of the causal effects can be obtained. In practice, each child with a disability/delay is given a weight of 1. The weighting of matched children who are typically developing depends on the method used for matching. In the case of caliper matching with replacement, each child who is typically developing can be used as a match for multiple children with disabilities/delays and may be one of any possible number of matches for a specific child with a disability/delay (e.g., 1 of 5, 1 of 12, etc.). For example, if a child who

is typically developing is matched to 3 children with disabilities/delays and is 1 of 5, 1 of 4, and 1 of 2 matches for the 3 children, then the child's weight will be 0.20 + 0.25 + 0.50 = 0.95.

Because the ECLS-B dataset already includes weights to make the sample nationally representative, the weighting procedure has to be adjusted slightly. Specifically, instead of being assigned a weight of 1, each child with a disability/delay is given the weight provided in the ECLS-B appropriate for these analyses (W3D0). For example, assume that a child who is typically developing is matched to 3 children with disabilities/delays and is 1 of 5, 1 of 4, and 1 of 2 matches for the 3 children and the 3 children have W3D0 weights of 320, 1450, and 2022, respectively. Then the typically developing child's weight would be 0.20*320 + 0.25*1450 + 0.50*2022 = 1437.5. By assigning weights to children who are typically developing in this fashion, the sample of children who are typically developing is reweighted to be similar to the sample of children with disabilities/delays. Estimates based on this weighting reflect the average treatment effect on the treated (ATT), or the effect of having a disability/delay on children who have disabilities/delays. Estimates are nationally representative of the effect of having a disability/delay *for children who have disabilities/delays.* The ATT is often what researchers are interested in and this practice has been used in previous studies using propensity score analysis and national data (e.g., Reardon, Cheadle, & Robinson, 2009). Once the matches are created, then the conditional independence assumption can be tested.

As previously noted, the conditional independence assumption holds that, conditional on **X**, treatment assignment is random. This assumption can be partially tested by checking for balance on the covariates in **X** in the matched sample of children with disabilities/delays and their typically developing controls. Balance between groups on the covariates in **X** is necessary, but not sufficient for the conditional independence assumption to hold. Thus, I must be confident that **X** includes all necessary covariates to account for the selection mechanism funneling children into the group of children with a disability/delay and the group of children who are typically developing. However, there is no way to confirm that I have done so. Balance checking is done by checking for mean differences between children with disabilities and their typically developing match(es) on each covariate. Although the two groups are matched using a propensity score, balance checks still need to be done for each covariate in **X** that was used to create the propensity score in order for the conditional independence assumption to hold. If treatment assignment is random conditional on **X**, there should be no mean differences between treatment groups on any covariate (although 1 in 20 covariates could have a significant difference just by chance using a cutoff of $\alpha = .05$). A significant mean difference indicates a lack of balance. If balance is not achieved, then the logit model creating the propensity score needs to be revisited. The addition of higher-order (e.g., quadratic, cubic, etc.) and interaction terms to the logit model may yield a model that better approximates the selection mechanism placing children into each of the two groups. For example, it may be that family

socioeconomic status interacts with race/ethnicity in determining child disability/delay status. Adding this interaction to the model may produce a matched sample that is balanced. After the logit model has been modified, the propensity score is again generated, matches are identified, and balance checks are repeated with the new propensity score and matched sample. This process is repeated until balance is achieved.

Once balance is achieved, the matched sample can be used to estimate the effect of child disability/delay status on father involvement for children with a disability/delay (the ATT). This is done by fitting a regression model on the matched sample with only disability status as the predictor (although covariates or the propensity score itself can be included to obtain more conservative estimates of the causal effect).

Strengths and Limitations of Propensity Score Analysis: Why Bother?

Given the complexity of propensity score analysis, one may ask why not just include all the relevant covariates in a regression model and estimate the causal effect of disability/delay status using this simpler approach? There are a number of advantages of propensity score matching over regression/covariate adjustment. First, as noted previously, to demonstrate causal effects one must control for all variables that are associated with any combination of disability/delay status, functional abilities, and father involvement. The addition of all "third variables" to a regression model leaves an increasingly smaller amount of variance left to be explained by the predictors of interest. Furthermore, each additional variable reduces the power of the model resulting in an increase in Type II errors. Propensity score matching does not require the inclusion of all covariates in the model estimating the effect of the treatment on the outcome because the matching process is done so that there are no mean differences between the treatment and control group on any covariates. (All of the covariates were already included in creating the propensity score and identifying the matches and so do not need to be included again in the final model estimating the effect of the treatment.) A limitation of including all of the covariates in the regression/covariate adjustment approach is that the inclusion of multiple covariates may lead to problems with multicollinearity if two or more covariates are highly correlated, which they often are. Although one of the offending covariates can be dropped or the covariates can be combined, this results in a loss of information. When covariates are used to generate the propensity score, multicollinearity is not a concern because the focus is not on interpreting the effect of any specific covariate on the probability of having a disability/delay, but rather on generating the best propensity score. In propensity score analysis, all variables can be included in generating the propensity score. In sum, propensity score analysis has greater flexibility to include all relevant covariates than the regression/covariate adjustment approach.

A second advantage of propensity score analysis over regression/covariate adjustment is that no assumption is made about the functional form of the relationship between the treatment and the outcome. The functional form of the association between the treatment and the outcome refers to whether the

association is linear, curvilinear, etc. For example, a predictor variable in a linear regression model is assumed to have a linear relationship with the outcome unless some other relationship is specified (e.g., quadratic or curvilinear). Often the accuracy of the functional form of the relationship between a predictor and the outcome goes unchecked in regression analysis because a linear functional form is assumed. Furthermore, researchers often do not have any theory to guide what functional form the relationship should take. In contrast, when conducting propensity score analysis the functional form of the association is not assumed at all because individuals are matched at similar levels of the covariates determining treatment status, whether the functional form of the covariate-outcome association is linear, quadratic, etc.

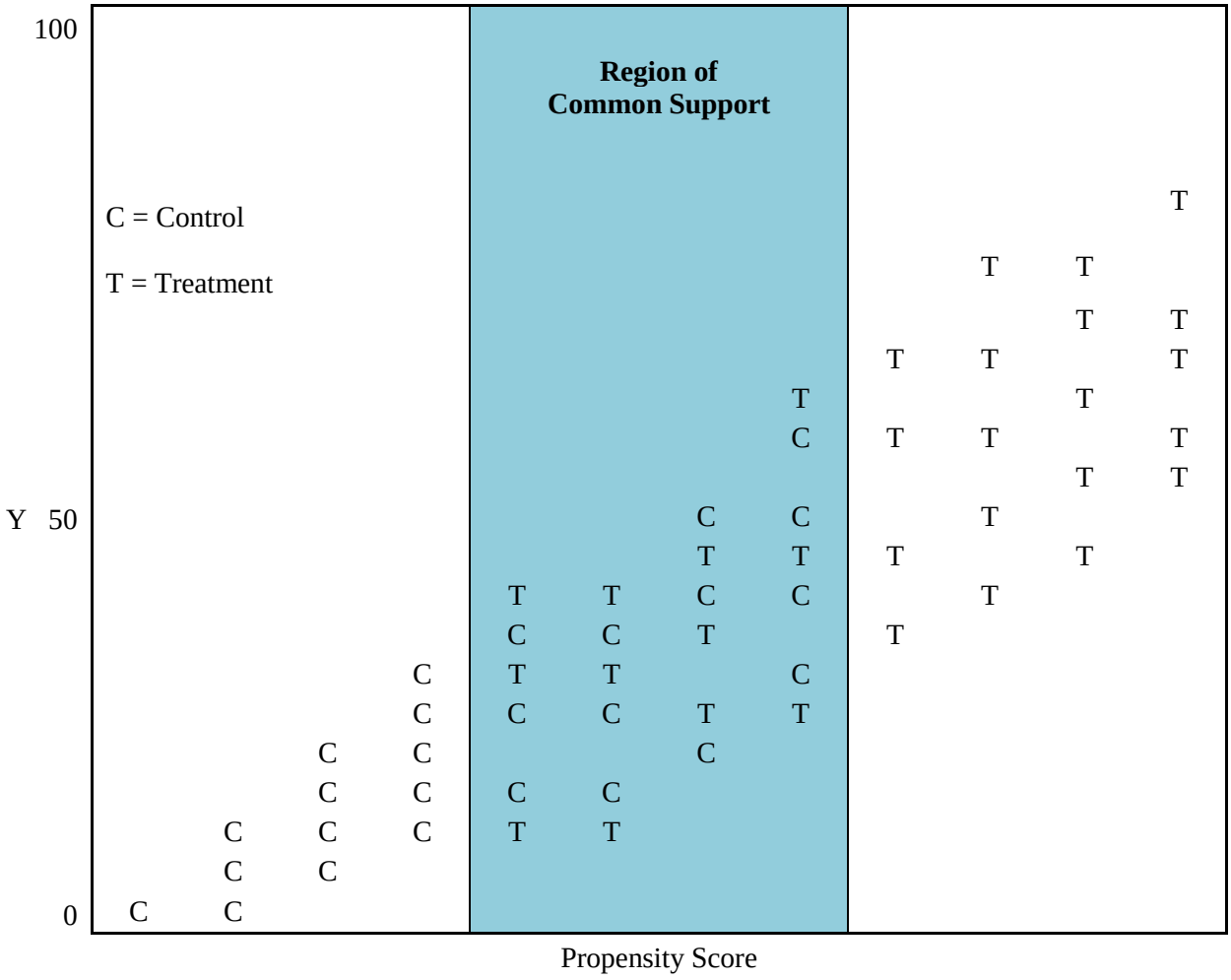
A third advantage of propensity score analysis is that it allows for partial testing of the conditional independence assumption. As previously noted, the conditional independence assumption holds that, conditional on **X**, treatment assignment is random. The assumption can be partially tested by checking whether there are mean differences between the treatment group and its matched control group on the covariates used to create the propensity score. The assumption cannot be tested in a regression model because balance on a covariate is conditional on all other variables included in the model.

A fourth advantage of propensity score analysis is that the region of common support is explicitly clarified. The region of common support refers to the values for the vector of covariates in **X** for which there are cases in both the treatment and control groups. If there are not cases in both conditions at a given value of **X**, then the causal effect of the treatment on the outcome cannot be estimated for that value of **X**. To illustrate this, we might consider only one covariate: family socioeconomic status. Suppose in this fictitious example that children with disabilities were only represented in the middle and lower levels of family SES while children who are typically developing were represented only in the middle and higher levels. If this were the case, then there is overlap between the two groups only for the middle levels of SES, but not at the higher or lower levels. The same logic applies when considering a vector of covariates, **X**, or the optimal combination of these covariates (i.e., the propensity score). This can be seen in Figure 9 created from fictitious data. In this figure, the outcome value for each case is plotted against the propensity score. Treatment cases are indicated by a T while control cases are indicated by a C. Only for the treatment cases inside the blue, center box is there at least one control match for the treatment case, and only for these cases can the effect of the treatment on the outcome be examined. The region inside the box is the region of common support, or the region where there is adequate support (i.e., representatives of both the treatment and control group) to estimate the causal effect. When conducting propensity score analysis, the values of **X** or the propensity score for which there is overlap between the groups is explicitly examined. In contrast, the regression/covariate adjustment approach does not explicitly identify the region of common support. Furthermore, regression analysis extrapolates the relationship between the treatment and the outcome beyond the region of common support into the

regions in which cases from only one of the groups is present. For example, using the example of family SES again, regression analysis extrapolates the relationship between disability/delay status and the outcome beyond the middle levels of family SES. While there is representation in the middle levels of family SES by both groups, only one group is represented in the high levels of family SES and the other in the lower levels of family SES. Traditional regression analysis extrapolates the estimated relationship across the full spectrum of family SES, which may or may not be justified. Propensity score analysis does not have this limitation.

Figure 9.

Region of Common Support.



Although propensity score analysis has several strengths, it has limitations as well. One limitation of propensity score analysis compared to the regression/covariate adjustment approach is that propensity score analysis may have larger standard errors because of a smaller sample size. Sample size is often smaller for analyses using propensity score analyses because not all cases in the control group will be a good match for cases in the treatment group. A smaller sample size and the resulting larger standard errors may elevate the risk of a Type II error. However, given the large sample size used in this study, this is not a significant concern. Propensity score analysis is also limited in comparison to a true experimental design. First, estimates of causal effects obtained using propensity score analysis can only be generalized to the region of common support. For example, returning to the fictitious family SES illustration, the results of propensity score analysis on such a data set would only generalize to children and fathers in the middle levels of family SES. In contrast, a true experimental design would generalize to all children and fathers in the population from which the sample was drawn.

Second, the conditional independence assumption that the outcome (father involvement) is independent of treatment status (children with disabilities/delays vs. typically developing) is not fully verifiable. As a result, I can never be sure that I am accurately estimating the effect of the treatment on the outcome. (It should be noted that this limitation likewise applies to the regression/covariate adjustment approach, which also has the additional limitations outlined previously).In spite of these limitations, given that random assignment to disability/delay status is not possible or ethical, propensity score analysis provides the best alternative for estimating the causal effect of child disability/delay status on father involvement and on children's functional abilities.

Measures

Disability/delay status of the child. Two groups were identified: children with disabilities/delays (CWD) and children who were typically developing (CTYP). As noted in the introduction, the logic of propensity score analysis requires restrictions on the types of disabilities/delays that can be used to place children in the disability/delay group. Specifically, this study can use only disabilities/delays for which the underlying health condition can safely be assumed to exist from birth, though symptoms may not be apparent at birth. Furthermore, it must be reasonable to assume that these disabilities/delays continue. While intervention services and other environmental factors may improve children's functional abilities, the underlying health condition is assumed to continue to exist. For example, intervention services may improve behaviors of children with ASD, but the child will continue to have ASD. The criteria used to identify disabilities/delays meeting these assumptions are listed and defined below.

Disabilities or delays whose causes are exclusively or primarily due to genetic or in utero development. Although environmental factors may affect the severity and type of symptoms manifest

across time for these disabilities or delays (e.g., medicine for epilepsy, interventions to promote functioning of children with Down syndrome), the disability/delay and or its cause are present at birth and continue to be exist indefinitely. Disabilities that meet these criteria could include:

1. Disabilities or delays that are exclusively the result of genetic mutation or inheritance.
2. Disabilities or delays that are exclusively the result of in utero development or, rather, the failure to develop appropriately in utero.
3. Disabilities or delays that are exclusively the result of an interaction between genetic mutation or inheritance and development in utero.

Disabilities or delays with primarily prenatal biological causes, but for which limited environmental causes exist (e.g., injury, interaction of genetic risk with illness). Although some environmental factors may exist, the influence of these factors should be limited. The following environmental factors would be considered limited:

1. Environmental factors that affect only a handful of children with these disabilities or delays (i.e., injury or illness may be the cause of the disability or delay for only a few cases, but prenatal biological factors are the cause of the vast majority of cases as in the case of blindness, deafness, and cerebral palsy).
2. Environmental factors that act on genetic predisposition to cause a disability or delay, but are not sufficient of themselves to cause the disability or delay (e.g., a viral infection is thought to interact with a genetic predisposition to cause diabetes).
3. Environmental factors that are a cause of a disability or delay, but are not likely to have sufficient influence to cause the disability or delay in the absence of prenatal biological factors. Although such environmental factors may affect the severity of the disability or delay, the disability or delay would still exist in the absence of these factors. In other words, the cause of the disability or delay is prenatal biological factors, but the severity of the disability/delay can be influenced by environmental factors (e.g., delay in learning to walk or severe cognitive impairment).

Identifying children with disabilities/delays in the sample meeting these criteria. When children were approximately 9 months, 2 years, and 4 years of age, each mother was given a list of disabilities/delays and asked to report for each disability/delay whether a doctor had diagnosed her child with the disability/delay. At 9 months and 2 years, mothers were asked if their child had *ever* been diagnosed with the disability/delay whereas at 4 years mothers reported if the child had been diagnosed *since* turning 2 years old. Additionally, each mother was also given a list of early intervention services and asked to report for each service whether her child was receiving that service to help with his or her disability/delay. At 9 months, a mother was only given the list to complete if she had reported that her

child had one or more of the disabilities/delays on the first list. At 2 years, all mothers reported whether or not their child had received the services listed since the last interview (at approximately 9 months). At 4 years, mothers reported on services received if the child was diagnosed with cerebral palsy, another developmental delay, epilepsy, a heart defect, intellectual disability, Autism or PDD, oppositional defiant disorder, ADHD, problem with non-food allergies, diabetes, or any impairment or health problem that requires the use of special equipment. At 4 years, mothers also reported whether or not the child had an IEP or IFSP if the child was diagnosed with a problem with ability to pay attention or learn, overall activity level, ability to communicate, or use of limbs, hearing, or vision. The list of disabilities/delays meeting the criteria listed above (hereafter "qualifying disabilities/delays") are included in Table 1. Children with qualifying disabilities/delays were placed in the disability/delay group.

It is possible that mothers may have been unsure about whether or not their child had a specific disability/delay, but may have been more comfortable reporting that their child was receiving intervention services. Consequently, services that would be received only for a qualifying disability/delay were used to place children in the disability/delay group. These services are listed in Table 2. Of the 2700 children in the sample, 350 had one or more of the qualifying disabilities/delays and or were receiving services for a qualifying disability/delay.

Table 1.

Disabilities/delays meeting study criteria for inclusion in the disability/delay group.

Disability or Delay	Reported at 9 Months	Reported at 2 Years	Reported at 4 Years
Cleft lip or palate	X	---	---
Down syndrome	X	---	---
Turner's syndrome	X	---	---
Spina bifida	X	---	---
Problem with mobility/using legs	X	---	---
A problem using arms or hands	X	---	---
Delay in learning to walk	---	X	---
Problem with mobility such as Cerebral Palsy	---	X	X
Blindness	X	X	---
Diagnosed problem with vision	---	---	X[a]
Difficulty hearing or deafness	X	X	---
Diagnosed problem with hearing	---	---	X[a]
Heart defect		X	X
Epilepsy	---	X	X
Intellectual disability	---	X	X
Autism or PDD	---	---	X
Diabetes	---	---	X
Oppositional Defiant Disorder (ODD)	---	---	X
ADHD	---	---	X

Note. Children with one or more of these disabilities/delays were included in the disability/delay group.
[a]If have IEP/IFSP

Table 2.

Services for disabilities/delays meeting study criteria for inclusion in the disability/delay group.

Service	Reported at 9 Months	Reported at 2 Years	Reported at 4 Years
Physical therapy	X	X	X[a]
Vision services	X	X	X[a]
Hearing services	X	X	X[a]
Occupational Therapy	---	X	X
Private tutoring or schooling for learning problems?	---	---	X
Instruction in Braille?	---	---	X
Instruction in sign language, cued speech, ASL, or TOCO	---	---	X

Note. Children receiving one or more of these services were included in the disability/delay group regardless of whether they had also been diagnosed with a disability/delay listed in Table 1.
[a]If have IEP/IFSP.

Several children had disabilities/delays that did not meet the study criteria or were receiving services for such disabilities/delays. These disabilities/delays included disabilities/delays that could be solely the result of environmental causes (e.g., failure to thrive, delay in learning to talk). Including children with these disabilities in the typically developing group along with children who have no diagnosed disabilities/delays could bias the analyses. As mentioned before, father involvement or environmental factors correlated with father involvement could be a cause of these disabilities/delays. Thus any estimate of the effect of environmentally-caused disabilities/delays on father involvement could include the effect of earlier father involvement on later father involvement through disability/delay status.

Furthermore, in carrying out propensity score analysis, children with qualifying disabilities/delays are matched with children who have a similar propensity for having a qualifying disability/delay, but do not. By including children with one or more of the excluded disabilities/delays in the typically developing group, they would be possible, and perhaps even likely, matches for children with qualifying disabilities/delays in the disability/delay group. This may bias the results because children with one or more of the excluded disabilities/delays could be thought of as having received the treatment (i.e., having a disability/delay, although not a "qualifying" disability/delay). Just as including individuals who received the treatment in the control group would bias the results of an experimental study, so would matching a child with a qualifying disability/delay to a child with a non-qualifying disability/delay bias the estimates in a propensity score analysis. The best option is to exclude these participants from the sample and note the implications of doing so when interpreting the results. Non-qualifying disabilities/delays and related services are listed in Tables 3 and Table 4. Children with one or more of the non-qualifying disabilities/delays or related services *who did not also have a qualifying disability/delay or related service* were dropped from the sample. Children with at least one qualifying disability/delay or related services were kept in the sample's disability/delay group even if they had one or more non-qualifying disabilities/delays and or related services.

Table 3.

Disabilities/delays not meeting study criteria for inclusion in the disability/delay group.

Disability or Delay	Reported at 9 Months	Reported at 2 Years	Reported at 4 Years
Failure to thrive	X	---	---
Any other types of special needs or limitations	X	---	---
Impairment or health problem requiring use of special equipment and/or limits child's ability to walk, run, or play and has lasted or is expected to last 12 months or longer	X		
Delay in learning to talk	---	X	---
Another developmental delay	---	X	X
Diagnosed problem with ability to pay attention or learn	---	---	X
Diagnosed problem with overall activity	---	---	X
Diagnosed problem with use of limbs	---	---	X
Diagnosed problem with ability to communicate	---	---	X

Note. Children with *only* these disabilities/delays were dropped from the sample.

Table 4.

Services for disabilities/delays not meeting study criteria for inclusion in the disability/delay group.

Service	Reported at 9 Months	Reported at 2 Years	Reported at 4 Years
Physical therapy			X[a]
Vision services			X[a]
Hearing services			X[a]
Social work services	X	X	X
Psychological services	X	X	X
Home visits	X	X	X
Parent support or training	X	X	X
Speech Therapy	---	X	X
Special classes with other children, some or all of whom also had special needs	---	X	X
Unspecified service	---	X[b]	---
IEP or IFSP?	---	---	X[c]

Note. Children receiving only these services were dropped from the sample.
[a]Included in study sample if child had an IEP or IFSP. [b]Mothers reported the child was receiving services, but not one of the services listed. [c]While having an IEP or IFSP was sometimes a requirement for including a disability in the study sample, the presence of an IEP or IFSP alone was not sufficient to place a child in the disability/delay group.

Information was also collected from mothers regarding health problems or other special needs of their children and is listed in Table 5. These health problems or special needs are relatively common or temporary experiences of children and or are not likely to bias the results of the propensity score analyses. Children with these conditions were placed in the typically developing group. After dropping 250 children with one or more of the non-qualifying disabilities/delays or related services from the sample, there were 2150 children in the typically developing group and 350 children in the disability/delay group. Table 6 includes additional information regarding the specific disabilities/delays and services experienced by this group.

Of those 350 children in the disability/delay group, 150 received at least one qualifying service at 9 months, 2 years, and or 4 years. For less than 50 of the 350 children in the disability/delay group, mothers reported that the children were receiving one or more of the qualifying services, but had not been diagnosed with any of the qualifying disabilities/delays. The remaining 300 plus children in the disability/delay group had one or more of the qualifying disabilities/delays with some receiving qualifying services and some not.

Reported timing of diagnosis was also examined by determining when children were first identified (i.e., reported by mothers) as having any qualifying disability/delay or receiving any qualifying services. Between birth and 9 months, 100 children were identified. Between 9 months and 2 years, 150 children were identified. Between 2 years and 4 years, 100 children were identified. It should be noted that when children were first diagnosed does not directly map on to when parents first become aware of their child having a disability, delay, or some unexplained problem. While one can safely assume that parents of a child identified by 9 months knew their child had some underlying health condition by 9 months, one cannot assume that parents of children identified later were not also aware of and concerned about some underlying health condition at 9 months or even earlier. Indeed, parents concerns about missed developmental milestones or other problems preceding diagnosis is in many cases what leads to the child being evaluated. For this reason, the effect of timing of diagnosis on father involvement is not explored in this study.

Some of the disabilities/delays meeting the criteria for inclusion in the sample could be the result of illness or injury, which in turn could be related to father involvement. Parents were not asked if their child's disability/delay was the result of illness or injury and it is unlikely in many cases that the parents would know. Mothers did report at each time point how often their child had been to a doctor, other medical professional, medical clinic, or emergency room since they were born or the previous visit for an injury. They also reported the cause of the injury and whether or not the child was hospitalized for at least one night. Given that injuries are common in childhood, neither the fact that a child was injured nor the cause of the injury are particularly helpful in identifying any children whose disability/delay was solely

the result of an injury. Furthermore, an injury could just as easily be the result of a disability/delay as its cause. Of the information collected, hospitalization is the most useful because any injury serious enough to cause a disability/delay would reasonably have resulted in hospitalization. Only 2.45% of children with disabilities/delays (vs. 1.35% of children who were typically developing) were hospitalized for an injury at least one night at any time between birth and 4 years. For two-thirds of the CWD who were hospitalized for an injury, hospitalization occurred *after* the diagnosis of their disability/delay. Therefore, the disability/delay could not be the result of the injury. Even if an injury caused the disability/delay in the remaining 0.61% of CWD, the risk of introducing bias into the analyses by including them is very, very minimal. There is no way to know with certainty that the CWD group only includes children with disabilities/delays that were present—in some form—from birth. However, the steps described above to include only children with disabilities/delays that meet the study criteria and the data screening just described should minimize any bias in the analyses.

Table 5.

Health problems or other special needs not meeting study criteria for inclusion in the disability/delay group or exclusion from sample.

Disability or Delay	Reported at 9 Months	Reported at 2 Years	Reported at 4 Years
Heart defect	X		
Difficulty seeing	X	X	---
A crossed eye, or a lazy or wandering eye	---	X	---
Difficulty seeing objects in the distance or letters on paper	---	---	X
Child wears glasses	---	---	X
Diagnosed problem with vision	---	---	X[b]
Diagnosed problem with hearing	---	---	X[b]
Any impairment or health problem that requires the use of special equipment, such as a brace, a wheelchair, a hearing aid, or corrective shoes? Do not include ordinary eyeglasses.	---	X[a]	X
Any impairment or health problem that limits child's ability to walk, run, or play?	---	X[a]	---
Lactose intolerance	---	X	X
Food allergies or sensitivities such as to peanuts	---	X	X
Problem with non-food allergies, such as to dust, animals, or medicine	---	---	X
Difficulty hearing and understanding speech in a normal conversation	---	---	X

Note. Children with only these problems or special needs were included in the typically developing group. [a]If impairment or health problem was expected to last 12 months or longer, then the child was dropped from the study sample. [b]If the child had an IEP, then he or she was placed in the typically developing group.

Table 6.

Disabilities/Delays and Services of the Disability/Delay (CWD) group.

Disabilities/Delays and Services	Reported at			% of CWD with *only* this disability/ delay or receiving *only* this service	% of CWD[c]
	9M	2Y	4Y		
Cleft lip or palate	X			0.31%	1.84%
Down syndrome	X			0.00%	1.53%
Turner's syndrome	X			0.00%	0.61%
Spina Bifida	X			0.31%	0.92%
A problem using arms or hands	X			0.31%	6.44%
Blindness	X	X		0.31%	3.07%
Difficulty hearing or deafness	X	X		4.91%	11.66%
Problem with mobility/using legs, C. palsy	X	X	X	1.53%	14.11%
Delay in learning to walk		X		4.91%	17.48%
Heart defect		X	X	16.26%	23.93%
Cognitive Disability		X	X	0.31%	3.68%
Epilepsy		X	X	7.36%	12.88%
Diabetes			X	1.23%	1.23%
Autism or PDD			X	0.92%	4.91%
Oppositional Defiant Disorder (ODD)			X	0.00%	1.84%
ADHD			X	0.61%	3.07%
Diagnosed problem with hearing			X[a]	2.45%	10.74%
Diagnosed problem with vision			X[a]	0.31%	8.28%
Physical therapy	X	X	X[a]	3.07%	30.67%
Vision services	X	X	X[a]	3.68%	21.17%
Hearing services	X	X	X[a]	0.92%	12.58%
Occupational therapy		X	X	1.23%	25.15%
Private tutoring or schooling for learning			X	1.23%	7.36%
Instruction in sign language, cued speech, ASL, or TOCO			X	0.00%	2.15%
Total *N*/% of children with disabilities/ delays meeting only 1 eligibility criteria[b]	---	---	---	150 / 52.15%	---
Total *N*/% of children with disabilities/ delays meeting > 1 eligibility criteria[b]	X	X	X	150 / 47.85%	---
Total *N* of children with disabilities/ delays and or receiving EI services	X	X	X	350	350

[a]If have IEP/IFSP. [b]Eligibility criteria refer to qualifying disabilities/delays or related services. [c]Percent of CWD with disability/delay or service alone *or* with other disabilities/delays or services.

Functional Abilities Index (FAI). Building upon a functional domains approach advocated by Buysse (e.g., Buysse et al., 1993), Daley (Daley et al., 2009), and Rosenberg (e.g., Rosenberg, Zhang & Robinson, 2010), the ECLS-B Functional Abilities Index (FAI) was developed as a post-hoc measure of children's general functioning within and across multiple developmental domains (McBride et al., 2013). The FAI consists of 8 domains of child functioning: health, hearing, vision, cognitive functioning, social skills, motor skills, communication skills, and self-regulation. Each domain has a possible score of 0, 0.5, or 1 with 0 indicating significant delay or limitation within that domain, 0.5 indicating a moderate delay or limitation, and 1 indicating typical functioning within the domain. Scores for each domain were computed for each of the first three waves of data collection (9 months, 2 years, and 4 years). Sources and scoring for each domain are detailed below. Included in Table 7 is a summary of the sources used to create the Functional Abilities Index and how scoring was calculated for each domain.

In general, for domains based on categorical measures (health, hearing, and vision) it was evident which categories represented delay or limitations. However, for domains based on continuous measures, there were no established cutoff scores for classifying a child's ability as limited or delayed. Previous post hoc indices of children's functional abilities (e.g., Rosenberg et al., 2010) used cutoffs based on means and standard deviations. Specifically, cutoffs scores identified children scoring 1 or 1.5 SD below the mean on measures of children's functional abilities. Several states also use means and standard deviations as (part of) their criteria for identifying children who are eligible for Part C. However, a limitation of the use of means and standard deviations is that measures of children's functional abilities may not be normally distributed and standard deviations may be larger or smaller for measures with skewed distributions. Consequently, cutoff scores identified based on these standard deviations will inconsistently identify a larger or smaller portion of the sample as delayed depending on the level and direction of the skewness. To avoid this limitation, percentiles were used in the present study in place of standard deviations. Specifically, we used as cutoffs the 15.87^{th} percentile and the 6.68^{th} percentile, which correspond to the percentage of individuals who are at or below 1 or 1.5 SD, respectively, below the mean of a normal distribution. It should be noted that the percentiles calculated for creating the FAI domains were estimated using the data weights and 8000 children from the full ECLS-B sample, including both children with disabilities/delays and typically developing children. These 8000 children were selected for inclusion in establishing cutoff scores because they represent the largest number of children in the ECLS-B who had the appropriate weights available. The weights used for creating the FAI (W3C0) were designed to be used in analyses examining children's scores on direct assessments (such as those used to create the FAI) at 9 months, 2 years, and 4 years. All 2500 children in the smaller sample used in this study had the W3C0 weight available (in addition to the W3D0 weight used in this study) and thus had a

FAI score. By using all 8000 children and their accompanying weights, the cutoffs used to create the FAI were based on a nationally representative sample.

At each wave, children's scores from each of the 8 domains were summed to form a total score with a possible range of 0 – 8. Higher FAI total scores indicate high functional abilities with no or few limitations, while lower FAI scores indicate severe developmental concerns that cut across multiple domains of functioning. Mean, standard deviation, and range of the FAI total score at 9 months, 2 years, and 4 years for each group is given in Table 8.

Health. At each time point, mothers reported on the child's health using 5 categories. Specifically, mothers reported if the child's health was excellent or very good (= 1), good (= 0.5), or fair or poor (= 0).

Hearing. At each time point, mothers reported whether their child experienced deafness, difficulty hearing or understanding conversation, and/or was receiving hearing services. Children who experienced at least one of these were given a score of 0, while all others were given a score of 1 for the *hearing index variable* at each time point.

Vision. Mothers reported at each time point whether their child was blind, had difficulty seeing, or was receiving vision services. Children who experienced one or more of these were given a score of 0 for the *vision index variable* at each time point. All other children were given a score of 1.

Cognitive functioning. At 9 months and 2 years, children were assessed using the Bayley Short Form-Research Edition (BSF-R; The Psychological Corporation, 2001), a shortened, equated (Andreassen & Fletcher, 2007) version of the Bayley Scales of Infant Development, 2nd Edition (BSID-II; Bayley, 1993). As part of the BSF-R assessment at 9 months and 2 years, children's cognitive functioning was measured. Children's percentile scores on the BSF-R mental scale were calculated. Children who scored at or below the 15.87th percentile were given a score of 0 for *cognitive functioning index variable*. All other children were given a score of 1.

At 4 years, children were administered a standardized math and reading assessment designed for use with the ECLS-B (Najarian, Snow, Lennon, & Kinsey, 2010). The math assessment examined such skills as children's number sense, counting, operations, etc. The reading assessment examined children's letter recognition, letter sound knowledge, receptive and expressive vocabulary, etc. Because the math and reading assessments were highly correlated (r = .77), the two variables were standardized, centered, and averaged to form a composite cognitive functioning score. Children who scored at or below the 15.87th percentile on the composite variable were given a score of 0 for the *cognitive functioning index variable* at 4 years. All other children were given a score of 1.

Table 7.

Functional Abilities Index: Sources and Scoring.

Domain	Source	Scoring
Health	Time 1 – 3 • Parent report of child health	• "Fair" or "Poor" = 0, "Good" = .5 • "Very Good" or "Excellent" = 1
Hearing	Time 1 – 3 • Parent report of hearing difficulties	• Deafness, difficulty hearing or understanding conversation, and/or receiving hearing services = 0 • Otherwise = 1
Vision	Time 1 – 3 • Parent report of vision difficulties	• Difficulty seeing or blindness or receiving vision services = 0 • Otherwise = 1
Cognitive Functioning	Time 1 & 2 • BSF-R Mental Developmental Index Score	• If scored ≤ 15.87th percentile, then T1/T2 Cognitive Functioning = 0 • Otherwise = 1
	Time 3 • Composite of math & reading assessments ($\alpha = 0.87$)	• If scored ≤ 15.87th percentile, then T3 Cognitive Functioning = 0 • Otherwise = 1
Social Skills	Time 1 • Videotaped parent-child teaching interaction coded using NCATS—TCS (i.e., ability to communicate with and respond adaptively to parent's cues)	• If scored ≤ 15.87th percentile, then T1 Social skills score = 0 • Otherwise = 1
	Time 2 - 3 • Videotaped parent-child structured interaction during The Two Bags Task. • Coded for child's engagement of parent and child's negativity toward parent (reverse coded)	• If scored ≤ 15.87th percentile on both the engage score AND the negativty score, then T2/T3 Social Skills = 0. • If scored ≤ 6.68th percentile on EITHER the engage scale OR the negativity scale, then T2/T3 Social Skills = 0. • Otherwise = 1

Note. BSF-R = Bayley Short Form-Research Edition. NCATS = Nursing Child Assessment Teaching Scale—Total Child Score

Table 7 (continued).

Functional Abilities Index: Sources and Scoring.

Domain	Source	Scoring
Motor Skills	Time 1 & 2 • BSF-R Motor Developmental Index Score	• If scored ≤ 15.87th percentile, then T1/T2 Motor Skills = 0 • Otherwise = 1

	Time 3
	• Fine Motor assessment
	  ○ Copying forms task
	  ○ Building blocks task
	• Gross Motor assessment
	  ○ 7 tasks: walking backwards, catching, jumping, balancing and hopping on left and right foot
		• If scored ≤ 15.87th percentile on the gross motor composite score AND ≤ 15.87th percentile on EITHER the copying forms score OR the building blocks score, then T3 Motor Skills = 0.
		• If scored ≤ 6.68th percentile the gross motor composite score, the forms composite score, OR the blocks composite score, then T3 Motor Skills = 0.
		• Otherwise = 1

Communication Skills	Time 1 • BSF-R: Proficiency probability of nonverbal communication	• If scored ≤ 15.87th percentile, then T1 Communication Skills = 0 • Otherwise = 1
	Time 2 • Composite variable of parent report of children's vocabulary, syntactic, and morphological language development (α = 0.88)	• If scored ≤ 15.87th percentile, then T2 Communication Skills = 0 • Otherwise = 1
	Time 3 • Composite variable of direct assessment of children's expressive language & parent report of children's vocabulary & communication skills (α = 0.71)	• If scored ≤ 15.87th percentile, then T3 Communication Skills = 0 • Otherwise = 1
Self-Regulation	Time 1 – 3 • Latent variables of parent report of child temper tantrums, delay of gratification, fussiness, etc.	• If scored ≤ 15.87th percentile, then T1/T2/T3 Self-Regulation = 0 • Otherwise = 1

Note. BSF-R = Bayley Short Form-Research Edition

Table 8.

FAI Descriptive Statistics.

FAI	Mean	SD	Possible Range	Actual Range
9 Months: CWD	6.74	1.29	0 – 8	2 – 8
9 Months: CTYP	7.24	0.91	0 – 8	3 – 8
2 Years: CWD	6.69	1.50	0 – 8	2 – 8
2 Years: CTYP	7.38	0.88	0 – 8	2.5 – 8
4 Years: CWD	6.31	1.76	0 – 8	0 – 8
4 Years: CTYP	7.29	1.00	0 – 8	2 – 8

Social skills. At 9 months, mother and child were videotaped while engaged in a parent-child teaching task. Interactions were videotaped and coded using the Nursing Child Assessment Teaching Scale (NCATS; Sumner & Spietz, 1994). The NCATS total child score reflects the child's ability to communicate with his or her parent and respond adaptively to the caregiver's cues. Children who scored at or below the 15.87th percentile on the total child score were assigned a score of 0 for the *social skills index variable*. All other children were assigned a score of 1.

At 2 and 4 years, mothers and children were videotaped while they engaged in the two-bag task, which consisted of pretend play and joint book reading tasks during the Two Bag Task (5 minutes each). The Two Bag Task is a shortened version of the Three Bag Task used in both the Early Head Start Research and Evaluation Project and the Study of Early Childcare (see Nord, Edwards, Andreassen, Green, & Wallner-Allen, 2006). Children were rated on their engagement of the parent (initiation and maintenance of interaction with parent, expressions of positive affect) and negative affect expressed toward the parent (expressions of anger, hostility, or dislike directed toward the parent). These scales were adapted from Fauth, Brady-Smith, and Brooks-Gunn (2003). Since engagement and negative affect were only moderately correlated, the scales were not combined (r = .20 and .28 at 2 and 4 years, respectively). Instead, both variables were used to create the *social skills index variable* as described below. Children's percentile scores were calculated for the engagement and negativity scales. The negativity scale was reverse coded before calculating percentiles so that high values represented better functioning (less negativity). Children received a 0 for the *social skills index variable* at 2 years or 4 years if they scored at or below 15.87th percentile on both the engage score and the negativity score. Children also received a 0 if they scored at or below the 6.68th percentile on either the engage score or the negativity score. All other children were assigned a score of 1.

The rationale for this scoring criterion for two moderately correlated, continuous measures is to avoid identifying individuals with moderate delays or limitations (i.e., $\leq$ 15.87th percentile, but not $\leq$ 6.68th percentile) as experiencing significant limitations in a domain unless the delay or limitation is present in more than one measure or a very significant delay (e.g., $\leq$ 6.68th percentile) is present in either. Similar criterion has been used with previous research on children with disabilities using the ECLS-B (e.g., Rosenberg et al., 2010).

Motor skills. As part of the BSF-R assessment at 9 months and 2 years, children's motor skills were assessed. Children's motor scores reflect their fine and gross motor ability. Children's percentile scores on the BSF-R motor scale were calculated. Children who scored at or below the 15.87th percentile were given a score of 0 for the *motor skills index variable*. All other children were given a score of 1.

At 4 years, children's performance on activities that required fine motor skills (copying forms and stacking blocks) was assessed. Children's percentile scores on the copying forms and stacking blocks

tasks were calculated. Children were also assessed on their ability to perform 7 activities requiring gross motor skills: walking backwards, catching, jumping, and balancing and hopping on their left and right foot. The variables for the 7 activities were standardized and centered. Factor analysis indicated the presence of a single construct, so the 7 standardized variables were averaged to form a gross motor skills composite variable ($\alpha = 0.74$). Children's percentile scores on the gross motor skills composite variable were calculated. Children's scores on the copying forms and stacking blocks tasks were only moderately correlated with one another ($r = .27$) and with gross motor skills composite variable ($r = .32$ and .25, respectively). Using this information, children were given a score of 0 on the T3 *motor skills index variable* if they scored at or below the 15.87th percentile on the gross motor composite score and at or below the 15.87th percentile on either the copying forms score OR the building blocks score. Children were also given a score of 0 if they scored at or below the 6.68th percentile on the gross motor composite variable, the forms variable, or the blocks variable. All other children were given a score of 1.

Communication skills. As part of the BSF-R assessment at 9 months, children's proficiency probabilities for different skills, including different communication skills were calculated. Proficiency probabilities represent the level of mastery of a specific skill. The scores range from 0 to 1 and indicate the likelihood that a child had reached a key milestone. Communication-related skills were non-verbal communication, naming objects, receptive and expressive vocabulary, and listening comprehension. At 9 months, very few children demonstrated any proficiency of these skills (e.g., less than 10% of children demonstrated more than 1% proficiency for several skills) with the exception of non-verbal communication. Consequently, only non-verbal communication was used to create the *communication skills index variable* at 9 months. The non-verbal communication proficiency assessed children's ability to communicate through diverse nonverbal sounds and gestures (e.g., vowel sounds, gesturing for an object, and jabbering expressively). Children's percentile scores on the non-verbal communication proficiency variable were calculated. Children who scored at or below the 15.87th percentile were given a score of 0 for the *communication skills index variable* at 9 months. All other children were given a score of 1.

At 2 years, mothers were given a list of 50 words and asked to indicate if the child could say the word in English (or in the family's language if other than English) or another word that means the same thing. For each word, children were given a score of 0 if they could not say the word and a score of 1 if they could. A vocabulary score was created by averaging children's scores across the 50 words ($\alpha = 0.95$). Mothers also reported on children's syntactic language development (whether the child combines words and how he or she usually communicates: 1 word sentence, 2-3 word sentences, short sentences, long sentences) and morphological language development (plurals, ownership, past tense, and present progressive). If children could not combine words, then mothers were not asked the other items about

syntactic and morphological language development and were missing data on these variables. Because these children were assumed to not have developed these skills, the missing values were recoded to indicate the child did not have the skill. All of these measures of children's communication development were selected from the MacArthur Communicative Development Inventory (M-CDI; Fenson et al., 1994) by its authors, but do not represent the M-CDI. The four morphological language development items were combined into a single morphological development score ($\alpha = 0.77$). Factor analysis of the vocabulary score, morphological language development score, and the two syntactic language development items indicated the presence of a single factor. These 4 variables were standardized, centered, and average to form a communication ability composite score at 2 years ($\alpha = 0.88$). Children's percentile scores on the composite variable were calculated. Children who scored at or below the 15.87^{th} percentile were given a score of 0 for the *communication skills index variable* at 2 years. All other children were given a score of 1.

At 4 years, mothers again reported on children's vocabulary development using a list of 25 words developed by the authors of the M-CDI. Because the list was included in a different survey instrument than at 2 years, only mothers who spoke English or Spanish completed the list. A vocabulary score was created in the same way as at 2 years ($\alpha = 0.77$). Mothers also reported on children's general communication skills using measures based on Leventhal (1998), including children's ability to speak clearly so a stranger can understand, to listen well, to wait his/her turn to speak, etc. Finally, children's expressive language ability was directly assessed using the Let's Tell Stories subtest of the PreLAS 2000 (Duncan & De Avila 1998). During the subtest, children were read two short stories. After each story, they were asked to retell the story using picture prompts. Recordings of their retelling were coded for expressive language ability. Only children who could speak at least some English's were administered the Let's Tell Stories subtest. Factor analysis of the vocabulary score, expressive language ability, and the 6 general communication skills indicated the presence of a single factor. These 8 variables were standardized, centered, and averaged to form a communication ability composite variable at 4 years ($\alpha = 0.71$). Children's percentile scores on the composite variable were calculated. Children who scored at or below the 15.87^{th} percentile were given a score of 0 for the *communication skills index variable* at 4 years. All other children were given a score of 1.

Self-regulation. At each time point, mothers reported on their child's self-regulation abilities. At 9 months and 2 years, items from the Infant/Toddler Symptoms Checklist (ITSC; DeGangi, Poisson, Sickel, & Wiener, 1995) were used (e.g., child fussiness, delay of gratification, sleep difficulties). At 4 years, mothers reported on items drawn from the Preschool and Kindergarten Behavior Scales–Second Edition (PKBS-2; Merrell, 2003) and from the Social Skills Rating System (SSRS; Gresham & Elliot, 1990) (e.g., child is impulsive, overactive, has temper tantrums). These indicators were used to form

latent variables. One latent variable was created for *self-regulation* at each time point. The factor scores for these latent variables were created simultaneously within the same model along with a latent variable for *communication* that was ultimately not used to create the FAI. Fitting multiple latent variables in a single model is preferred to fitting a separate model for each latent variable when possible because the information used to create one latent variable also informs the creation of the other latent variables. Model fit for the measurement model in which the 4 latent variables were created was good: RMSEA = 0.020, CFI = 0.949, SRMR = 0.026, $\chi^2_{142} = 592.85$. Factor scores from the latent variables were saved and included with the other variables in the data. Children's percentile scores on the self-regulation factor scores were calculated. Children who scored at or below the 15.87^{th} percentile were given a score of 0 for the *self-regulation index variable* at each time point. All other children were given a score of 1.

Accounting for variation in age of assessment. In the ECLS-B, children were not assessed exactly at 9 months, 2 years, or 4 years of age and, consequently, age might influence children's scores in 5 domains: motor skills, cognitive functioning, self-regulation, social skills, and communication skills. Therefore, before creating index scores for these domains, residual scores were generated by fitting regression models in which performance on relevant variables was predicted by children's age of assessment (i.e., age in months at the time of the assessment). Doing so removed from these variables the variance explained by age. The residual scores were then used to create the functional abilities index scores. However, the motor and mental scale scores from the BSF-R used to create the 9 month and 2 year motor skills and cognitive functioning index scores were already age-normed (Andreassen & Fletcher, 2007) and thus did not need to have the variance explained by age removed.

Father involvement. Fathers were asked to report on how frequently they engaged in several different activities with their children. Specifically, fathers reported on routine caregiving involvement, play involvement, literacy involvement, and responsive caregiving involvement. Routine caregiving involvement refers to routine activities fathers engage in to provide care for their child's functional needs. Play involvement refers to play activities fathers engage in with their child, including sedentary and activity play. Literacy involvement refers to fathers' engagement in reading or other language activities with their children. Responsive caregiving involvement measures how often fathers were the parent to respond to the less routine caregiving needs of their child. Fathers were asked, "When the following things happen or need to be done, how often are you the one who does them?" regarding such tasks as "taking their child to the doctor" and "soothing their child when upset." Responsive caregiving differs from the other types of involvement in two ways. First, responsive caregiving is not an "absolute" measure of father involvement in that it does not measure the frequency a father engages in specific caregiving activities. Rather, it measures how often fathers responded to the needs of their child relative to the mother. Second, responsive caregiving measures how often fathers engaged in these activities when these activities needed to be done. In contrast to the routine caregiving involvement items, these activities tend to consist of responding to less routine needs of children. The items for each involvement type are given in Table 9.

Rescaling of father involvement. Response categories were based on 4 point or 6 point scales ranging from never to every day or more than once a day. In the original metric, a one point difference may represent the difference between once a week and 2 or 3 times a week or between 7 times a week and 14 or more times per week. Thus, the original metric is not equal-interval. To address this limitation, father involvement items were rescaled to represent the approximate number of times per week the father engaged in the activity (see Table 9). Although this approach is not perfect given that number of times per week has to be approximated (e.g., it is assumed that "few" means twice a month and "more than once a day" means twice a day), rescaling results in a more meaningful scale that is closer to reality. A similar approach to rescaling can be found in previous research using ECLS-B data with items that use these same response categories (e.g., Dyer et al., 2009).

Accounting for variation in age of assessment. As was the case with the measures used to create the FAI, data on father involvement were not collected exactly at 9 months, 2 years, or 4 years. The procedure used to remove bias in the FAI measures due to variation in assessment age was also used to remove the bias in father involvement variables. This was done prior to the creation of the father involvement latent variables described next.

Table 9.

Measures of Father Involvement.

	In the past month, how often did you do the following things with your child? Was it: More than once a day (14)[a], About once a day (7), A few times a week (4.5), A few times a month (1), Rarely (0.25), or Not at all (0)? Other responses: Refused and Don't Know.		
	Father 9 Months	Father 2 Years	Father Preschool
Routine Caregiving Involvement	Change your child's diaper?	Change your child's diapers or help your child use the toilet?	
	Feed your child or give your child a bottle?	Assist your child with eating?	
	Prepare meals or bottles for your child?	Prepare meals for your child?	Prepare meals for your child?
	Dress your child?	Help your child get dressed?	Help your child dress him/herself?
	Wash or bathe your child?	Give your child a bath?	Help your child bathe him/herself?
	Put your child to sleep?	Help your child to bed?	Help your child to bed?
		Help your child brush his or her teeth?	Help your child brush his/her teeth?
Play Involvement	Take your child outside for a walk or to play in the yard, a park, or a playground?	Take your child outside for a walk or to play in the yard, a park, or a playground?	Take him/her outside for a walk or to play in the yard, a park, or a playground?
	Play peek-a-boo with your child?		
	Do things like tickle your child, blow on his/her belly, or move his/her arms and legs around in a playful way?		
		Play chasing games with your child?	
		Take your child for a ride on your shoulders or back?	
		Play with games or toys indoors with your child?	Play together with toys for building things like blocks, tinker toys, Lincoln logs, or Legos?

[a]Numbers in parentheses are the rescaled values representing approximate weekly occurrence.

Table 9 (continued).

Measures of Father Involvement.

	Father 9 Months	Father 2 Years	Father Preschool
Literacy Involvement	*In a typical week, how often do you do the following things with your child? Would you say: Not at all (0)[a], Once or twice (1.5), 3 to 6 times (4.5), or Every day (7)? Other responses: Refused and Don't Know.*		
	Read books to your child?	Read books to your child?	Read books to your child?
	Tell stories to your child?	Tell stories to your child?	Tell stories to your child?
	Sing songs with your child?	Sing songs with your child?	Sing songs with your child?
Responsive Caregiving Involvement	*When the following things happen or need to be done, how often are you the one who does them? Do you Always (4), Often (3), Sometimes (2), Rarely (1), or Never (0) do them? Other responses: Refused and Don't Know.*		
	Take your child to the doctor?		
	Get up with your child when he/she wakes up during the night?		
	Soothe your child when he/she is upset?	Soothe your child when he/she is upset?	
	Stay home to care for your child when he/she is ill?	Stay home to care for your child when he/she is ill?	

[a]Numbers in parentheses are the rescaled values representing approximate weekly occurrence for literacy involvement.

Creation of father involvement latent variables and measurement invariance testing. For each type of father involvement, latent variables of 9 month, 2 year, and 4 year father involvement were created in the same model using Mplus 7.11. As noted previously, it is often useful to create latent variables in a single model (i.e., fit one model to create latent variables for *all 4* types of father involvement) so that the information in the items used to create one latent variable can also inform the creation of others. However, the number of parameters that must be estimated grows exponentially as more variables (observed and latent) are added. In the present study, model complexity is further increased because the latent variables were created for both children with disabilities/delays and children who were typically developing using multiple group analysis. To reduce model complexity, latent variables were fit for each type of father involvement separately. However, for each type of father involvement, latent variables for involvement at 9 months, 2 years, and 4 years were created in the same model.

Multiple group analysis was used to create the father involvement latent variables for both groups in the same model. An advantage of multiple group analysis is that it allows researchers to test for measurement invariance. Conceptually, measurement invariance refers to whether items represent the same underlying latent variable for different groups (e.g., CWD and CTYP). "Strong" or "strict" measurement invariance is a necessary assumption for making direct comparisons between groups (Vandenberg & Lance, 2000; Widaman & Reise, 1997). In assessing measurement invariance, I followed Meredith's (1993) outline for testing factorial invariance and tested for configural (items load on the same factors across groups), weak factorial (factor loadings equivalent across groups), strong factorial (intercepts of indicator variables also equivalent), and strict factorial invariance (residual variances of indicator variables also equivalent). Configural factorial invariance is a prerequisite for weak factorial invariance, weak for strong, and strong for strict.

Strict invariance across groups held for routine caregiving, literacy, play, and responsive caregiving involvement within each time point. These measures of father involvement appear to mean the same thing for both fathers of children with disabilities/delays and fathers of typically developing children. Consequently, comparison of levels of father involvement between groups can be made. Because the same number of identical items was used for literacy involvement at each time point, I also tested for and found that strict invariance held across time for literacy involvement meaning that literacy involvement was the same construct across time as well as across groups within a time point.

Using the results of the measurement invariance testing, models were fit for each type of father involvement assuming strict invariance across groups (and across time for literacy involvement). Model fit for routine caregiving involvement was good: RMSEA = 0.035, CFI = 0.951, SRMR = 0.045, and χ^2_{300} = 746.60, p = .000. Model fit for play involvement was good: RMSEA = 0.023, CFI = 0.978, SRMR =

0.038, and $\chi^2_{63} = 104.25$, $p = .001$. Model fit for literacy involvement was good: RMSEA = 0.028, CFI = 0.978, SRMR = 0.039, and $\chi^2_{71} = 139.24$, $p = .000$. Model fit for responsive caregiving involvement was good: RMSEA = 0.021, CFI = 0.991, SRMR = 0.028, and $\chi^2_{26} = 40.02$, $p = .039$. Mean and standard deviation of father involvement latent variables are given in Table 10. Although structural equation modeling was used to create the latent variables of father involvement, propensity score analysis in SEM is a largely unexplored country. However, latent variables are still preferred to observed variables because they eliminate measurement error (Bollen, 1998). Consequently, the latent variables were saved as factor scores to be used in subsequent analyses. Correlations between father involvement and children's functional abilities for children who are typically developing and children with disabilities/delays are given in Table 11.

Table 10.

Father Involvement Descriptive Statistics

Father Involvement	CWD		CTYP	
	Mean	SD	Mean[a]	SD
Routine Caregiving Involvement 9 Months	-0.21	2.03	0.00	1.90
Routine Caregiving Involvement 2 Years	-0.27	2.10	0.00	2.12
Routine Caregiving Involvement 4 Years	-0.28	1.98	0.00	2.02
Play Involvement 9 Months	-0.26	1.23	0.00	1.14
Play Involvement 2 Years	-0.19	1.68	0.00	1.74
Play Involvement 4 Years	0.10	1.41	0.00	1.41
Literacy Involvement 9 Months	-0.02	1.17	0.00	1.06
Literacy Involvement 2 Years	-0.17	1.22	0.02	1.07
Literacy Involvement 4 Years	0.01	1.05	-0.01	1.01
Responsive Caregiving 9 Months	-0.02	0.29	0.00	0.30
Responsive Caregiving 2 Years	-0.01	0.35	0.00	0.34

Note. Only comparisons between groups within the same time point are valid. With the exception of literacy involvement, each involvement variable represents a slightly different construct at each time point making comparisons across time invalid.

[a]For the model to be identified, the means of the CTYP group latent variables were constrained to 0 at all time points for routine caregiving, play, and responsive caregiving involvement. For literacy involvement, only the mean of the 9 month latent variable for the CTYP group was constrained to 0.

Table 11.

Correlations between father involvement and children functional abilities for CTYP and CWD.

	CTYP	1	2	3	4	5	6	7	8	9	10	11	12	13
1	F. Literacy Inv. (9M)	1.00												
2	F. Literacy Inv. (2Y)	.72	1.00											
3	F. Literacy Inv. (4Y)	.56	.72	1.00										
4	F. Routine Caregiving Inv. (9M)	.24	.22	.21	1.00									
5	F. Routine Caregiving Inv. (2Y)	.24	.32	.27	.70	1.00								
6	F. Routine Caregiving Inv. (4Y)	.18	.22	.31	.52	.60	1.00							
7	F. Play Inv. (9M)	.35	.32	.30	.53	.47	.33	1.00						
8	F. Play Inv. (2Y)	.32	.37	.32	.42	.62	.38	.68	1.00					
9	F. Play Inv. (4Y)	.27	.31	.36	.33	.42	.51	.53	.63	1.00				
10	F. Responsive Caregiving Inv. (9M)	.23	.16	.12	.49	.41	.30	.34	.29	.25	1.00			
11	F. Responsive Caregiving Inv. (2Y)	.18	.16	.13	.41	.48	.30	.29	.34	.26	.79	1.00		
12	Child FAI (9M)	.03	.04	.02	.01	-.01	-.02	.04	.02	-.01	-.03	-.03	1.00	
13	Child FAI (2Y)	.06	.11	.05	-.06	-.02	-.04	-.02	.00	-.04	-.09	-.10	.21	1.00
14	Child FAI (4Y)	.04	.10	.06	-.04	-.07	-.07	-.04	-.08	-.06	-.07	-.11	.15	.33

Note. CTYP = Children who are typically developing.

Table 11 (continued).

Correlations between father involvement and children functional abilities for CTYP and CWD.

	CWD	1	2	3	4	5	6	7	8	9	10	11	12	13
1	F. Literacy Inv. (9M)	1.00												
2	F. Literacy Inv. (2Y)	.74	1.00											
3	F. Literacy Inv. (4Y)	.52	.73	1.00										
4	F. Routine Caregiving Inv. (9M)	.22	.22	.23	1.00									
5	F. Routine Caregiving Inv. (2Y)	.21	.29	.26	.70	1.00								
6	F. Routine Caregiving Inv. (4Y)	.21	.22	.37	.56	.61	1.00							
7	F. Play Inv. (9M)	.30	.36	.31	.44	.36	.26	1.00						
8	F. Play Inv. (2Y)	.32	.42	.37	.34	.54	.36	.62	1.00					
9	F. Play Inv. (4Y)	.32	.33	.41	.28	.38	.51	.43	.65	1.00				
10	F. Responsive Caregiving Inv. (9M)	.18	.14	.21	.51	.42	.33	.32	.25	.22	1.00			
11	F. Responsive Caregiving Inv. (2Y)	.13	.13	.18	.38	.50	.36	.28	.36	.28	.73	1.00		
12	Child FAI (9M)	-.03	-.01	-.05	-.12	-.09	-.09	-.09	-.07	-.04	-.11	-.02	1.00	
13	Child FAI (2Y)	.07	.15	.08	-.06	-.12	-.13	-.06	-.09	-.04	-.15	-.11	.39	1.00
14	Child FAI (4Y)	.04	.10	.08	-.06	-.06	-.04	-.12	-.08	-.09	-.12	-.10	.30	.60

Note. CWD = Children with disabilities/delays

Predictors of disability/delay status. Covariates that predict child disability/delay status, children's functional abilities, and father involvement were used to generate the propensity score. These covariates must be "pre-treatment" variables, meaning that they must precede the child having a disability/delay; otherwise, these variables could be influenced by the disability/delay itself and thus would not be appropriate for use in predicting a child's likelihood of having a disability/delay. Many disabilities are not diagnosable until later in early childhood (e.g., Autism). However, the symptoms may be present from very early in the child's life, even before they can be recognized as such by medical professionals or other practitioners. Consequently, "pretreatment" is operationalized as prior to the child's birth. In the ECLS-B, information about mothers' and fathers' pre-birth characteristics was obtained from the child's birth certificate and other hospital information. Additional data were not collected before children were approximately 9 months of age. However, several retrospective measures are available for assessing pretreatment risk for disability/delay status, such as mother's report of preconception weight or cigarette use during pregnancy. Additionally, several variables are time-invariant, such as the sex and race/ethnicity of the child. Finally, if variables were not measured prenatally and are not retrospective, but demonstrate significant stability after birth, then they may be adequate proxies of the prenatal measures not available in the data set.

Recommendations of the types of covariates should be included in the propensity score model have varied over time (Hosmer, Lemeshow, & Sturdivant, 2013). Using simulations, Austin, Grootendorsi, and Anderson (2007) found that propensity score models formed from covariates predicting treatment assignment and covariates confounding the effect of treatment assignment on the outcome performed the best. Included below is a list of covariates available in the ECLS-B data set that may be associated with the likelihood of a child having a disability. These variables may also be theoretically associated with the outcome variables in this study (children's functional abilities and father involvement), though the association may be through yet another variable.

The list also includes covariates for which there is not necessarily strong theoretical or empirical evidence that the covariates are related to disability/delay status, functional abilities, or father involvement. One of the advantages of propensity score analysis over the regression/covariate approach is that the researcher is not concerned about multicollinearity or loss of power in creating the propensity score. Thus, the researcher has more flexibility to include covariates in the model that may only be weakly associated with the treatment or outcome or for which an association has only been speculated. This is particularly important given that research identifying the risk factors for having a child with a disability/delay is ongoing and the hypothesized association between many factors and disability/delay status has yet to be empirically demonstrated. Research into the antecedents of father involvement is likewise an ongoing process. Therefore, for many of the factors included below, there may not be strong

empirical justification for expecting that they be associated with disability status, children's functional abilities, or father involvement. However, if a reasonable argument can be made for their inclusion as possible predictor of treatment status or the outcomes, there is no harm in including them. It should be noted that many, if not all, of the predictors of disability status are also likely predictors of children's functional abilities. A summary of the predictors in provided in Table 12.

Table 12.

List of predictors of disability/delay status.

Variable	Scoring
Demographics	
Child race/ethnicity	Binary indicator variables: African American, Latino(a), Pacific Islander or Asian, Native American or Alaskan Native, Multi-racial/multiethnic. Comparison group = Caucasian, non-Latino(a)
Sex of child	Female (1) Male (0)
Mothers' age	Continuous variable
Fathers' age	Continuous variable
Parents' highest level of education	Binary indicator variables: 12th grade or below, High school diploma or equivalent, Vocational/technological program or some college, Graduate or professional schooling/degree. Comparison group = Bachelor's degree.
Others in home with special need, delay, or disability	Yes (1) No (0)
Mother was employed during the 12 months prior to birth	Yes (1) No (0)
Mother was born outside of the U.S. (foreign)	Yes (1) No (0)
Maternal History	
No previous births.	No previous births (1) Previous births (0)
No prior terminations (miscarriages or abortions)	No prior terminations (1) Prior terminations (0)
Quality of Prenatal Care	
Prenatal care information to mom	Continuous composite variable
Adequacy of Prenatal Care Utilization	Binary indicator variables: Inadequate, Intermediate adequacy, Adequate Plus. Comparison group = Adequate.
Vitamin use 3 months before conception.	Yes (1) No (0)
Vitamin use 3 months after conception.	Yes (1) No (0)

Table 12 (continued).

List of predictors of disability/delay status.

Variable	Scoring
Maternal Medical Risk Factors for Pregnancy	
Anemia	Present (1) Absent (0)
Lung Disease	Present (1) Absent (0)
Diabetes	Present (1) Absent (0)
(Oligo)hydramnios	Present (1) Absent (0)
Hypertension (chronic or gestational)	Present (1) Absent (0)
Eclampsia	Present (1) Absent (0)
Previous > 4 kg birth	Yes (1) No (0)
Previous preterm birth	Yes (1) No (0)
Other medical risk factors	Present (1) Absent (0)
Other Risk Factors for Pregnancy	
Mother smoked during pregnancy	Yes (1) No (0)
Mother consumed alcohol during pregnancy	Yes (1) No (0)
Preconception BMI	Binary indicator variables: Underweight, Overweight, Obese. Comparison group = Recommended weight.
Pregnancy weight gain	Binary indicator variables: No gain or loss, Undergain, Overgain. Comparison group = Recommended weight gain.
Obstetric Procedures	
Amniocentesis	Yes (1) No (0)
Fetal monitor	Present (1) Absent (0)
Tocolysis	Yes (1) No (0)
Ultrasound during first half of pregnancy	Yes (1) No (0)
Complications of Labor and Delivery	
Febrile (> 100 degree fever)	Present (1) Absent (0)
Membrane rupture > 12 hours	Present (1) Absent (0)
Placenta previa or abruptio placenta	Present (1) Absent (0)
Fetal distress	Present (1) Absent (0)
Other labor complications	Present (1) Absent (0)

Table 12 (continued).

List of predictors of disability/delay status.

Variable	Scoring
Method of Delivery	
Method of delivery	Binary indicator variables: C-section, Use of birth instruments. Comparison group = Vaginal birth without instruments.
Child Health Status at Birth	
5-minute Apgar score	Continuous
Birth weight category	Binary indicator variables: Extremely low, Very low, Moderately low, High. Comparison group = Normal
Preterm status	Yes (1) No (0)
Intrauterine growth category	Binary indicator variables: Small-for-gestational age, Normal-for-gestational age. Comparison group = Large-for-gestational age.
Abnormal Conditions of the Newborn	
Infant respiratory distress syndrome	Present (1) Absent (0)
Assisted ventilation	Binary indicator variables: > 30 minutes, < 30 minutes. Comparison group = 0 minutes
Other abnormal newborn condition	Present (1) Absent (0)
Prenatal Father Involvement	
Discussed pregnancy with spouse	Yes (1) No (0)
Saw a sonogram or ultrasound of the baby	Yes (1) No (0)
Listened to the baby's heartbeat	Yes (1) No (0)
Felt the baby move	Yes (1) No (0)
Attended childbirth classes or Lamaze classes	Yes (1) No (0)
Bought things for the child	Yes (1) No (0)
Present in the delivery room at birth	Yes (1) No (0)

Demographic predictors. Eight demographic predictors were identified: child race/ethnicity, sex of child, biological mother's age, biological father's age, family SES (family income, family poverty level, parents' highest level of education), others in the household with a disability, mother's prenatal employment, and mother's place of birth (Foreign vs. US).

Child race/ethnicity. Child's race/ethnicity has been linked to risk of having a disability/delay (e.g., Flores & the Committee on Pediatric Research, 2010; Hastings et al., 2005; Herring et al., 2006, Mandell et al., 2009). Race/ethnicity has also been linked to levels of father involvement (see Pleck & Masciadrelli, 2004). Six race/ethnicity groups were identified: Caucasian, non-Latino(a); African American, non-Latino(a); Asian, Native Hawaiian or other Pacific Islander, non-Latino(a); American Indian or Alaskan Native, non-Latino(a); Multiple races/ethnicities, non-Latino(a); and Latino(a). Four dichotomous variables were created for each group except Caucasian, non-Latino(a), which was the reference group.

Sex of child. A dichotomous variable with 0 = Male and 1 = Female. Males are more likely to have a disability/delay than females, particularly for specific disabilities/delays such as ASD (e.g., Jeans, Santos, Laxman, McBride, & Dyer, 2014). There is also some evidence, though mixed, that fathers may be more involved with sons than with daughters (see Pleck & Masciadrelli, 2004).

Mothers' and fathers' age at child's birth. This information was obtained from the birth certificate. Advanced maternal age has been linked to increased risk for having a child with a disability/delay, including chromosomal abnormalities, such as Down syndrome (Morris, Mutton, & Alberman, 2002), intellectual disabilities (Croen, Grether, & Selvin, 2001), Autism spectrum disorders (Croen, Grether, & Selvin, 2002; Glasson et al., 2004). Advanced paternal age has also been linked to increased risk for having a child with a disability, including congenital disorders such as a cleft lip or cleft palate (e.g., Savitz, Schwingl, & Keels, 1991), and Autism spectrum disorders (Reichenberg et al., 2006).

Young maternal age has also been identified as a possible predictor of disability/delay status (e.g., Gueorguieva et al., 2001). However, the association is not straightforward because young maternal age is often confounded with other risk factors for disability status including poverty, education, poor prenatal care, etc. (Hueston, Geesey, Diaz, 2008; Milne & Glasier, 2008). In some studies indicating an association between young maternal age and disability/delay status, the association disappears once these confounding factors are controlled for in analyses (e.g., Gueorguieva et al., 2001). In short, it is not known if the association between young maternal age and increased risk of having a child with a disability/delay is due solely to incomplete maternal biological development or is instead a result of comorbid risk factors. Even if young maternal age (via incomplete biological development) were not a significant predictor of disability/delay status in and of itself, young maternal age at birth should still be considered a risk factor for predicting disability/delay status in this study as a proxy for those other

factors. Some of these other factors may not have been included in the dataset while others may not have even been identified yet by researchers.

Since both young and advanced maternal age are risk factors for having a child with a disability, the association between maternal age and likelihood of having a child with a disability is not expected to be linear. To account for greater risk at both ends of the distribution of maternal age, the functional form of the association between maternal age and disability/delay status was assumed to be quadratic (curvilinear).

Young paternal age has not been linked to an increased risk of having a child with a disability/delay, although advanced paternal age has been as noted above. It is possible, however, that many of the risk factors associated with young maternal age may also be associated with young paternal age since mother's and father's age were strongly correlated (r = .72). For this reason, both linear and quadratic functional forms of the association between father involvement and disability/delay status were explored.

For both father's and mother's age, continuous measures were favored over categorical measures. Although, there may not be significant change in risk of having a child with a disability for parents throughout their 20s and early 30s, the exact cut points (e.g., age 30, age 32, age 37) are not known. Age categories derived from cut points chosen without theoretical guidance may obscure the effects of age on the likelihood of having a child with a disability/delay and or indicate a significant effect of an age category that simply capitalizes on chance variation in the data. Allowing the functional form of the association between parents' age and disability/delay status to take a quadratic form may be particularly useful in that it allows for the estimated risk to be fairly stable in the 20s and early 30s but increase quickly as it approaches the younger and older ages at the ends of the parabola.

Mothers' and fathers' age may also be associated with father involvement. For example, using data drawn from the 1997 Child Development Supplement to the Panel Study of Income Dynamics, Yeung, Sandberg, Davis-Kean, and Hofferth (2001) found that older fathers differed in their levels of involvement from younger fathers. However, the direction of the difference (i.e., higher or lower levels) varied by father involvement type.

Family socioeconomic status. Socioeconomic status (SES) has been repeatedly linked to child well-being, including the association between low SES and increased risk of having a disability (Bradley & Corwyn, 2002). Additionally, parents' level of education may be a protective factor against behaviors that could lead to having a child with a disability. Specifically, more educated parents may theoretically be more knowledgeable about and thus avoid risk factors, such as substance abuse, poor prenatal diet, etc. Socioeconomic status has also been linked to father involvement, though the association is often weak

(Dyer et al., 2009; Pleck & Hofferth, 2008; Pleck & Masciadrelli, 2004). Several variables assessing different aspects of family SES were available in the ECLS-B.

Family SES Measure 1: Family income. At 9 months, family income was reported. This variable was highly correlated with reported family income at 2 years and 4 years (r = .81, .76, respectively) suggesting that a family's level of income relative to that of other families in the sample is stable across time and may be an adequate proxy for pretreatment (pre-birth) income. However, it is possible that family income may have changed for some families as a result of having a child with a disability. For example, often mothers will leave the workforce or reduce their hours to take care of a child with a disability (Brandon, 2011; Olsson & Hwang, 2006), which decreases family earnings.

Family SES Measure 2: Family poverty level. Also included in the ECLS-B were variables indicating family poverty levels. These variables were derived from family income and household size and compared to census poverty thresholds. Three dichotomous variables were available indicating whether the family lived at 100%, 130%, or 185% of the poverty level. These variables were available at 9 months and were moderately to strongly correlated with the same measures at 2 and 4 years (r = .52, .50 respectively for 100% poverty level, r = .64, .59 respectively for 130% poverty level, r = .71, .65 respectively for 185% poverty level). As with family income, it is possible that some families may have entered into poverty if mothers (or fathers) left employment to care for their child with a disability.

Family SES Measure 3: Mother and father education. Mothers and fathers reported on their education levels at 9 months. Although this may have changed since the birth of their child for a few participants, it would only be moving them into the category that best describes them. For example, suppose a parent had some college before the child was born, but by 9 months he or she held a Bachelor's degree. For this to happen, he or she had to be close to graduation and could be reasonably placed in either category to reflect their level of knowledge, earning ability, and general socioeconomic status. Five categories of level of education were identified: completed 12th grade or below, high school diploma or equivalent, vocational/technology program or some college, bachelor's degree, and graduate or professional schooling/degree. Four dichotomous variables were created indicating the parent's level of education with parents holding a bachelor's degree as the comparison group. In addition to variables indicating mothers' and fathers' individual level of education, the ECLS-B includes a family-level parent education variable indicating the highest level of education of either parent. Because highest level of education was a better predictor of disability/delay status than parent's individual level of education (Pseudo $R^2_{\text{Highed Ed.}}$ = 1.09% vs. Pseudo $R^2_{\text{Mother \& Father Ed.}}$ = 0.59%) and may more adequately capture family-level SES, parents' highest level of education was used. To verify the stability of highest education and therefore its appropriateness for use as a proxy of parents' prenatal highest education levels, stability and change in education category was examined across time. From 9 months to 2 years,

86.38% of the sample stayed in the same education category and an additional 12.32% were within an adjacent category (e.g., a respondent family who had been in vocational/technology or some college category at 9 months was in the bachelor's degree category at 4 years). From 9 months to 4 years, 83.24% of the sample stayed in the same education category and an additional 15.31% were within an adjacent category. Some change was expected due to parents continuing their education and reporter error. Parents' highest education level showed remarkable stability indicating that parents' education levels at 9 months is an adequate proxy of parents' education levels prenatally.

Comparing measures of family SES. All measures of family SES were collected post-birth, but parents' highest level of education is probably the best proxy for pre-birth SES because it would be the most stable. Parents' highest level of education has greater stability because it changes more slowly over time than family income and can only increase when it does change. In contrast, family income could change frequently and decrease as well as increase. Consequently, family income (and therefore measures of poverty) was likely far more susceptible to being affected by child disability/delay status than parents' highest level of education. Parents' highest level of education was the least biased measure of pre-birth SES.

To examine the appropriateness of using parents' highest level of education as a measure of family SES, its relationship with other measures of SES was explored. Parents' highest education explained 31.72% of the variance in 9 month income. Parents' highest education also explained 22.71% of the variance in the 100% poverty level variable, 25.44% of the variance in the 130% poverty level variable, and 26.75% of the variance in the 185% poverty level variable at 9 months. The relative effectiveness of measures of family SES in predicting disability/delay status was also explored. Family income at 9 months added very little above highest education to the predictive ability of a model predicting disability/delay status from parents' highest level of education (ΔPseudo $R^2_{\text{Family Income}} = 0.11\%$, Pseudo $R^2_{\text{Highest Ed.}} = 1.09\%$). Measures of family's level of poverty likewise explained little beyond parents' highest level of education (ΔPseudo $R^2_{\text{100\% Poverty}} = 0.01\%$, ΔPseudo $R^2_{\text{130\% Poverty}} = 0.22\%$, ($\Delta$Pseudo $R^2_{\text{185\% Poverty}} = 0.02\%$).Given that parents' highest education (1) was the best predictor of child disability/delay status, (2) was likely the least biased measure of pre-birth family SES, and (3) explained a substantial amount of the variance in other measures of family SES and (4) other measures of family SES explained little beyond parents' highest level of education in predicting disability/delay status, only parents' highest level of education was used as a measure of family SES.

Others in household with a disability. At 9 months, mothers reported if any household members had a special need, delay, or disability (other than the target child if the child had a disability/delay). Although a crude proxy, this variable was included as a predictor because it might capture the genetic risk for having a child with a disability/delay found in some households. Even though this variable was

assessed post-birth, it is not likely that child disability/delay status influenced it during the first 9 months of life.

Mother's prenatal employment. Mothers reported whether or not they were employed at any time during the 12 months prior to their child's birth. Employment could potentially be a weak predictor of child disability/delay status if employed mothers experienced greater stress during the pregnancy. They could also be potential be exposed to various environmental teratogens as part of their employment. This variable is one of a handful that does not have strong empirical or theoretical evidence indicating a link with disability/delay status. Its inclusion in the model, even if incorrect, should not negatively impact the results.

Mother's place of birth (Foreign vs. US). Foreign born women have been shown to have better perinatal health than their US-born counterparts, including a lower occurrence of low birth weight babies, preterm delivery, and perinatal mortality (e.g., Forna, Jamieson, Sanders, & Lindsay, 2003; Kealher & Jessop, 2002). Thus, foreign born women may be at lower-risk for having a child with a disability/delay. A dichotomous variable was created with 1 indicating the mother was foreign-born and 0 indicating she was born in the US.

Maternal history. Two variables assessing aspects of the mother's history of births and terminations were identified.

Prior births. Birth-order may be a risk factor for having a child with a disability/delay. For example, risk of ASD is higher for first-born children than for later born children (e.g., Durkin et al., 2008). Mothers reported on the number of previous births. The distribution was skewed and a review of the association between prior births and child disability/delay status did not indicate any logical cut-point other than to dichotomize on no prior births (1) or 1 or more prior births (0).

Prior terminations. Mothers reported on the number of terminations (miscarriage or abortion) prior to the conception of the target child. It is possible that mothers who had an abortion did so because they discovered that the fetus had a disability/delay. However, there could be many other reasons as well. A miscarriage could be the result of the fetus having a severe disability that did not allow it to survive until birth. However, again, there could be other reasons for a miscarriage. Both of these scenarios could indicate genetic risk for having a child with a disability/delay. The distribution of prior terminations was skewed and no logical cut-point could be identified other than to dichotomize on no prior terminations (1) and 1 or more prior terminations (0).

Prenatal care. There is some evidence that no/inadequate prenatal care increases the risk of having a baby who is preterm, low-birth-weight, or small-for-gestational age (e.g., Heamann, Newburn-Cook, Green, Elliott, & Helewa, 2008), which in turn are associated with increased risk of the child having a disability/delay (e.g., Bolisetty, Bajuk, Me, Vincent, Sutton, & Lui, 2006; Hediger, Overpeck,

Ruan, & Troendle, 2002; Morse, Zheng, Tang, & Roth, 2009). Expectant mothers who receive no prenatal care are also more likely to engage in other behaviors that increase the risk of having a child with a disability, such as smoking or substance abuse (Maupin et al., 2004). They are also more likely to be less educated and not have insurance (Maupin et al., 2004). To the author's knowledge, there are no studies linking quality of prenatal care and father involvement. However, some of the covariates of prenatal care may be associated with father involvement. Therefore, measures of prenatal care merit inclusion as a predictor even though they may not be directly associated with father involvement. Several measures assessing various aspects of prenatal care were available in the ECLS-B data set: prenatal care information, adequacy of prenatal care utilization, and vitamin use 3 months before conception and 3 months after conception.

Prenatal care information. Mother's reported whether a doctor, nurse, or other healthcare worker provided them with information about appropriate prenatal care. Specifically, mothers reported whether or not they were given information about (1) what to eat during pregnancy, (2) how smoking could affect their baby, (3) how drinking alcohol could affect their baby, (4) what kinds of medicines were safe to take when pregnant, (5) how the baby grows and develops during pregnancy, (6) what to do if labor starts early, and (7) getting blood tested for HIV. A dichotomous score was formed for each item, with 1 indicating that the mother was given information about the topic and 0 indicating she was not. Less than 1% of the sample did not have any prenatal care visits and were not asked this question. Their scores were recoded as 0. Factor analysis indicated the presence of a single factor. These 7 items were averaged to form a single scale (Cronbach's $\alpha = .77$). Receiving information about appropriate prenatal care may increase a parent's knowledge and thus decrease the risk of having a child with a disability/delay.

*Adequacy of Prenatal Care Utilization Index (APNCU).*The APNCU (Kotelchuck, 1994a) is a function of the month that prenatal care began (or did not begin) and the "proportion of the number of visits recommended by the American College of Obstetricians and Gynecologists (ACOG) received from the time prenatal care began until delivery" (Kotelchuck, 1994a, p. 1418). One of the advantages of the APNCU over other indices of prenatal care is that the APNCU includes a category for "adequate plus" care that includes expectant mothers whose use of prenatal care exceeded ACOG's standards. Kotelchuck (1994b) found that these mothers and the mothers with "inadequate" prenatal care were more likely to have low-birth-weight babies (an important risk-factor for childhood disabilities/delays) than mothers who had "intermediate" or "adequate" care. These results suggest that a particularly high number of prenatal care visits may indicate pregnancy complications that may elevate the risk of having a child with a disability/delay. On the other hand, a particularly low number of visits and or late initiation of prenatal care may also be a risk factor for having a child with a disability/delay.

Vitamin use 3 months before and 3 months after conception. Mothers reported at 9 months whether they took prenatal vitamins during the 3 months before they found out they were pregnant. They also reported whether they took vitamins during the 3 months after they found out. In addition to promoting positive prenatal development and reducing risk of birth defects, such as a neural tube defect, the use of vitamins may indicate that mothers were aware of the pregnancy or potential pregnancy and were more likely to be avoiding harmful environmental teratogens.

Medical risk factors for pregnancy. From the birth certificate, information on several medical risk factors for the pregnancy was available. These risk factors may contribute to the risk of having a child with a disability/delay as they may indicate a less than ideal environment for prenatal development (e.g., lower oxygen supply, higher blood pressure, etc.). A dichotomous variable was formed for each risk factor indicating the presence (1) or absence (0) of the risk factor.

Anemia. Maternal anemia is a risk factor for low birth weight and preterm delivery (Levy, Fraser, Katz, Mazor, & Sheiner, 2005), which in turn are risk factors for having a child with a disability/delay.

Maternal lung disease. Mothers with lung disease may have difficulty getting enough oxygen for themselves and their fetus, which could negative affect the fetus' development.

(Oligo)hydramnios. Oligohydramnios is a condition of having too little or too much amniotic fluid and is a risk factor for the fetus' development.

Hypertension. Hypertension can put a mother had higher risk of other pregnancy complications. Birth certificates included information on gestational and chronic hypertension. The occurrence of gestational hypertension was low, so a dichotomous indicator variable was created indicating whether the mother had gestational *or* chronic hypertension (1) or no hypertension (0).

Eclampsia. Eclampsia is a serious condition consisting of seizures following preeclampsia that is dangerous for the health of both the mother and the baby.

Previous 4 kg + baby. A mother who has previously given birth to baby over 4 kg may be at risk for having another large baby (Heiskanen, Raatikainen, & Heinonen, 2006; Walsh, Mahony, Foley, Daly, & O'Herlihy, 2007). Although large babies can be delivered safely, there can be risk of complications such as shoulder dystocia.

Previous preterm baby. Mothers who have previously given birth to a preterm or small baby may be at risk for having another. Preterm birth status is associated with greater risk of having a child with a disability.

Other medical risk factors. Birth certificates included an "other medical risk factors" category for risk factors not listed above or included in the list that follows: cardiac disease, herpes, hemoglobinopathy, incompetent cervix, kidney disease, rh sensitization, uterine bleeding. The medical risk factors in the list were not included because either they were not thought to be a risk factor for having

a child with a disability/delay or the occurrence of the risk factor in the sample was too low to be used as a predictor.

Other risk factors for pregnancy. Additional risk factors for a healthy pregnancy included cigarette use during pregnancy, alcohol use during pregnancy, preconception BMI category, and pregnancy weight gain category.

Cigarette use. Cigarette use is a commonly warned against teratogen that can have harmful effects on children's prenatal development. A dichotomous variable was created with 0 indicating no prenatal cigarette use and 1 indicating prenatal cigarette use reported either on the birth certificate or retrospectively at the 9-month parent interview.

Alcohol use. Alcohol use is another commonly warned against teratogen, though what constitutes a harmful dose is not clear and may vary from person to person. A dichotomous variable was created with 0 indicating no prenatal alcohol use and 1 indicating prenatal alcohol use reported either on the birth certificate or retrospectively at the 9-month parent interview.

Preconception BMI. At 9 months, mothers reported their preconception weight as well as their current height. This information was used to calculate mothers' preconception body mass index (BMI). Based on their BMI, mothers were placed into one of four categories: underweight, normal weight, overweight, and obese. Being underweight or overweight or obese is associated with increased risk of pregnancy complications and having a low birth weight baby (Moos et al., 2008).

Pregnancy weight gain. Information on weight gained during the pregnancy was reported both on the birth certificate and retrospectively by mothers at 9 months. Birth certificate data was used for all cases except those for which it was missing. For these cases, self-report data was used. However, a review of the self-report data identified 8 cases that were outliers relative to the other cases. Examination of these cases suggested that these mothers misunderstood the question and reported their preconception weight rather than the weight gain. Consequently, these cases were recoded to missing. Excluding the outliers, the correlation between self-report and birth-certificate-reported pregnancy weight gain was large (r = .71). The Institute of Medicine release new guidelines in 2009 for recommended weight gain based on preconception BMI (Rasmussen & Yaktine, 2009). These guidelines recommend that women who are unweight, normal weight, overweight, or obese should gain 28 – 40 lbs., 25 – 35 lbs., 15 – 25 lbs., or 11 – 25 lbs., respectively. Based on these recommendations mothers were placed into four pregnancy weight gain categories: No weight gain or weight loss, Gained too little, Gained the recommended amount, or Gained too much. Gaining too little or gaining too much during pregnancy has been associated with a number of risk factors children's development, such as low birth weight, macrosomia, small-for-gestational age, large-for-gestational age, preterm birth, and infant mortality (Rassmussen, Catalano, & Yaktine, 2009).

***Obstetric Procedures*.** Information on obstetric procedures was available on the birth certificates. A dichotomous indicator was created for each procedure with 1 indicating the procedure was done and 0 indicating it was not.

Amniocentesis. Amniocentesis is a procedure for identifying chromosomal abnormalities. Mothers who elect to perform amniocentesis may have reason to suspect their child has a disability and thus may be at greater risk for having a child with a disability/delay.

Fetal monitor. Elective fetal monitors were used by 84.5% of the sample. Children who were delivered without the use of a fetal monitor may have been at greater risk for experiencing fetal distress and other complications that went undetected and so were not addressed. These complications could, in turn, increase the risk of disability/delay, though the association is likely weak.

Tocolysis. Tocolysis refers to the inhibition of labor contractions to prevent preterm delivery. Children threatened with preterm delivery may be at greater risk for poorer development.

Ultrasound. Mothers reported at 9 months if they had an ultrasound during the first half of the pregnancy. Over 90% of the sample had an ultrasound. Failure to have an ultrasound could indicate lower quality prenatal care, which is a risk factor for having a child with a disability/delay.

Complications of labor and or delivery. Data on several complications of labor and delivery were available from the birth certificates. Each of these complications has the potential to negatively affect children's health and thus increase their risk of having a disability/delay.

Febrile. A dichotomous variable was created indicating whether (1) or not (0) the mother had a fever above 100 ºF during labor and delivery.

Premature rupture of the membrane. A dichotomous variable was created indicating if the membrane had been ruptured for more than 12 hours, but labor had not started (1) or the membranes were not prematurely ruptured (0).

Problems with the placenta. A dichotomous variable was created indicating whether there was a problem with the placenta (placenta previa or abruptio placenta) (1) or there were no such problems (0).

Fetal distress. A dichotomous variable was created indicating the absence (0) or presence (1) of signs indicating fetal hypoxia, or a deficiency of oxygen.

Other labor complications. A dichotomous variable was created indicating the absence (0) or presence (1) of other labor complications not listed above or in the following list: moderate/heavy meconium, other excessive bleeding, seizures during labor, precipitous delivery, prolonged labor, dysfunctional labor, breech delivery, cephalopelvic disproportion, cord prolapse, or anaesthetic complications. The labor complications in the list were not included because either they were not thought to be a risk factor for having a child with a disability/delay or the occurrence of the risk factor in the sample was too low to be used as a predictor.

Method of delivery. The method of delivery was available from the birth certificate. Three categories were formed into two dichotomous variables: C-section and vaginal delivery with instruments. The comparison group was vaginal delivery without instruments. Although all 3 methods can be safely used to deliver a baby, C-section and instrumental deliveries may indicate some complication in labor and delivery that might be related to the risk of disability/delay status.

Child health status at birth. Several measures of children's health at birth were available. These measures may indicate current health problems that are a result of labor and delivery complications or problems during prenatal development, which in turn may place children at greater risk for having a disability/delay.

5-minute APGAR score. The APGAR test is a commonly used assessment of children's health and adaption to the external environment. Lower APGAR scores may indicate development difficulties.

Child birth weight status. Four dichotomous variables indicating extremely low, very low, moderately low, and high birth weight were created. The comparison category was normal birth weight. Birth weight status has been linked to risk of disability/delay (e.g., Hediger et al., 2002).

Preterm birth status. The birth certificate included data on length of gestation based on both clinical estimates and last menstrual period. The two estimates differ somewhat in identifying children as preterm. Consequently, children were identified as being preterm if either method indicated they were. Preterm birth status has also been linked to risk of disability/delay (e.g., Bolisetty et al., 2006; Morse et al., 2009).

Intra-Uterine Growth (IUG). An NCES generate composite variable included in the ECLS-B data set assessed children's intrauterine growth retardation. Three groups were identified: small for gestational age (< 10^{th} percentile for birth weight for gestational age), large for gestational age (> 90^{th} percentile for birth weight), and normal for gestational age (10 – 90^{th} percentile). Normal for gestational age was the comparison category. Intrauterine growth has been linked to risk of having a child with a disability/delay (e.g., Hediger et al., 2002).

Abnormal conditions of the newborn. These variables also provide assessments of children's health at birth.

Infant Respiratory Distress Syndrome (IRDS). Birth certificate data included information on the presence of IRDS at birth. IRDS is a syndrome in premature infants that requires them to need oxygen and help breathing.

Assisted ventilation. Birth certificate data included information on whether the child needed assisted ventilation for over 30 minutes, less than 30 minutes, or not at all. Two dichotomous variables were created for assisted ventilation for over 30 minutes and for under 30 minutes with "no assisted ventilation" as they comparison category.

Other abnormal newborn conditions. A dichotomous variable was created indicating the absence (0) or presence (1) of other abnormal newborn conditions not listed above or in the following list: anemia, birth injury, fetal alcohol syndrome, meconium aspiration syndrome, seizures. The abnormal conditions in the list were not included because they were not thought to be a risk factor for having a child with a disability/delay, the occurrence of the risk factor in the sample was too low to be used as a predictor, or the condition was itself a disability/delay.

Prenatal father involvement. Several measures of prenatal father involvement reported retrospectively by fathers at 9 months were available. While these measures may not be directly associated with the risk of having a disability/delay, they may be indirectly associated. Furthermore, they are likely associated with later father involvement as well and should therefore be including in estimating the propensity score. Seven items assessing whether (1) or not (0) fathers engaged in prenatal father involvement behaviors were identified: discussed pregnancy with partner, saw a sonogram or ultrasound of the baby, listened to the baby's heartbeat, felt the baby move, attended childbirth classes or Lamaze classes, bought things for the child, and presence in the delivery room at birth.

Missing Data

Missing data in the FAI. The amount of missing data was less than 6% for most domains. Specifically, the composite was created by averaging across items. The amount of missing data for the domains of motor skills at T3, communication skills at T3, and self-regulation at T2 and T3 ranged from 6.00% to 10.0%. The amount of missing data for the domains of Social Skills at T1 – T3 ranged from 12.00% to 17.00%. Missing values for the 60 + variables used to create the FAI were imputed using Stata's *ice* command (StataCorp, 2011; see Royston, 2005, 2007, 2009 for details on the *ice* command), which uses regression switching to impute missing values. van Buuren, Boshuizen, and Knook (1999) recommend twenty cycles or iterations, but note that 5 are probably sufficient. Twenty cycles were run in creating a single, fully imputed data set. For the regression models imputing missing values for variables used to create the FAI, predictors were the scores for the other domains. Additionally, 110 other available in the ECLS-B variables (e.g., birth weight status, prenatal substance use, demographic variables, disability/delay status indicators, measures of father and mother involvement, and other measures of children's development) were included in the *ice* command during the imputation process. Including these variables improves the quality of the estimates of missing values by providing supplementary information and by helping account for the mechanisms responsible for missing data (Acock, 2012). The data set with the imputed values was then used to create the index variables for each domain of the FAI as outlined above. This process resulted in complete FAI scores for all children in the sample.

Missing father involvement data. The amount of missing data for father involvement ranged from 2.26% to 12.00% within each father involvement type. However, for every type of father

involvement, less than 5.00% of the sample was missing more than 1 item. Because father involvement was a latent variable, missing data were addressed in creating the father involvement latent variables by using full information maximum-likelihood (FIML) in the Mplus 7.11 program. FIML is a preferred approach to other common methods (Acock, 2005) including listwise deletion, which is considered one of the least effective methods for dealing with missing data (see Wilkinson & the Task Force on Statistical Inference, 1999).

Missing data on disability/delay status. There were only a handful of cases that could not be assigned due to missing data to the disability/delay group, the typically developing group, or the group of excluded environmental/acquired disabilities. They represented only 0.28% of the sample and thus there was little risk that their exclusion could bias the analyses. Consequently, these cases were dropped from the sample.

Missing data on predictors of disability/delay status. Missing data for predictors of disability/delay status was limited. Most variables had less than 3.52% of cases missing. Exceptions were fetal distress which was missing 7.52% and the 5-minute APGAR score, which was missing for 19.38% of the sample. Missing data for variables used to create the propensity score were imputed using Stata's *ice* command similar to how missing data on the FAI was imputed. A dichotomous variable indicating that data was missing on a variable for a particular case was created for each variable containing missing data used in creating the propensity score. Balance on this variable must also be achieved to avoid violating the conditional independence assumption discussed previously. Because of multicollinearity, some missing data indicators were combined. For example, two missing data indicators overlapped on all but 1 case, and so were combined.

Analytic Plan Part 2: Analyses for each Question

The present study explored each of the following four research questions:

1. How does the presence of a child's disability or developmental delay influence father involvement across the first four years of life?
2. How are children's functional abilities associated with father involvement?
3. Do children's functional abilities mediate the association between disability/delay status and father involvement?
4. How is father involvement associated with children's functional abilities and does father involvement mediate the effect of disability/delay status on functional abilities?

Propensity score analysis was used to obtain estimates of the causal effects of disability/delay status on father involvement (Question 1) and on functional abilities (Questions 3 and 4). Four types of father involvement were examined in this study: routine caregiving involvement, play involvement, literacy involvement, and responsive caregiving involvement. Because propensity score analysis has been

primarily used with a single outcome per model, the research questions of this study were examined for each of the four involvement types separately. As mentioned previously, I focused on specific types of positive engagement (i.e., literacy, play, routine caregiving, and responsive caregiving involvement) available in the data set. I examined these specific types of father engagement as opposed to overall father involvement because they addresses different needs of children, such as routine physical needs through routine caregiving, physical and cognitive stimulation through play and literacy activities, and less-routine physical needs through responsive caregiving. Furthermore, specific types of father involvement may be more strongly affected by disability/delay status or functional abilities and some may not be affected at all.

Question 1. The first question addresses the effect of child disability/delay status on father involvement across the first four years of life. Propensity score analysis was used to address this question. Once an adequate matched sample was created, a regression model was fit using the matched sample in which disability/delay status is a predictor of father involvement, providing an estimate of the causal effect of disability/delay status on father involvement. A separate model was fit for each type of father involvement at each time point.

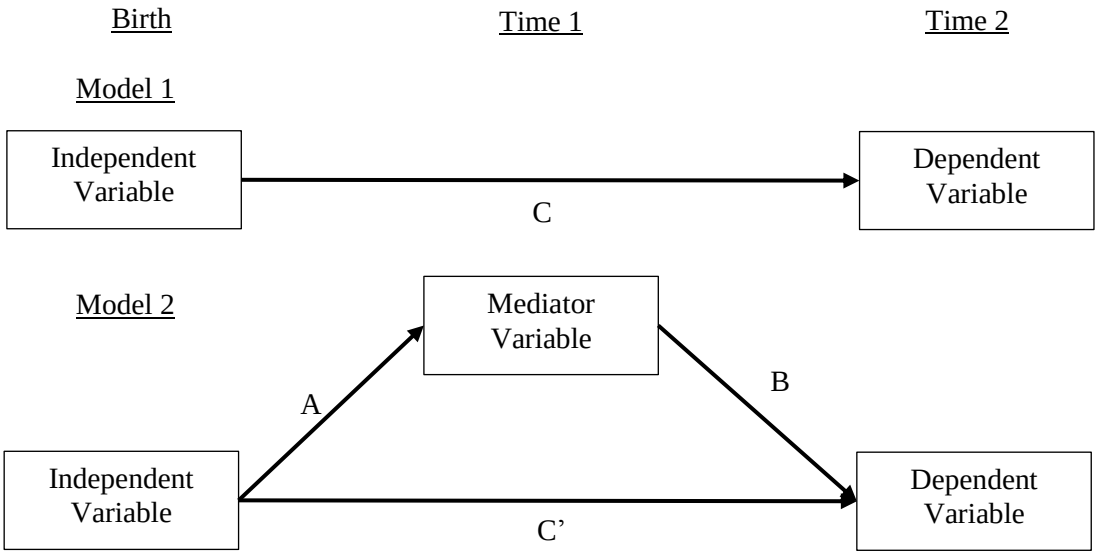
The presence of a child with a disability/delay may affect father involvement not only directly, but also through father involvement at earlier time points. That is, earlier father involvement may mediate the association between child disability/delay status and later father involvement. Before further discussing this potential mediation effect, it may be useful to review the mediation model and relevant terms. Figure 10 presents two models. In Model 1, the dependent variable (DV) is affected only by the independent variable (IV). In Model 2, two new paths have been added allowing the IV to affect the DV through a mediator variable (MV) as well. Path C is referred to as the total effect. Path C' is referred to as the direct effect. Paths A and B combine to form the indirect effect (AB). Baron and Kenny (1986), Judd and Kenny (1981), and James and Brett (1984) outlined four steps in establishing mediation: First, show that the IV is associated with the DV (Path C). Second, show that the IV is associated with the MV (Path A). Third, show that the MV is associated with the DV (Path B). Fourth, show that when controlling for the MV the association between the IV and the DV is 0 (C' = 0) for complete mediation or decreases (C' < C) for partial mediation. The first step has faced some criticism as a requirement for mediation because suppression effects may mask a significant Path C (MacKinnon, 2008). Additionally, inconsistent mediation can occur when the indirect effect and the total effect have opposite signs. In this situation, the MV behaves as a suppressor in that the effect of the IV on the DV is larger once the mediator is included (e.g., C' < C; MacKinnon, Krull, & Lockwood, 2000; Rucker, Preacher, Tormla, & Petty, 2011). In such cases, there may not be a significant association between the IV and the DV in the absence of the MV (i.e., a non-significant total effect; Path C) due to the suppression effect. Furthermore, Rucker and

colleagues (2011) demonstrated with a simulation study and experimental results that there can be significant indirect effects even when significant total effects (Path C) are absent (see Rucker et al., 2011, for several reasons why total effects may not be detected). Following their recommendations, a significant total effect was not a requirement for testing mediation in this study.

The use of the terms full and partial mediation has also been criticized in recent years. The term partial mediation is used to describe mediation models in which Path C' < Path C, but still significantly different from 0. In this situation, the MV is thought to partially mediate the effect of the IV on the DV. The term full mediation is used to describe mediation models in which Path C was significant, but Path C' was not. In this situation the MV is thought to fully mediate the effect of the IV on the DV and no other mediators exist. Rucker and colleagues (Rucker et al., 2011) demonstrated with a simulation study and experimental results that there can be significant indirect effects even when significant direct effects (Path C') are absent (see Rucker et al., 2011, for several reasons why direct effects may not be detected).The implication of their demonstration is that a mediation model in which a significant Path C becomes a non-significant Path C' in the presence of an MV is not necessarily full mediation. Perhaps for extremely well-studied, simple phenomenon, a researcher can claim full mediation. In general, however, Rucker and colleagues suggest that the term "full mediation" be avoided. Additionally, they recommend focusing on the significance of the indirect effect regardless of the significance of the total or direct effects.

Figure 10.

Mediation Model.



The analysis of earlier father involvement as a mediator may provide insight into how child disability/delay status influences father involvement. However, it also has the potential to introduce bias into the model because father involvement occurs after the child's birth and the hypothesized initial presence of the disability/delay. If unobserved factors influence father involvement differently for fathers of children who are typically developing than fathers of children with disabilities/delays, then model estimates may not be accurate (see Figure 11). In short, the addition of earlier father involvement as a mediator changes the model from a causal model to an observational/correlational model. However, the insights gained from the mediation analyses may still be useful. To reduce bias in the model, additional control variables can be used: child race/ethnicity, fathers' age, sex of the child, and family SES. All but family SES were included in the formation of the propensity score and do not need to be controlled for again because they did not change in value or they changed similarly for all participants (e.g., fathers' age). Although parents' highest education was included as a proxy for family SES in creating the propensity score, family SES may have changed since birth and should be included as a model covariate. Family SES is a quintile variable generated by NCES that incorporates family income, mothers' and fathers' level of education, and mothers' and fathers' occupation prestige. Mediation analyses were conducted using the *sgmediation* program in Stata 12 (Ender, 2006). This program performs the Sobel-Goodman tests to check if a mediator carries the effect of an independent variable to the dependent variable. It can also account for complex survey design.

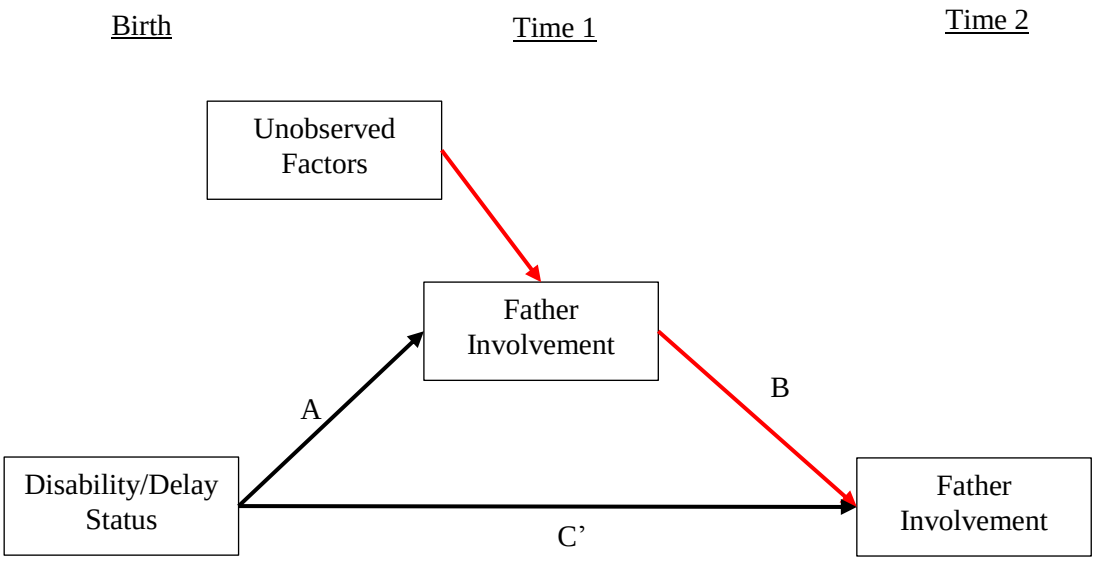
Sobel-Goodman tests face criticism because they assume a normal distribution of the indirect effect (AB), which may not be accurate. Preacher and Hayes (2004) suggest creating an empirical distribution through bootstrapping to create confidence intervals. This approach does not assume a normal distribution and thus provides a better alternative for testing the significance of the direct effect. For all mediation analyses, I used the Preacher-Hayes bootstrapping approach with 1000 resamples to generate a 95% confidence interval. If the confidence interval did not include 0, then the indirect effect is significantly different from 0 at the $\alpha = .05$ level.

Mediation is best demonstrated if the independent variable, the mediator, and the dependent variable are at different time points. Given that disability/delay status is conceptualized as being present from birth, the following mediation models were examined for each type of father involvement:

1. Child disability/delay status → Father involvement (9 months) → Father involvement (2 years)
2. Child disability/delay status → Father involvement (9 months) → Father involvement (4 years)
3. Child disability/delay status → Father involvement (2 years) → Father involvement (4 years)

Figure 11.

Mediation Model of the Effect of Disability/Delay Status on Father Involvement Mediated by earlier Father Involvement with Unobserved Factors.



Question 2. The second question addressed the association between children's functional abilities and later father involvement. Because I matched on the propensity to have a disability and not the propensity to have a specific value on the Functional Abilities Index (FAI), causal estimates of the effect of children's functional abilities on father involvement cannot be obtained. However, I still examined the association between children's functional abilities and father involvement using the matched sample to 1) facilitate comparisons of this association with the effect of disability status on father involvement and 2) capitalize on a possible reduction in bias of the estimate of the association between the FAI and father involvement due to having "artificially" created a randomized sample. To elaborate, in observational/correlational studies it is common practice to control for variables associated with both independent (functional abilities) and dependent variables (father involvement) to reduce bias introduced into model estimates by these "third variables." Because the sample was matched on many of these variables (e.g., the sex of the child, parents' education, age of the parents), bias in model estimates due to these variables is reduced. Additionally, as was the case for the mediation analyses for Question 1, controlling for family SES can also reduce bias. It is possible, though, that there are yet other variables that are associated with both children's functional abilities and father involvement that were not included in estimating the propensity score. Additionally, some of the pre-birth variables accounted for through propensity score matching may have changed over time. The inclusion of prior values of the endogenous (outcome) variable ("a lagged endogenous variable") in a model helps control for the unmeasured third variables (Menard, 2002). Thus, by including prior levels of father involvement in the models addressing Question 2, bias in the estimates can be further reduced.

Furthermore, the inclusion of a lagged endogenous variable can also reduce bias due to changes in control variables from their pre-birth values. For example, control variables that have new values at 9 months may be associated with both children's functional abilities and later father involvement and may bias the estimates of the association of functional abilities and father involvement. However, the control variables are likely also associated with father involvement at 9 months, which is in turn associated with father involvement at later time points. Therefore, including father involvement at 9 months in the model predicting father involvement at a later time point helps address some of this bias.

It should be noted that the measures of father involvement vary across time in terms of item type and number. An exception is father literacy involvement which was measured using the same 3 items at each time point. The measurement invariance analyses mentioned previously confirmed that father literacy involvement is indeed the same construct across time (and across groups). In contrast, routine caregiving, responsive caregiving, and play involvement differed across time in the number and type of items used and thus could not be tested for full measurement invariance across time. For these types of involvement, including prior father involvement in the model does not represent a perfect lagged

endogenous variable. This is because the outcome measure of father involvement and the measure of prior father involvement are not the same—at least in terms of how they were measured. However, in consideration of the dramatic changes in children's needs and abilities across the first four years of life, it may not be reasonable to expect that measures of father involvement using the same items across time would show strong or strict invariance. Indeed, the strict invariance of father literacy involvement across time may be an exception. Would one expect play, routine caregiving, or even responsive caregiving involvement to look the same for fathers of infants as for fathers of toddlers? That being said, prior "imperfect" lagged endogenous variables may still accomplish the goal of reducing bias in our model if these prior, "imperfect" measures are correlated with the target "third variables" and later father involvement. Although many of the third variables are unknown and/or unmeasurable, correlations between earlier and later father involvement can be examined. In the present study measures of father involvement were significantly correlated across time. The average of the 9 month to 2 year, 9 month to 4 year, and 2 to 4 year correlations for each type of father involvement are as follows: play involvement, $r_{avg} = .60$; routine caregiving involvement, $r_{avg} = 0.78$; and literacy involvement, $r_{avg} = 0.67$, and $r_{avg} = 0.62$; responsive caregiving involvement. The strength of these correlations suggests that prior measures of father involvement are adequate lagged endogenous variables.

A consequence of including prior levels of father involvement in the model is that doing so may change the meaning of the estimated association between the independent and dependent variables. Specifically, it has been suggested (see Finkel, 1995; Selig & Little, 2012) that when a prior measure of the outcome variable is included in the model, then the estimate of the association between the independent variables and the dependent variables should be interpreted as the amount that the outcome variable *changes* for a one unit change in the independent variables (conditional on the control variables). Such models are known as lagged-regression models (also called "residual change models"). However, this type of change is different from the type of change examined by using growth models or change scores which estimate within-person change. In contrast, residual change models are actually looking at rank-order instability or interindividual variability in the outcome (Selig & Little, 2012). Unlike growth curve models, residual change models tell us nothing about within-person change or intraindividual variability. In short, the inclusion of prior levels of father involvement for the analyses addressing Question 2 leads to a change in interpretation of model estimates, which interpretation is not as readily understood (or agreed upon; see Newsom, 2012) as a model without prior levels of father involvement. Consequently, when examining the association between children's functional abilities and later father involvement, I fit both models with and without prior levels of father involvement. A separate model was fit for each type of father involvement at each time point.

Question 3. The third question explored whether children's functional abilities mediate the relationship between disability/delay status and father involvement. Mediation was tested by fitting a series of regression models in which the associations between disability/delay status, children's functional abilities, and father involvement were explored to see if there was an indirect effect of disability/delay status on father involvement through children's functional abilities (see Figure 6). As noted before, mediation is best demonstrated if the independent variable, the mediator, and the dependent variable are at different time points. For this question, the following mediation models were fit separately for each type of father involvement:

1. Child disability/delay status → Functional abilities (9 months) → Father involvement (2 years)
2. Child disability/delay status → Functional abilities (9 months) → Father involvement (4 years)
3. Child disability/delay status → Functional abilities (2 years) → Father involvement (4 years)

It should be noted that for the same reasons that estimates of the causal effect could not be obtained for the mediation analyses for Question 1, neither could causal mediation estimates be estimated for Question 3. Estimates of the causal effect of disability/delay status on functional abilities can be obtained, but estimates of the causal effect of functional abilities on father involvement cannot due to the presence of "third variables" associated with both children's functional abilities and father involvement. However, because the sample was matched on many of these variables, bias in model estimates due to these variables is reduced by fitting the model in the matched sample. As was the case with Question 2, including prior values of father involvement in the model can help control for any additional unmeasured third variables and further reduce bias in model estimates. As previously noted, since the inclusion of prior levels of father involvement in the model changes the interpretation, the mediation models were fit both with and without prior levels of father involvement. Family SES was also included as a control variable. Mediation analyses were conducted using the *sgmediation* program in Stata 12 (Ender, 2006).

Question 4. The fourth question addresses how father involvement is associated with children's functional abilities and whether father involvement mediates the effect of disability status on functional abilities (see Figure 7). The causal effect of father involvement on children's functional abilities cannot be estimated since the sample was not matched on propensity for specific levels of father involvement. However, by fitting a regression model in which children's functional abilities are predicted by earlier father involvement using the matched sample, some of the bias in the estimates can be reduced. Furthermore, controlling for prior levels of children's functional abilities can help reduce bias in the estimates of the association between father involvement and children's functional abilities. Similar to what was done for the previous questions, models controlled for family SES and were fit with and without prior levels of children's functional abilities. These analyses were done for each type of involvement separately. Exploration of the mediation of the effect of disability/delay status on functional

abilities by father involvement was explored in a similar manner as outlined for Question 3. However, the mediator was father involvement and the outcome was children's functional abilities. Specifically, for this question, the following mediation models were fit separately for each type of father involvement:

1. Child disability/delay status → Father involvement (9 months) → Functional abilities (2 years)
2. Child disability/delay status → Father involvement (9 months) → Functional abilities (4 years)
3. Child disability/delay status → Father involvement (2 years) → Functional abilities (4 years)

Chapter Three: Results

Propensity Score Matching

Overview. The process of propensity score matching is as follows: Step 1. The propensity score is estimated using logistic regression predicting child disability/delay status from the identified covariates. Step 2. Each child with a disability/delay is matched to one or more children who are typically developing. Children who are typically developing are assigned weights using the procedure previously described to reflect how often they were matched adjusted for how many other matched children there were for each child with a disability. Step 3. The conditional independence assumption is partially tested by checking for imbalance on the covariates. If imbalance is detected, the analyst returns to Step 1 and adds higher order terms and interactions term to the propensity score model try to address the balance. The propensity score is re-estimated and Steps 2 and 3 are repeated. This cycle continues until there is no imbalance on the covariates. In practice, it may not be possible to balance on all covariates. This is of greater concern for covariates that are principal predictors of disability/delay status than for other covariates. If balance is not achieved, one option is to control for the imbalanced covariates in the propensity score analyses.

Creating the Propensity Score. The procedure describe above was followed and the cycle repeated until the final model was identified. The final propensity score model was generated using all of the variables listed in Table 12 in a logistic regression model predicting disability/delay status. All variables were assumed to have a linear association with the logit of disability/delay status with the exception of maternal age, which was allowed to have a curvilinear (quadratic) association. To attempt to address the imbalance on covariates resulting from initial models, higher order terms were added to the model. Specifically, interactions between mothers' age and the presence of hypertension during pregnancy or chronic hypertension and preterm status of the child were added. Additionally, interactions between mothers' age and parents' highest education and child race/ethnicity were added. Interactions between fathers' age and parents' highest education and child race/ethnicity were also added.

The model was fit with these additional variables and the resulting propensity scores were saved. Review of the range of propensity scores for the two groups revealed that children who were typically developing had a range of 0.17% to 94.25% probability of having disability/delay and children who were later diagnosed with disabilities/delays had a range of 2.26% to 99.28% probability of having disability/delay. Hosmer, Lemeshow, & Sturdivant (2013) recommend that control cases (i.e., children who are typically developing) with propensity scores values lower than the lowest value in the treatment group (i.e., children with disabilities/delays) or higher than the highest value in the treatment group should be dropped from the sample and the propensity score model be refit. Accordingly, approximately 150 cases (5.82%) were dropped from the typically developing group and the propensity score model was

refit. The new model did not have any "out of range" cases in the typically developing group. The refit, final model is reported in Table 13.

It should be emphasized that there is likely considerable multicollinearity in the model given the large number of correlated predictors. This is not a problem for creating the propensity score. In fact, all relevant variables *should* be included in the propensity score model regardless of issues of multicollinearity in order to obtain the best estimate of each child's propensity for having a disability/delay. However, the significance of predictors should not be interpreted.

Figures 12 and 13 display the kernel density estimates of propensity scores of children who are typically developing (CTYP) and children with disabilities/delays (CWD) using the full, unmatched sample on the initial and refit propensity score models, respectively. Note that in Figure 12, there are a number of children who are typically developing who have propensity scores lower than the minimum propensity score of the group of children with a disability/delay. Such is not the case in Figure 13. These figures graphically represent how the likelihood of having a disability/delay is unequally distributed between the two groups and the need for propensity score matching.

Table 13.

Propensity Score Model Predicting Child Disability/Delay Status from Pre-birth Covariates.

Outcome: Child Disability/Delay Status	Odds Ratio	Linearized Std. Errors	*t*	*p*	95% Confidence Interval of Odds Ratio	
Demographics						
Child is African-American	0.39	0.23	-1.63	.107	0.12	1.23
Child is Latino(a)	0.42	0.15	-2.49	.015	0.21	0.84
Child is Pacific Islander or Asian	0.37	0.16	-2.32	.022	0.16	0.86
Child is Native American or Alaskan Native	1.00	0.51	0.00	.999	0.36	2.75
Child is Multi-racial/multi-ethnic	0.52	0.25	-1.36	.178	0.20	1.36
Comparison Group: Child is Caucasian, non-Latino(a)						
Child is female vs. male	0.56	0.09	-3.42	.001	0.40	0.79
Mothers age	0.96	0.06	-0.58	.564	0.84	1.10
(Mothers' age)^2	1.00	0.00	0.31	.756	1.00	1.01
Fathers' Age	1.07	0.05	1.48	.144	0.98	1.17
Highest Education: 12th grade or below	2.05	1.12	1.31	.194	0.69	6.07
Highest Education: High school diploma or equivalent	1.04	0.36	0.10	.921	0.52	2.07
Highest Education: Vocational/Technology program or some college	1.35	0.34	1.18	.240	0.82	2.24
Highest Education: Graduate or professional schooling/degree	1.34	0.37	1.05	.296	0.77	2.34
Comparison Group: Highest Education: Bachelor's degree						
Others in home with special need, delay, or disability	1.13	0.38	0.38	.706	0.59	2.19
Mother was employed during the 12 months prior to birth	0.84	0.18	-0.82	.415	0.55	1.29
Mother was born outside the U.S. (foreign-born)	0.92	0.31	0.29	.775	0.52	1.64

Table 13 (continued).

Propensity Score Model Predicting Child Disability/Delay Status from Pre-birth Covariates.

Outcome: Child Disability/Delay Status	Odds Ratio	Linearized Std. Errors	*t*	*p*	95% Confidence Interval of Odds Ratio	
Maternal History						
No previous births.	0.99	0.22	-0.02	.981	0.65	1.53
No terminations (miscarriages or abortions)	1.18	0.24	0.80	.428	0.78	1.77
Quality of Prenatal Care						
Prenatal care information to mother	1.48	0.60	0.98	.329	0.67	3.30
Adequacy of prenatal care utilization: Inadequate	0.88	0.41	-0.27	.784	0.35	2.21
Adequacy of prenatal care utilization: Intermediate adequacy	1.07	0.31	0.25	.801	0.61	1.90
Adequacy of prenatal care utilization: Adequate plus	1.31	0.29	1.21	.229	0.84	2.05
Comparison Group: Adequacy of prenatal care utilization: Adequate						
Prenatal vitamin use 3 months before conception.	1.22	0.28	0.87	.386	0.77	1.92
Prenatal vitamin use 3 months after conception.	1.26	0.44	0.66	.514	0.63	2.51
Maternal Medical Risk Factors for Pregnancy						
Anemia	0.80	0.56	-0.31	.754	0.20	3.18
Lung Disease	1.99	1.89	0.72	.473	0.30	13.22
Diabetes	0.73	0.30	-0.76	.450	0.33	1.65
(Oligo)hydramnios	0.96	0.58	-0.06	.950	0.29	3.18
Hypertension (Chronic or gestational)	0.51	0.20	-1.71	.091	0.23	1.12
Eclampsia	3.15	2.84	1.27	.206	0.53	18.88
Previous > 4 kg birth	1.74	0.78	1.23	.222	0.71	4.23
Previous preterm birth	1.48	0.80	0.73	.465	0.51	4.32
Other medical risk factors	1.19	0.30	0.69	.492	0.72	1.97

Table 13 (continued).

Propensity Score Model Predicting Child Disability/Delay Status from Pre-birth Covariates.

Outcome: Child Disability/Delay Status	Odds Ratio	Linearized Std. Errors	*t*	*p*	95% Confidence Interval of Odds Ratio	
Other Risk Factors for Pregnancy						
Mother smoked during pregnancy	1.35	0.35	1.19	.238	0.82	2.25
Mother consumed alcohol during pregnancy	0.69	0.27	-0.93	.354	0.31	1.52
Preconception BMI: Underweight	1.34	0.49	0.79	.430	0.64	2.78
Preconception BMI: Overweight	1.20	0.21	1.05	.295	0.85	1.71
Preconception BMI: Obese	1.37	0.34	1.26	.211	0.83	2.26
Comparison group: Preconception BMI: Recommended weight						
Pregnancy weight gain: No gain or loss	1.37	0.74	0.59	.558	0.47	4.02
Pregnancy weight gain: Undergain	1.14	0.24	0.61	.542	0.75	1.74
Pregnancy weight gain: Overgain	1.60	0.30	2.49	.015	1.10	2.33
Comparison group: Pregnancy weight gain: Recommended weight gain						
Obstetric Procedures						
Amniocentesis	0.73	0.33	-0.70	.488	0.30	1.79
Fetal monitor	1.54	0.51	1.30	.196	0.80	2.97
Tocolysis	3.99	1.51	3.65	.000	1.88	8.48
Ultrasound	0.90	0.26	-0.37	.713	0.50	1.61
Complications of Labor and Delivery						
Febrile (> 100 degree fever)	1.49	0.97	0.62	.540	0.41	5.43
Membrane rupture > 12 hours	1.76	0.78	1.27	.207	0.73	4.25
Placenta previa or abruptio placenta	0.15	0.09	-3.11	.003	0.04	0.51
Fetal distress	2.83	1.16	2.53	.013	1.25	6.41
Other labor complications	1.00	0.26	0.00	.998	0.59	1.69

Table 13 (continued).

Propensity Score Model Predicting Child Disability/Delay Status from Pre-birth Covariates.

Outcome: Child Disability/Delay Status	Odds Ratio	Linearized Std. Errors	*t*	*p*	95% Confidence Interval of Odds Ratio	
Method of Delivery						
C-section	1.26	0.26	1.13	.261	0.84	1.90
Use of birth instruments	0.66	0.24	-1.15	.253	0.32	1.36
Comparison group: vaginal birth with no instruments						
Child Health Status at Birth						
5-minute Apgar score	0.73	0.11	-2.15	.034	0.54	0.98
Birth weight category: Extremely low	7.53	4.43	3.43	.001	2.34	24.22
Birth weight category: Very low	2.18	1.10	1.54	.128	0.80	5.96
Birth weight category: Moderately low	1.22	0.48	0.52	.605	0.56	2.66
Birth weight category: High	1.08	0.35	0.24	.809	0.57	2.06
Comparison group: Birth weight category: normal birth weight						
Child is preterm	1.36	0.45	0.92	.360	0.70	2.64
Child is small for gestational age (SGA)	1.06	0.32	0.20	.844	0.58	1.93
Child is large for gestational age (LGA)	0.98	0.33	-0.05	.957	0.50	1.91
Comparison group: Child is normal size for gestation age						
Abnormal Conditions of the Newborn						
Infant respiratory distress syndrome	0.56	0.34	-0.97	.334	0.17	1.85
Child needed assistant ventilation for < 30 min	1.39	0.72	0.63	.528	0.49	3.91
Child needed assistant ventilation for > 30 min	2.39	1.38	1.52	.132	0.76	7.50
Comparison group: Child did not need assisted ventilation						
Other abnormal newborn condition	1.55	0.59	1.15	.254	0.73	3.31

Table 13 (continued).

Propensity Score Model Predicting Child Disability/Delay Status from Pre-birth Covariates.

Outcome: Child Disability/Delay Status	Odds Ratio	Linearized Std. Errors	*t*	*p*	95% Confidence Interval of Odds Ratio	
Prenatal Father Involvement						
Discuss pregnancy	1.30	0.54	0.64	.524	0.57	2.95
See a sonogram or ultrasound of the baby	1.48	0.63	0.92	.361	0.63	3.46
Listen to the baby's heartbeat	1.51	0.66	0.94	.349	0.63	3.59
Feel the baby move	0.61	0.37	-0.81	.420	0.19	2.03
Attend childbirth classes or Lamaze classes	1.12	0.19	0.68	.497	0.80	1.58
Buy things for the child	1.23	0.48	0.54	.592	0.57	2.68
Present in the delivery room at birth	1.06	0.53	0.12	.908	0.39	2.87

Table 13 (continued).

Propensity Score Model Predicting Child Disability/Delay Status from Pre-birth Covariates.

Outcome: Child Disability/Delay Status	Odds Ratio	Linearized Std. Errors	*t*	*p*	95% Confidence Interval of Odds Ratio	
Missing Data Indicators						
Missing fathers' age	1.87	1.14	1.03	.308	0.56	6.28
Missing mothers' birth location (Foreign vs. US)	5.36	8.75	1.03	.307	0.21	137.66
Missing previous births	0.03	0.04	-2.75	.007	0.00	0.39
Missing terminations	34.10	30.14	3.99	.000	5.88	197.61
Missing Adequacy of Prenatal Care Utilization	0.82	0.43	-0.38	.704	0.29	2.32
Missing other medical risk factors	1.82	1.49	0.73	.469	0.35	9.31
Missing pregnancy weight gain	13.32	15.48	2.23	.028	1.32	134.19
Missing pregnancy preconception weight	0.11	0.11	-2.18	.032	0.01	0.82
Missing obstetrics	29.84	52.83	1.92	.058	0.88	1006.86
Missing labor and delivery complications	0.04	0.05	-2.78	.007	0.00	0.40
Missing fetal distress	0.75	0.36	-0.62	.540	0.29	1.93
Missing birth method	0.17	0.13	-2.40	.019	0.04	0.74
Missing 5-minute Apgar score	0.86	0.28	-0.48	.634	0.45	1.62
Missing child birth weight status	1.09	1.23	0.08	.940	0.12	10.20
Missing preterm status	0.47	0.65	-0.55	.586	0.03	7.26
Missing Intra-Uterine Growth status	1.19	1.26	0.16	.870	0.14	9.78
Missing newborn abnormal condition	0.62	0.51	-0.58	.563	0.12	3.21
Missing prenatal father involvement	0.42	0.54	-0.68	.499	0.03	5.30

Table 13 (continued).

Propensity Score Model Predicting Child Disability/Delay Status from Pre-birth Covariates.

Outcome: Child Disability/Delay Status	Odds Ratio	Linearized Std. Errors	*t*	*p*	95% Confidence Interval of Odds Ratio	
Interactions						
Hypertension X Mothers' age	1.07	0.06	1.22	.227	0.96	1.20
Preterm X Mothers' age	0.95	0.05	-1.07	.289	0.85	1.05
Highest Ed.: 12th grade or less X Mothers' age	0.93	0.12	-0.53	.597	0.72	1.21
Highest Ed.: High school diploma X Mothers' age	1.03	0.09	0.35	.725	0.87	1.22
Highest Ed.: Vocc./Tech prog. or some college X Mothers' age	0.95	0.07	-0.73	.466	0.82	1.09
Highest Ed.: Graduate or professional schooling X Mothers' age	1.01	0.08	0.13	.899	0.87	1.17
Highest Ed.: 12th grade or less X Fathers' age	0.97	0.09	-0.33	.743	0.80	1.17
Highest Ed.: High school diploma X Fathers' age	1.04	0.07	0.52	.607	0.91	1.18
Highest Ed.: Vocc./Tech prog. or some college X Fathers' age	1.03	0.06	0.61	.541	0.93	1.15
Highest Ed.: Graduate or professional schooling X Fathers' age	0.94	0.05	-1.22	.224	0.84	1.04
African American X Mothers' age	1.01	0.09	0.14	.892	0.84	1.22
Latino(a) X Mothers' age	1.06	0.09	0.74	.460	0.90	1.24
Asian or Pacific Islander X Mothers' age	1.03	0.08	0.33	.744	0.87	1.21
Native American X Mothers' age	1.16	0.15	1.15	.253	0.90	1.49
Multiple race/ethnicities X Mothers' age	1.05	0.07	0.78	.439	0.92	1.21
African American X Fathers' age	0.90	0.06	-1.46	.147	0.78	1.04
Latino(a) X Fathers' age	0.91	0.06	-1.55	.124	0.81	1.03
Asian or Pacific Islander X Fathers' age	0.95	0.07	-0.74	.460	0.82	1.09
Native American X Fathers' age	1.01	0.12	0.07	.947	0.80	1.27
Multiple race/ethnicities X Fathers' age	0.99	0.06	-0.21	.836	0.87	1.12
Constant	0.25	0.40	-0.88	.382	0.01	5.65

Figure 12.

Initial model kernel density estimates of propensity scores of children who are typically developing (CTYP) and children with disabilities/delays (CWD) using unmatched sample.

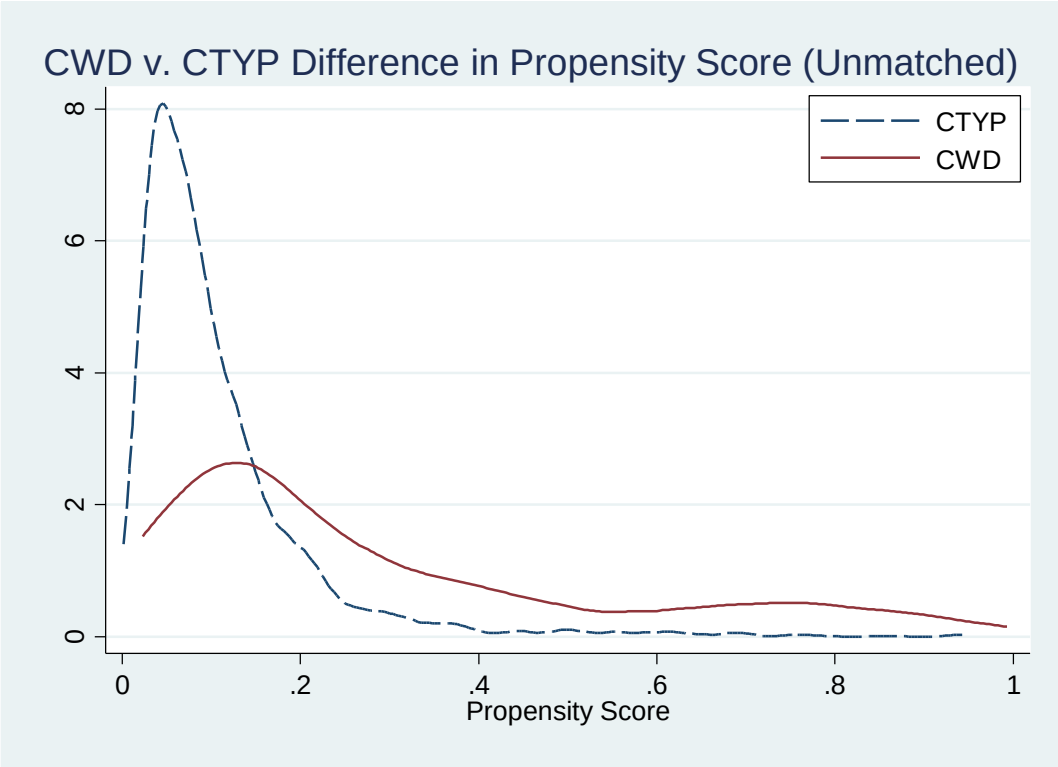
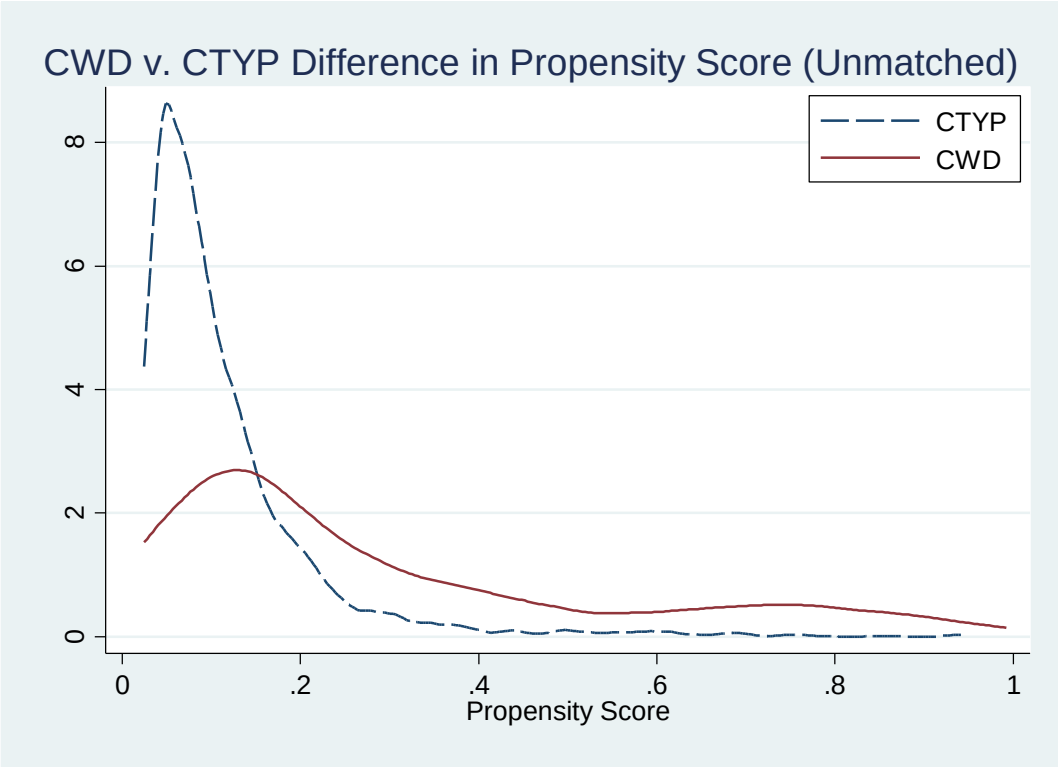


Figure 13.

Refit model kernel density estimates of propensity scores of children who are typically developing (CTYP) and children with disabilities/delays (CWD) using unmatched sample.



Matching on the propensity score. For this study, caliper matching with replacement was used. Caliper matching consists of matching each child with a disability/delay with every child in the typically developing group who has a propensity score within a certain range, or caliper. Rosenbaum and Rubin (1985) suggest that a caliper of 0.25 standard deviations may provide an effective range. The standard deviation of the propensity score in the refit model was 0.093. Children who were typically developing who were within 0.25*0.093 = 0.02 SD of the propensity score of a child with a disability/delay were eligible matches for that child. Matching with replacement means that a child who is typically developing can be used as a match more than once. Matching with replacement has the advantage over matching without replacement in that higher quality (i.e., closer) matches can be made for each child with a disability/delay (Abadie & Imbens, 2002). This is because all children who are typically developing are eligible to be used as a match rather than only those who have not already been matched.

All 2000 children in the typically developing group were an acceptable match for at least one child with a disability/delay. Of the children with a disability/delay, 10 children did not have any acceptable matches among the children who were typically developing. All 10 of these children had higher propensity scores as shown in Table 12. The effect of disability/delay status cannot be estimated for these children. However, these children represent small gaps in the matched range of the propensity score: 0.81 – 0.82, 0.89 – 0.91, and 0.97 – 0.99. As can be seen in Table 14, few children in the CTYP group had a propensity score above 0.75 and were available for matching. Because only a few cases were available, estimates may be less reliable for the effect of disability/delay status for children with propensity scores in the upper quartile and caution should be used in generalizing the results of this study to children with high propensity scores. Figures 14 displays the kernel density estimates of propensity scores of children who are typically developing (CTYP) and children with disabilities/delays (CWD) using the matched sample.

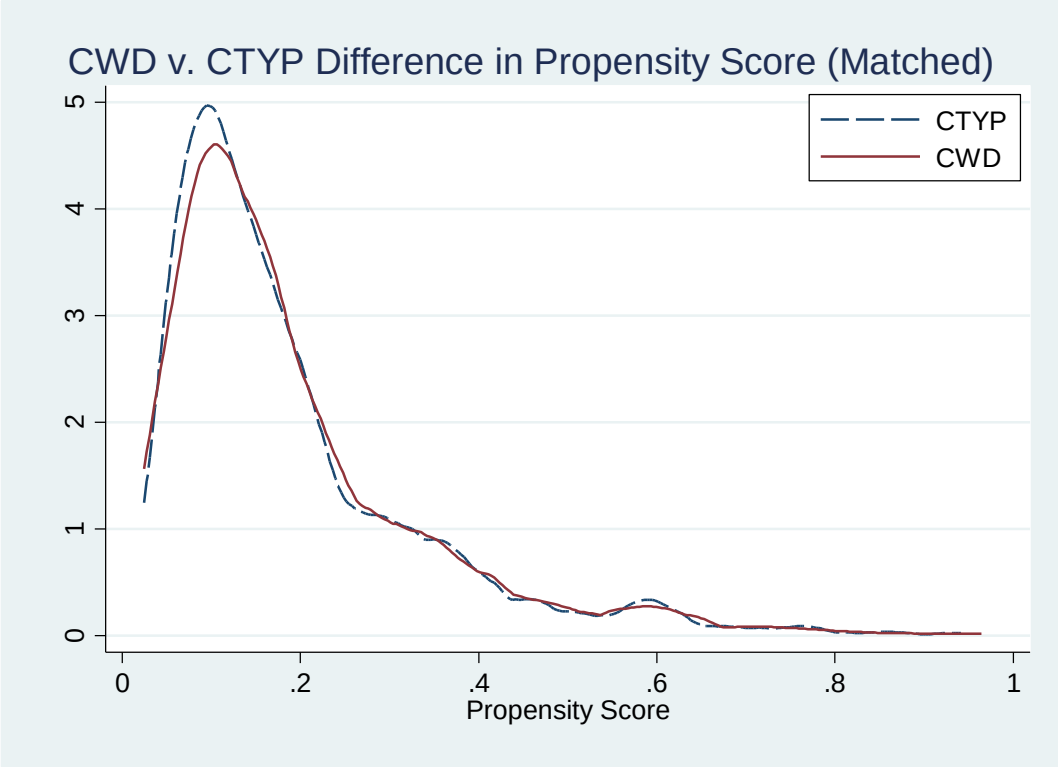
Table 14.

Propensity scores in the upper quartile by group.

CWD without an adequate match (N)	CWD with 1 or more adequate matches (N)	CTYP (N)
	0.75 (1)	0.75 (1)
	0.76 (1)	
	0.77 (1)	
	0.78 (3)	0.78 (1)
	0.80 (2)	
0.81 (4)		
0.82 (1)		
	0.83 (1)	
	0.85 (3)	
		0.86 (1)
	0.87 (2)	
	0.88 (1)	
0.89 (2)		
0.91 (1)		
	0.92 (1)	
		0.94 (2)
	0.95 (1)	
	0.96 (3)	
0.97 (1)		
0.99 (1)		

Figure 14.

Refit model kernel density estimates of propensity scores of children who are typically developing (CTYP) and children with disabilities/delays (CWD) using the matched sample.



Conditional independence assumption and balance checks. As previously noted, the conditional independence assumption holds that, conditional on **X** (i.e., the covariates used to create the propensity score), treatment assignment is random. This assumption can be partially tested by checking for balance on the covariates in **X** in the matched sample of children with disabilities/delays and their typically developing controls. Balance between groups on the covariates in **X** is necessary, but not sufficient for the conditional independence assumption to hold. Balance checking was done by testing for mean differences between children with disabilities and their typically developing match(es) on each covariate using the matched sample. For multi-category covariates, balance checks were accomplished by fitting a logistic regression model with child disability/delay status as the outcome and the multi-category covariate as the predictor using the matched sample. All possible comparisons between categories were made. Some imbalance on predictors might be expected when using a $p = .05$ cutoff because 1 in 20 covariates could have a significant difference just by chance. With 88 covariates in the model, 4 or 5 might be significantly different by chance. Chronic or gestational hypertension and parents' highest level of education were imbalanced in initial propensity score models. Additionally, child race/ethnicity, preterm birth status, placenta previa or abruption placenta, birth weight status, and the missing data indicator for birth weight status were imbalanced. The addition of higher-order interactions terms (hypertension X mothers' age, preterm status X mothers' age, parents' highest education X mothers' age, parents' highest education X fathers' age, child race/ethnicity X mothers' age, and child race/ethnicity X fathers' age) addressed the imbalance in parents' highest education and chronic or gestational hypertension. Imbalance in the other predictors was not corrected (see Tables 15 and 16). Although most differences were quite small, some differences were larger and the total number of significant differences was more than might be expected by chance. Because balance could not be achieved on these covariates, all analyses using the matched sample controlled for child race/ethnicity, preterm status, placenta previa or abruption placenta, birth weight status, and the missing data indictor for birth weight status.

Table 15.

Mean Difference in Imbalanced Covariates.

	CTYP *M*	CWD *M*	|M Difference|	*t*	*p*
Binary Variables					
Placenta previa or abruptio placenta	0.02	0.00	0.01	2.48	.015
Preterm	0.24	0.15	0.09	3.06	.003
Missing birth weight category	0.01	0.00	0.01	2.31	.023
Child Race/Ethnicity					
Caucasian, non-Latino(a)	0.63	0.76	0.13	3.68	.000
African American	0.04	0.03	0.01	0.57	.569
Latino(a)	0.15	0.17	0.02	0.62	.531
Asian/Pacific Islander	0.08	0.02	0.07	7.62	.000
Native American or Alaskan Native	0.03	0.00	0.03	3.30	.000
Multiple races/ethnicities	0.07	0.03	0.05	3.64	.001
Child Birth Weight Status					
Extremely low birth weight	0.03	0.01	0.02	1.82	.072
Very low birth weight	0.06	0.01	0.05	4.19	.000
Moderately low birth weight	0.11	0.06	0.05	2.47	.016
Normal birth weight	0.68	0.79	0.11	3.07	.003
High birth weight	0.12	0.12	0.00	0.14	.878

Note. Significance tests for multi-category variables of child race/ethnicity and child birth weight status refer to within category differences between groups.

Table 16.

Significant (Imbalanced) Multi-category Predictors of Disability/Delay Status .

Outcome: Child disability/delay status	*b*	*SE*	*t*	*p*
Child Race/Ethnicity				
African American	-0.46	0.53	-0.86	.392
Latino(a)	-0.08	0.20	-0.38	.701
Asian/Pacific Islander	-1.91	0.28	-6.88	.000
Native American or Alaskan Native	-2.27	0.38	-6.03	.000
Multiple races/ethnicities	-1.16	0.38	-3.00	.004
Comparison Group: Caucasian, non-Latino(a)				
Latino(a)	0.38	0.54	0.70	.484
Asian/Pacific Islander	-1.45	0.55	-2.66	.009
Native American or Alaskan Native	-1.81	0.59	-3.06	.003
Multiple races/ethnicities	-0.70	0.64	-1.08	.282
Comparison Group: African American				
Asian/Pacific Islander	-1.83	0.29	-6.27	.000
Native American or Alaskan Native	-2.19	0.36	-6.13	.000
Multiple races/ethnicities	-1.08	0.41	-2.60	.011
Comparison Group: Latino(a)				
Native American or Alaskan Native	-0.36	0.37	-0.96	.341
Multiple races/ethnicities	0.75	0.45	1.67	.099
Comparison Group: Asian or Pacific Islander				
Multiple races/ethnicities	1.11	0.50	2.21	.029
Comparison Group: Native American/Alaskan Native				
Child Birth Weight Status				
Very low birth weight	-0.83	0.46	-1.82	.073
Moderately low birth weight	0.19	0.39	0.48	.635
Normal birth weight	0.86	0.35	2.44	.017
High birth weight	0.76	0.40	1.90	.061
Comparison Group: Extremely low birth weight				
Moderately low birth weight	1.02	0.39	2.63	.010
Normal birth weight	1.69	0.30	5.69	.000
High birth weight	1.59	0.37	4.28	.000
Comparison Group: Low birth weight				
Normal birth weight	0.68	0.27	2.52	.014
High birth weight	0.57	0.33	1.71	.091
Comparison Group: Moderately low birth weight				
High birth weight	-0.11	0.26	-0.41	.685
Comparison Group: Normal birth weight				

Results by Question

Results for Question 1: Effect of child disability/delay status on father involvement. For the first question, the causal effect of disability/delay status was estimated using the matched sample and fitting a regression model with father involvement as the outcome variables and child disability/delay status as the predictor variable. To compare approaches, four different models were fit. First, a regression model was fit using the unmatched sample with a single predictor of father involvement: child disability/delay status. Second, a regression model was fit again using the unmatched sample predicting father involvement from child disability/delay status, but for this model all of the controls used to create the propensity score were included. This model follows the traditional regression/covariate adjustment approach used with most observational/correlational studies. Third, a regression model was fit using the matched sample with father involvement predicted by child disability/delay status. Because there was imbalance on some covariates after matching on the propensity score, those covariates were included as control variables. Fourth, because it is possible that there may be some undetected imbalance on some of the covariates, the propensity score is added to the third model as an additional control. By adding the propensity score to the fourth model, a more conservative estimate of the effect of disability/delay status on father involvement is obtained. These four models were fit for each type of father involvement at each time point.

The results are reported in Table 17. In terms of significant findings, the traditional regression/covariate adjustment approach tends to give the same results (i.e. the effect of disability/delay status is significant for the same types of involvement), though the estimated size of the effect may vary. It also appears that there may be some residual imbalance on the propensity score that once controlled for (in the fourth model) increases the size of the estimated effect. As can be seen in Table 17 using the most conservative model (the fourth model), child disability/delay status has a negative effect on father literacy involvement at 2 years, father routine caregiving and play involvement at 9 months and 2 years, and has no effect on father responsive caregiving involvement.

Table 17.

Estimates of the Effect of Child Disability/Delay Status on Father Involvement.

Outcome: Father Literacy inv. at 9 months	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.02	-0.02	0.08	.800
Unmatched, OLS regression + controls	-0.06	-0.06	0.10	.500
Matched, OLS regression + unbalanced controls	-0.07	-0.06	0.08	.456
Matched, OLS regression + unbalanced controls + p-score	-0.07	-0.07	0.09	.423
Outcome: Father Literacy inv. at 2 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.20	-0.21	0.11	.067
Unmatched, OLS regression + controls	-0.24	-0.21	0.11	.036
Matched, OLS regression + unbalanced controls	-0.26	-0.22	0.10	.012
Matched, OLS regression + unbalanced controls + p-score	-0.25	-0.22	0.10	.018
Outcome: Father Literacy inv. at 4 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	0.01	0.01	0.09	.934
Unmatched, OLS regression + controls	-0.05	-0.05	0.09	.612
Matched, OLS regression + unbalanced controls	-0.03	-0.02	0.07	.808
Matched, OLS regression + unbalanced controls + p-score	-0.01	-0.03	0.07	.737
Outcome: Father Routine Caregiving inv. at 9 months	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.21	-0.11	0.16	.189
Unmatched, OLS regression + controls	-0.35	-0.18	0.17	.046
Matched, OLS regression + unbalanced controls	-0.25	-0.13	0.15	.090
Matched, OLS regression + unbalanced controls + p-score	-0.32	-0.17	0.15	.036
Outcome: Father Routine Caregiving inv. at 2 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.28	-0.14	0.18	.124
Unmatched, OLS regression + controls	-0.45	-0.22	0.19	.022
Matched, OLS regression + unbalanced controls	-0.30	-0.14	0.17	.089
Matched, OLS regression + unbalanced controls + p-score	-0.35	-0.17	0.17	.045
Outcome: Father Routine Caregiving inv. at 4 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.29	-0.15	0.18	.121
Unmatched, OLS regression + controls	-0.36	-0.19	0.19	.053
Matched, OLS regression + unbalanced controls	-0.28	-0.15	0.19	.140
Matched, OLS regression + unbalanced controls + p-score	-0.29	-0.15	0.18	.119

Note. cwd = child disability/delay status. β_{StdY} = Standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP.

Table 17 (continued).

Estimates of the Effect of Child Disability/Delay Status on Father Involvement.

Outcome: Father Play inv. at 9 months	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.26	-0.22	0.12	.028
Unmatched, OLS regression + controls	-0.36	-0.31	0.11	.002
Matched, OLS regression + unbalanced controls	-0.28	-0.24	0.11	.009
Matched, OLS regression + unbalanced controls + p-score	-0.31	-0.26	0.11	.005

Outcome: Father Play inv. at 2 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.19	-0.11	0.14	.178
Unmatched, OLS regression + controls	-0.30	-0.18	0.14	.038
Matched, OLS regression + unbalanced controls	-0.24	-0.14	0.13	.066
Matched, OLS regression + unbalanced controls + p-score	-0.26	-0.16	0.13	.042

Outcome: Father Play inv. at 4 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	0.09	0.07	0.11	.424
Unmatched, OLS regression + controls	-0.07	-0.05	0.12	.554
Matched, OLS regression + unbalanced controls	0.02	0.02	0.10	.850
Matched, OLS regression + unbalanced controls + p-score	-0.01	-0.01	0.11	.939

Outcome: Father Responsive Caregiving inv. at 9 months	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.02	-0.06	0.02	.433
Unmatched, OLS regression + controls	-0.02	-0.09	0.02	.321
Matched, OLS regression + unbalanced controls	-0.00	-0.01	0.02	.924
Matched, OLS regression + unbalanced controls + p-score	-0.01	-0.04	0.02	.577

Outcome: Father Responsive Caregiving inv. at 2 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.01	-0.03	0.03	.714
Unmatched, OLS regression + controls	-0.01	-0.04	0.03	.647
Matched, OLS regression + unbalanced controls	0.00	0.04	0.03	.674
Matched, OLS regression + unbalanced controls + p-score	0.00	0.01	0.03	.914

Note. cwd = child disability/delay status. β_{StdY} = Standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP.

In addition to examining the direct effects of child disability/delay status on father involvement, mediation analyses were conducted to explore whether the effect of child disability/delay status on later father involvement was mediated by earlier father involvement. The following mediation models were examined for each type of father involvement:

1. Child disability/delay status → Father involvement (9 months) → Father involvement (2 years)
2. Child disability/delay status → Father involvement (9 months) → Father involvement (4 years)
3. Child disability/delay status → Father involvement (2 years) → Father involvement (4 years)

Mediation models with significant indirect effects are reported in Tables 18 – 24. In interpreting mediation models, I will refer to the paths in Figure 10. Also, the terms direct and total effects can be confusing because both terms refer to the effect the IV has *directly* on the DV. The only difference is that the total effect (Path C in Figure 10) is the estimated effect of the IV on the DV when the MV *is not* included in the model while the direct effect (Path C' in Figure 10) is the estimated effect when the MV *is* included in the model. To avoid confusion, I will also note the relevant Path from Figure 10.

Significant indirect effects were detected for routine caregiving, play, and literacy involvement, but not for responsive caregiving involvement. The effect of disability/delay status on routine caregiving involvement at 2 years was mediated by routine caregiving involvement at 9 months (see Table 18). Similarly, the effect of disability/delay status on routine caregiving involvement at 4 years was mediated by the routine caregiving at 9 months (see Table 20) and 2 years (see Table 23). In all three models, child disability/delay status was linked to lower father routine caregiving involvement at 9 months or 2 years. Because father involvement was strongly and positively correlated across time, the lower levels of father involvement at 9 months and 2 years were carried forward resulting in lower levels of father involvement at 2 years or 4 years. Thus the indirect effect of child disability/delay status on later father involvement was negative, meaning that child disability/delay status was linked to lower levels of later father involvement through earlier father involvement. For these three models, the effect of child/disability delay status on later father routine caregiving involvement decreased after earlier routine caregiving father involvement was included in the model (i.e., Path C' < Path C in Figure 10). This is known as consistent mediation. Based on the recommendations of Rucker and colleagues (2011), I do not classify the first model as full mediation although the effect of disability/delay status on father routine caregiving involvement at 2 years becomes non-significant once father involvement at 9 months is added. They caution against using the term full mediation because, as they demonstrated, other significant indirect effects can still be detected when the direct effect is not significant. The non-significance may be due to low power or unobserved suppression effects. Of interest, as reported in Table 17, child disability/delay status did not have a significant total effect on father routine caregiving involvement at 4 years (Path C in Figure 10), but the results of the mediation analyses indicate that child disability/delay status did have a

significant indirect effect through earlier father involvement leading to lower levels of routine caregiving involvement at 4 years (Paths A and B in Figure 10). There was not a direct effect (Path C' in Figure 10) of child disability/delay status on father routine caregiving involvement at 2 years or 4 years.

The effect of child disability/delay status on play involvement at 2 years was mediated by play involvement at 9 months (see Table 19) in that child disability/delay status was linked to lower father involvement at 9 months which was, in turn, linked to father involvement at 2 years. Similarly, there was a significant indirect effect of disability/delay status on play involvement at 4 years through play involvement at 9 months (see Table 21) and 2 years (see Table 24). The same pattern observed for routine caregiving involvement was repeated for play involvement. Specifically, child disability/delay status had a negative effect on earlier father play involvement (Path A in Figure 10). Because father play involvement was positively correlated across time (Path B), the lower levels of father play involvement resulting from child disability/delay status were carried forward. Thus, child disability/delay status was indirectly linked to lower levels of father play involvement at 2 years and 4 years through earlier father play involvement (Paths A and B). In regard to the total and direct effects of child disability/delay status on later father play involvement, the effect of child/disability delay status on father play involvement at 2 years decreased after earlier father play involvement was included in the model (Path C' < Path C). For the models predicting father involvement at 4 years, the effect of disability/delay status *increased* when earlier father play involvement at 9 months or 2 years was included in the model, though the effect of disability/delay status did not reach statistical significance (Path C' > Path C). MacKinnon and colleagues (2000) suggest that when the indirect effect (Path AB in Figure 10) is opposite in sign to the total effect (Path C) inconsistent mediation (suppression) has occurred. They do not provide guidance when the total effect is close to 0 and the direction of the effect, or sign, is thus not clear. However, given that the effect of disability/delay status increased upon adding earlier father involvement, inconsistent mediation is probably the best description. Of interest, child disability/delay status did not have a significant total effect (Path C) on father play involvement at 4 years (see Table 17). However, the results of the mediation analyses indicate that child disability/delay status did have a significant indirect effect through earlier father involvement leading to lower levels of play involvement at 4 years (Paths A and B). There were no significant direct effects (Path C') of child disability/delay status on father play involvement at 2 years and 4 years.

The effect of child disability/delay status on literacy involvement at 4 years was mediated by literacy involvement at 2 years (see Table 22). Child disability/delay status was linked to a lower father literacy involvement at 2 years (Path A, Figure 10). Because father literacy involvement was correlated across time (Path B), these lower levels of involvement were carried forward leading to lower levels of father literacy involvement at 4 years (Path B). Thus, there was a negative indirect effect of child

disability/delay status on father involvement at 4 years (Paths A and B). In regard to total and direct effects, the total effect of child/disability delay status on father literacy involvement at 4 years (Path C; see Table 17) increased and became a significant direct effect (Path C', see Table 22) after earlier father literacy involvement was included in the model.

Table 18.

Mediation of the effect of disability/delay status on father routine caregiving involvement at 2 years by father routine caregiving involvement at 9 months.

Regression model with DV regressed on IV: Path C					
Father Routine Caregiving (2 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.35	-0.17	0.16	-2.15	.034

Regression model with M regressed on IV: Path A					
Father Routine Caregiving (9 months)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.33	-0.17	0.15	-2.24	.028

Regression model with DV regressed on M and IV: Paths B and C'					
Father Routine Caregiving (2 years)	*b*	β	*SE*	*t*	*p*
Father Routine Caregiving (9 months)	0.75	0.70	0.03	26.21	.000
Child disability/delay status	-0.10	-0.05	0.12	-0.85	.397

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	-0.25	-0.12	0.11	-2.23	.026
Goodman-1 (Aroian)	-0.25	-0.12	0.11	-2.23	.026
Goodman-2	-0.25	-0.12	0.11	-2.23	.026

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	-0.25	0.00	0.11	-0.46	-0.04

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.33	-0.17	0.15	-2.24	.025
B coefficient	0.75	0.70	0.03	26.21	.000
Indirect Effect	-0.25	-0.12	0.11	-2.23	.026
Direct Effect	-0.10	-0.05	0.12	-0.85	.395
Total Effect	-0.35	-0.17	0.16	-2.15	.032
Proportion of total effect that is mediated:				0.71	
Ratio of indirect to direct effect:				2.51	
Ratio of total to direct effect:				3.51	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Table 19.

Mediation of the effect of disability/delay status on father play involvement at 2 years by father play involvement at 9 months.

Regression model with DV regressed on IV: Path C					
Father Play Involvement (2 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.28	-0.17	0.12	-2.24	.027

Regression model with M regressed on IV: Path A					
Father Play Involvement (9 months)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.31	-0.27	0.10	-3.13	.002

Regression model with DV regressed on M and IV: Paths B and C'					
Father Play Involvement (2 years)	*b*	β	*SE*	*t*	*p*
Father Play Involvement (9 months)	0.93	0.65	0.05	18.94	.000
Child disability/delay status	0.01	0.01	0.10	0.12	.905

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	-0.29	-0.17	0.09	-3.09	.002
Goodman-1 (Aroian)	-0.29	-0.17	0.09	-3.08	.002
Goodman-2	-0.29	-0.17	0.09	-3.09	.002

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	-0.29	0.00	0.08	-0.46	-0.12

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.31	-0.27	0.10	-3.13	.002
B coefficient	0.93	0.65	0.05	18.94	.000
Indirect Effect	-0.29	-0.17	0.09	-3.09	.002
Direct Effect	0.01	0.01	0.10	0.12	.905
Total Effect	-0.28	-0.17	0.12	-2.24	.025
Proportion of total effect that is mediated:				1.04	
Ratio of indirect to direct effect:				-24.97	
Ratio of total to direct effect:				-23.97	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Table 20.

Mediation of the effect of disability/delay status on father routine caregiving involvement at 4 years by father routine caregiving involvement at 9 months.

Regression model with DV regressed on IV: Path C					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.26	-0.14	0.17	-1.52	.131
Regression model with M regressed on IV: Path A					
Father Routine Caregiving (9 months)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.33	-0.17	0.15	-2.24	.028
Regression model with DV regressed on M and IV: Paths B and C'					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Father Routine Caregiving (9 months)	0.55	0.56	0.03	16.38	.000
Child disability/delay status	-0.08	-0.04	0.14	-0.58	.566

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	-0.18	-0.10	0.08	-2.22	.027
Goodman-1 (Aroian)	-0.18	-0.10	0.08	-2.21	.027
Goodman-2	-0.18	-0.10	0.08	-2.22	.027

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	-0.18	0.00	0.08	-0.33	-0.02

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.33	-0.17	0.15	-2.24	.025
B coefficient	0.55	0.56	0.03	16.38	.000
Indirect Effect	-0.18	-0.10	0.08	-2.21	.027
Direct Effect	-0.08	-0.04	0.14	-0.58	.564
Total Effect	-0.26	-0.14	0.17	-1.52	.127
Proportion of total effect that is mediated:				0.70	
Ratio of indirect to direct effect:				2.30	
Ratio of total to direct effect:				3.30	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Table 21.

Mediation of the effect of disability/delay status on father play involvement at 4 years by father play involvement at 9 months.

Regression model with DV regressed on IV: Path C					
Father Play Involvement (4 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.02	-0.02	0.10	-0.26	.799
Regression model with M regressed on IV: Path A					
Father Play Involvement (9 months)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.31	-0.27	0.10	-3.13	.002
Regression model with DV regressed on M and IV: Paths B and C'					
Father Play Involvement (4 years)	*b*	β	*SE*	*t*	*p*
Father Play Involvement (9 months)	0.53	0.45	0.08	6.86	.000
Child disability/delay status	0.14	0.10	0.10	1.41	.163

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	-0.17	-0.12	0.06	-2.85	.004
Goodman-1 (Aroian)	-0.17	-0.12	0.06	-2.82	.005
Goodman-2	-0.17	-0.12	0.06	-2.87	.004

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	-0.16	-0.00	0.05	-0.27	-0.07

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.31	-0.27	0.10	-3.13	.002
B coefficient	0.53	0.45	0.08	6.86	.000
Indirect Effect	-0.16	-0.12	0.06	-2.85	.004
Direct Effect	0.14	0.10	0.10	1.41	.160
Total Effect	-0.03	-0.02	0.10	0.26	.799
Proportion of total effect that is mediated:				6.24	
Ratio of indirect to direct effect:				-1.19	
Ratio of total to direct effect:				-0.19	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Table 22.

Mediation of the effect of disability/delay status on father literacy involvement at 4 years by father literacy involvement at 2 years.

Regression model with DV regressed on IV: Path C					
Father Literacy Involvement (4 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	0.01	0.01	0.07	0.12	.908
Regression model with M regressed on IV: Path A					
Father Literacy Involvement (2 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.22	-0.19	0.10	-2.27	.026
Regression model with DV regressed on M and IV: Paths B and C'					
Father Literacy Involvement (4 years)	*b*	β	*SE*	*t*	*p*
Father Literacy Involvement (2 years)	0.65	0.73	0.02	31.97	.000
Child disability/delay status	0.15	0.15	0.04	3.48	.001

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	-0.14	-0.14	0.06	-2.26	.024
Goodman-1 (Aroian)	-0.14	-0.14	0.06	-2.26	.024
Goodman-2	-0.14	-0.14	0.06	-2.26	.024

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	-0.14	0.00	0.06	-0.25	-0.02

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.22	-0.19	0.10	-2.27	.023
B coefficient	0.65	0.73	0.02	31.97	.000
Indirect Effect	-0.14	-0.14	0.06	-2.26	.024
Direct Effect	0.15	0.15	0.04	3.48	.000
Total Effect	0.01	0.01	0.07	0.12	.910
Proportion of total effect that is mediated:				-17.82	
Ratio of indirect to direct effect:				-0.95	
Ratio of total to direct effect:				0.05	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Table 23.

Mediation of the effect of disability/delay status on father routine caregiving involvement at 4 years by father routine caregiving involvement at 2 years.

Regression model with DV regressed on IV: Path C					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.25	-0.13	0.17	-1.47	.145

Regression model with M regressed on IV: Path A					
Father Routine Caregiving (2 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.34	-0.17	0.16	-2.09	.039

Regression model with DV regressed on M and IV: Paths B and C'					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Father Routine Caregiving (2 years)	0.59	0.64	0.03	19.73	.000
Child disability/delay status	-0.05	-0.03	0.12	-0.44	.660

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	-0.20	-0.11	0.10	-2.08	.038
Goodman-1 (Aroian)	-0.20	-0.11	0.10	-2.08	.038
Goodman-2	-0.20	-0.11	0.10	-2.08	.038

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	-0.20	0.01	0.09	-0.35	-0.03

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.34	-0.17	0.16	-2.09	.037
B coefficient	0.59	0.64	0.03	19.73	.000
Indirect Effect	-0.20	-0.11	0.10	2.08	.038
Direct Effect	-0.05	-0.03	0.12	-0.44	.659
Total Effect	-0.25	-0.13	0.17	-1.47	.141
Proportion of total effect that is mediated:				0.80	
Ratio of indirect to direct effect:				3.89	
Ratio of total to direct effect:				4.89	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Table 24.

Mediation of the effect of disability/delay status on father play involvement at 4 years by father play involvement at 2 years.

Regression model with DV regressed on IV: Path C					
Father Play Involvement (4 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.02	-0.01	0.10	-0.18	.855

Regression model with M regressed on IV: Path A					
Father Play Involvement (2 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.27	-0.16	0.12	-2.19	.031

Regression model with DV regressed on M and IV: Paths B and C'					
Father Play Involvement (4 years)	*b*	β	*SE*	*t*	*p*
Father Play Involvement (2 years)	0.53	0.65	0.03	15.42	.000
Child disability/delay status	0.13	0.09	0.07	1.70	.092

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	-0.14	-0.11	0.07	-2.17	.030
Goodman-1 (Aroian)	-0.14	-0.11	0.07	-2.16	.031
Goodman-2	-0.14	-0.11	0.07	-2.17	.030

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	-0.14	-0.00	0.07	-0.28	-0.02

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.27	-0.16	0.12	-2.19	.029
B coefficient	0.53	0.65	0.03	15.42	.000
Indirect Effect	-0.14	-0.10	0.07	-2.17	.030
Direct Effect	0.13	0.09	0.07	1.70	.089
Total Effect	-0.02	-0.01	0.10	-0.18	.854
Proportion of total effect that is mediated:				7.62	
Ratio of indirect to direct effect:				-1.15	
Ratio of total to direct effect:				-0.15	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Results for Question 2: Association between children's functional abilities and father involvement. The association between children's functional abilities and later father involvement was examined by fitting regression models using the matched sample controlling for the propensity score, unbalanced covariates, and family SES. Models were fit with and without prior father involvement as an additional control. Children's functional abilities at 9 months were not associated with father involvement at 2 years or at 4 years, with or without controlling for prior levels of father involvement. However, children's functional abilities at 2 years were associated with lower levels of father literacy and routine caregiving, but not play involvement at 4 years (see Table 25). Children's functional abilities at 2 years were only associated with father literacy involvement at 4 years when controlling for prior levels of literacy involvement.

Table 25.

Association between Children's Functional Abilities at 2 Years and Father Involvement at 4 Years.

Outcome: Father Literacy Involvement at 4 Years					
Model without prior father involvement	*b*	β	*SE*	*t*	*p*
Functional Abilities Index 2Y	-0.01	-0.01	0.03	-0.22	.830
Model with prior father involvement	*b*	β	*SE*	*t*	*p*
Functional Abilities Index 2Y	-0.05	-0.65	0.02	-2.12	.037
Father literacy involvement 2Y	0.65	0.73	0.02	30.57	.000

Outcome: Father Routine Caregiving Involvement at 4 Years					
Model without prior father involvement	*b*	β	*SE*	*t*	*p*
Functional Abilities Index 2Y	-0.24	-0.16	0.07	-3.43	.001
Model with prior father involvement	*b*	β	*SE*	*t*	*p*
Functional Abilities Index 2Y	-0.14	-0.09	0.06	-2.47	.016
Father routine caregiving involvement 2Y	0.58	0.63	0.03	19.55	.000

Outcome: Father Play Involvement at 4 Years					
Model without prior father involvement	*b*	β	*SE*	*t*	*p*
Functional Abilities Index 2Y	-0.03	-0.03	0.04	-0.74	.464
Model with prior father involvement	*b*	β	*SE*	*t*	*p*
Functional Abilities Index 2Y	-0.03	-0.03	0.04	-0.79	.432
Father play involvement 2Y	0.53	0.64	0.03	15.31	.000

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement (X).

Results for Question 3: Mediation of the effect of disability/delay status on father involvement by children's functional abilities. Analyses examining the mediation of the effect of disability/delay status on later father involvement by children's functional abilities were conducted using the matched sample and controlling for unbalanced covariates, the propensity score, and family SES. Models were fit both with and without prior levels of father involvement as a control variable. Part of the mediation analyses consists of estimating the effect of child disability/delay status on children's functional abilities. Because these causal estimates can be obtained and are of interest, I fit each of the four models described for the Question 1 results to compare approaches to estimating the effect of child disability/delay status on children's functional abilities at 9 months, 2 years, and 4 years. Results of these analyses are reported in Table 26. The four models are consistent in their estimates of the effect of disability/delay status on children's functional abilities and that the size of the effect increases across time.

The following mediation models were fit separately for each type of father involvement:

1. Child disability/delay status → Functional abilities (9 months) → Father involvement (2 years)
2. Child disability/delay status → Functional abilities (9 months) → Father involvement (4 years)
3. Child disability/delay status → Functional abilities (2 years) → Father involvement (4 years)

For the mediation analyses, significant indirect effects were detected only for the mediation of the effect of child disability/delay status on father routine caregiving involvement at 4 years by children's functional abilities at 2 years. Both the model with and the model without prior routine caregiving involvement as a control indicate the same general pattern (see Tables 27 and 28): child disability/delay status was linked to lower levels of functional abilities at 2 years (Path A in Figure 10). Children's functional abilities were negatively associated with father routine caregiving involvement at 4 years, meaning that fathers were more involved with children who with lower levels of functional abilities. Thus, the indirect effect of child disability/delay status on father routine caregiving involvement at 4 years through children's functional abilities at 2 years is an increase in father involvement. Additionally, in the model controlling for prior involvement, the direct effect (Path C') of child disability/delay status on father routine caregiving involvement at 4 years was not significant. When prior father involvement was not controlled for, the direct effect was significant. In the analyses for Question 1, the total effect (Path C) of disability/delay status on father involvement had not been significant indicating a suppression effect or inconsistent mediation.

In regard to father literacy involvement, a change in the significance of a path should be noted. By including child disability/delay status in the model predicting father literacy involvement at 4 years from children's functional abilities at 2 years and controlling for father literacy involvement at 2 years, children's functional abilities were no longer significantly associated with father literacy involvement. In

the analysis of this association for Question 2 without controlling for child disability/delay status, children's functional abilities at 2 years had been significantly associated with father literacy involvement at 4 years.

Table 26.

Estimates of the Effect of Child Disability/Delay Status on Functional Abilities.

Children's Functional Abilities at 9 months	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.50	-0.44	0.09	.000
Unmatched, OLS regression + controls	-0.44	-0.38	0.08	.000
Matched, OLS regression + unbalanced controls	-0.51	-0.45	0.08	.000
Matched, OLS regression + unbalanced controls + p-score	-0.49	-0.43	0.08	.000

Children's Functional Abilities at 2 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-0.70	-0.56	0.09	.000
Unmatched, OLS regression + controls	-0.67	-0.54	0.08	.000
Matched, OLS regression + unbalanced controls	-0.76	-0.61	0.09	.000
Matched, OLS regression + unbalanced controls + p-score	-0.74	-0.59	0.09	.000

Children's Functional Abilities at 4 years	*b*(cwd)	β_{StdY}(cwd)	*SE*	*p*
Unmatched, OLS regression	-1.00	-0.68	0.12	.000
Unmatched, OLS regression + controls	-0.98	-0.67	0.11	.000
Matched, OLS regression + unbalanced controls	-1.04	-0.72	0.12	.000
Matched, OLS regression + unbalanced controls + p-score	-1.01	-0.70	0.12	.000

Note. cwd = child disability/delay status. β_{StdY} = Standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP.

Table 27.

Mediation of the effect of disability/delay status on father routine caregiving involvement at 4 years by children's functional abilities at 2 years: Not controlling for prior father involvement.

Regression model with DV regressed on IV: Path C					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.25	-0.13	0.17	-1.47	.145

Regression model with M regressed on IV: Path A					
Children's FAI (2 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.71	-0.57	0.09	-7.87	.000

Regression model with DV regressed on M and IV: Paths B and C'					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Children's FAI (2 years)	-0.29	-0.19	0.07	-4.01	.000
Child disability/delay status	-0.46	-0.24	0.17	-2.66	.009

Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	0.21	0.11	0.06	3.57	.000
Goodman-1 (Aroian)	0.21	0.11	0.06	3.55	.000
Goodman-2	0.21	0.11	0.06	3.60	.000

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	0.21	-0.00	0.06	0.10	0.32

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.71	-0.57	0.09	-7.87	.000
B coefficient	-0.29	-0.19	0.07	-4.01	.000
Indirect Effect	0.21	0.11	0.06	3.57	.000
Direct Effect	-0.46	-0.24	0.17	-2.66	.008
Total Effect	-0.25	-0.13	0.17	-1.47	.141
Proportion of total effect that is mediated:				-0.82	
Ratio of indirect to direct effect:				-0.45	
Ratio of total to direct effect:				0.55	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement or children's FAI, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement/children's FAI (X).

Table 28.

Mediation of the effect of disability/delay status on father routine caregiving involvement at 4 years by children's functional abilities at 2 years: Controlling for prior father involvement.

Regression model with DV regressed on IV: Path C					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.05	-0.03	0.12	-0.44	.660
Father routine caregiving (2 years)	0.59	0.64	0.03	19.73	.000
Regression model with M regressed on IV: Path A					
Children's FAI (2 years)	*b*	β	*SE*	*t*	*p*
Child disability/delay status	-0.74	-0.59	0.09	-8.02	.000
Father routine caregiving (2 years)	-0.07	-0.12	0.03	-2.63	.010
Regression model with DV regressed on M and IV: Paths B and C'					
Father Routine Caregiving (4 years)	*b*	β	*SE*	*t*	*p*
Children's FAI (2 years)	-0.17	-0.11	0.06	-2.75	.007
Child disability/delay status	-0.17	-0.09	0.12	-1.50	.138
Father routine caregiving (2 years)	0.58	0.63	0.03	19.07	.000
Sobel-Goodman Mediation Tests					
	b	β	*SE*	*t*	*p*
Sobel	0.12	0.06	0.05	2.60	.009
Goodman-1 (Aroian)	0.12	0.06	0.05	2.59	.010
Goodman-2	0.12	0.06	0.05	2.62	.009

Preacher-Hayes Bootstrap Results for Indirect Effect					
	Observed *b*	Bias	Bootstrap *SE*	95% Conf. Interval	
Indirect Effect	0.12	-0.00	0.05	0.04	0.21

Summary of Effects					
	b	β	*SE*	*t*	*p*
A coefficient	-0.74	-0.59	0.09	-8.02	.000
B coefficient	-0.17	-0.11	0.06	-2.75	.006
Indirect Effect	0.12	0.06	0.05	2.60	.009
Direct Effect	-0.17	-0.09	0.12	-1.50	.134
Total Effect	-0.05	-0.03	0.12	-0.44	.659
Proportion of total effect that is mediated:				-2.34	
Ratio of indirect to direct effect:				-0.70	
Ratio of total to direct effect:				0.30	

Note. Not shown are control variables. If predictor variable in regression model is D/D status, $\beta = \beta_{StdY}$ = standardized difference in outcome (Y) associated with a 1-unit difference in disability/delay status (X), or the mean standardized difference between CWD and CTYP. If predictor variable in regression model is father involvement or children's FAI, $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement/children's FAI (X).

Results for Question 4: Association between father involvement and children's functional abilities and mediation of the association between child disability/delay status and children's functional abilities by father involvement. The association between father involvement and children's later functional abilities was examined by fitting a regression model using the matched sample and controlling for the propensity score, imbalanced covariates, family SES, and child disability/delay status. These models were fit separately for each type of father involvement with and without controlling for prior levels of children's functional abilities. Mediation of the effect of child disability/delay status on functional abilities by father involvement was explored using the same controls both with and without prior levels of children's functional abilities. For this question, the following mediation models were fit separately for each type of father involvement:

1. Child disability/delay status → Father involvement (9 months) → Functional abilities (2 years)
2. Child disability/delay status → Father involvement (9 months) → Functional abilities (4 years)
3. Child disability/delay status → Father involvement (2 years) → Functional abilities (4 years)

Mediation analyses found no significant indirect effects in which father involvement mediated the effect of child disability/delay status on later functional abilities. Significant associations between father involvement and children's later functional abilities were found for father routine caregiving and responsive caregiving involvement as reported in Tables 29 and 30, respectively. Specifically, father routine and responsive caregiving involvement at 9 months were associated with lower levels of children's functional abilities at 2 years and at 4 years. This was true for both the models controlling for children's prior functional abilities and the models not controlling for it. In regard to 2 year predictors, only father routine caregiving involvement was associated with lower levels of children's functional abilities at 4 years. However, this association was not significant in the model controlling for children's functional abilities at 2 years.

Table 29.

Association between Father Routine Caregiving Involvement and Children's Later Functional Abilities.

Outcome: Children's Functional Abilities at 2 Years					
Regression model without prior FAI	*b*	β	*SE*	*t*	*p*
Father routine caregiving involvement 9M	-0.08	-0.13	0.03	-3.20	.002
Regression model with prior FAI	*b*	β	*SE*	*t*	*p*
Father routine caregiving involvement 9M	-0.09	-0.14	0.03	-3.30	.001
Functional Abilities Index 9M	0.29	0.26	0.05	6.02	.000

Outcome: Children's Functional Abilities at 4 Years					
Regression model without prior FAI	*b*	β	*SE*	*t*	*p*
Father routine caregiving involvement 9M	-0.08	-0.11	0.03	-2.71	.008
Regression model with prior FAI	*b*	β	*SE*	*t*	*p*
Father routine caregiving involvement 9M	-0.08	-0.09	0.03	-2.70	.008
Functional Abilities Index 9M	0.21	0.23	0.07	3.19	.002

Outcome: Children's Functional Abilities at 4 Years					
Regression model without prior FAI	*b*	β	*SE*	*t*	*p*
Father routine caregiving involvement 2Y	-0.07	-0.10	0.02	-2.88	.005
Regression model with prior FAI	*b*	β	*SE*	*t*	*p*
Father routine caregiving involvement 2Y	-0.03	-0.04	0.02	-1.27	.208
Functional Abilities Index 2Y	0.55	0.47	0.06	9.76	.000

Note. Not shown are control variables. $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement/children's FAI (X).

Table 30.

Association between Father Responsive Caregiving Involvement and Children's Later Functional Abilities.

Outcome: Children's Functional Abilities at 2 Years					
Regression model without prior FAI	*b*	β	*SE*	*t*	*p*
Father responsive caregiving involvement 9M	-0.50	-0.12	0.19	-2.59	.011
Regression model with prior FAI	*b*	β	*SE*	*t*	*p*
Father responsive caregiving involvement 9M	-0.52	-0.12	0.19	-2.68	.009
Functional Abilities Index 9M	0.29	0.26	0.05	6.02	.000

Outcome: Children's Functional Abilities at 4 Years					
Regression model without prior FAI	*b*	β	*SE*	*t*	*p*
Father responsive caregiving involvement 9M	-0.51	-0.10	0.19	-2.61	.011
Regression model with prior FAI	*b*	β	*SE*	*t*	*p*
Father responsive caregiving involvement 9M	-0.52	-0.09	0.19	-2.66	.009
Functional Abilities Index 9M	0.21	0.23	0.07	3.18	.002

Outcome: Children's Functional Abilities at 4 Years					
Regression model without prior FAI	*b*	β	*SE*	*t*	*p*
Father responsive caregiving involvement 2Y	-0.17	-0.04	0.17	-1.02	.311
Regression model with prior FAI	*b*	β	*SE*	*t*	*p*
Father responsive caregiving involvement 2Y	-0.11	-0.02	0.13	-0.82	.414
Functional Abilities Index 2Y	0.55	0.47	0.05	10.09	.000

Note. Not shown are control variables. $\beta = \beta_{StdYX}$ = standardized difference in outcome (Y) associated with a 1-standard-deviation difference in prior father involvement/children's FAI (X).

Chapter Four: Discussion

Study Overview

This study examined the associations between child disability/delay status, children's functional abilities, and father involvement in a nationally representative sample. Children with disabilities/delays thought to be present from birth were matched to one or more children who were typically developing and had a similar propensity score. As part of the matching process, children were assigned weights such that analyses conducted using the matched sample produced estimates of the effect of child disability/delay status on father involvement or on children's functional abilities for children with a disability/delay. In other words, analyses produced estimates of the treatment (i.e., disability/delay status) on the treated (i.e., children who have a disability/delay). Four research questions were explored in this study:

1. How does the presence of a child's disability or developmental delay influence father involvement across the first four years of life?
2. How are children's functional abilities associated with father involvement?
3. Do children's functional abilities mediate the association between disability/delay status and father involvement?
4. How is father involvement associated with children's functional abilities and does father involvement mediate the effect of disability/delay status on functional abilities?

Below, study findings for the four questions are first summarized for each type of father involvement. Second, the extent to which hypothesized associations were found in the data is reviewed. Third, study limitations and guidance on generalization of findings is given. Finally, implications of findings for interventions are discussed.

Summary of Findings by Father Involvement Type

Father literacy involvement. Propensity score analyses indicated that child disability/delay status had a direct, causal effect on father literacy involvement at 2 years, but not at 9 months or 4 years. Child disability/delay status was linked with lower levels of father literacy involvement at 2 years. Mediation analyses testing whether the effect of child disability/delay status on father literacy involvement was mediated by earlier father involvement revealed a significant indirect effect. Specifically, child disability/delay status was linked to lower levels of father literacy involvement at 2 years which in turn was linked to father literacy involvement at 4 years. Thus, child disability/delay status led to lower levels of father literacy involvement at 4 years through its effect on father literacy involvement at 2 years. Additionally, after including father literacy involvement at 2 years in the model, child disability/delay status had a significant, positive, direct effect on father literacy involvement at 4 years leading to higher levels of father literacy involvement. In summary, child disability/delay status was

linked to both higher and lower levels of father literacy involvement at 4 years through its indirect and direct effects, respectively.

Although analyses for Question 3 revealed no evidence that children's functional abilities mediated the association between child disability/delay status and father literacy involvement, analyses for Question 2 indicated that children's functional abilities at 2 years were associated with lower levels of father literacy involvement at 4 years when controlling for father literacy involvement at 2 years. However, this association was no longer significant once disability/delay status was included in the model. This change in significance indicates that this association may have been simply an artifact of the significant association between children's functional abilities and disability/delay status, which did have an effect on father literacy involvement. There was no evidence that father literacy involvement mediated the association between child disability/delay status and children's functional abilities or that father literacy involvement was associated with later functional abilities.

Father routine caregiving involvement. Propensity score analyses indicated that child disability/delay status had a direct, causal effect on father routine caregiving involvement at 9 months and 2 years, but not at 4 years. Child disability/delay status was linked with lower levels of father routine caregiving involvement at 9 months and 2 years. Mediation analyses testing whether the effect of child disability/delay status on father routine caregiving involvement was mediated by earlier father involvement revealed 3 significant indirect effects. Specifically, child disability/delay status was linked to lower levels of father routine caregiving involvement at 9 months which in turn was linked to father routine caregiving involvement at 2 years. Thus, child disability/delay status led to lower levels of father routine caregiving involvement at 2 years through its effect on father routine caregiving involvement at 9 months. Once father routine caregiving involvement at 9 months was included in the model, child disability/delay status no longer significantly predicted father routine caregiving at 2 years. Similarly, father routine caregiving at 9 months and at 2 years mediated the effect of child disability/delay status on father routine caregiving involvement at 4 years. In both cases, child disability/delay status led to lower levels of father routine caregiving involvement at 4 years through routine caregiving involvement at 9 months and at 2 years. In summary, child disability/delay status had a direct effect on father routine caregiving involvement at 9 months, but appeared to have only indirect effects on father routine caregiving involvement at 2 years and 4 years. In all cases, child disability status was linked to lower levels of father routine caregiving involvement. It should be noted that it is possible that direct effects of child disability/delay status on father routine caregiving involvement at 2 years and 4 years exist, but suppression effects and/or other mediators with opposite indirect effects are hiding the direct effect (see Rucker et al., 2011).

Analyses for Question 2 revealed that children's functional abilities at 2 years were associated with lower levels of father routine caregiving at 4 years. Furthermore, analyses for Question 3 indicated that children's functional abilities at 2 years mediated the effect of child disability/delay status on father routine caregiving involvement at 4 years. This was true whether or not analyses controlled for father routine caregiving involvement at 2 years. Analyses with and without prior father involvement differed on whether there was still a significant direct effect of child disability/delay status on father routine caregiving involvement at 4 years. If one interprets the model with prior father involvement as representing change (i.e. "rank-order instability"), there is only evidence that child disability/delay status had an effect on "change" in father routine caregiving involvement at 4 years indirectly through father routine caregiving involvement at 2 years. In contrast, the model without prior father involvement indicated that child disability/delay status had both a direct and indirect effect on father routine caregiving involvement at 4 years. Which model is correct? The inclusion of prior father involvement changes the interpretation of the coefficients, but it also helps control for some unobserved variables that may be biasing the results of the model without prior father involvement (Menard, 2002). Given that the model with prior father involvement gives a more conservative estimate, it is probably safer to conclude that child disability/delay status only has an indirect effect on father involvement at 4 years. That being said, it is possible that the direct effect might be hidden by suppression effects of unobserved variables. In general, it is better to interpret only significant effects in mediation models given that additional mediating mechanisms may be yet unaccounted for.

Although analyses for Question 4 provided no evidence that father routine caregiving involvement mediated the effect of child disability/delay status on children's functional abilities, father routine caregiving involvement at 9 months was associated with lower levels of children's functional abilities at 2 years and at 4 years. Analyses controlling for prior levels of children's functional abilities indicated that father routine caregiving involvement at 9 months was associated with a decrease in functional abilities from 9 months to 2 years and from 9 months to 4 years, meaning a decrease in rank order relative to others in the sample even if the entire sample increased in functional abilities. Technically, participants may not have changed rank order, but the stability of their rank order may have changed, meaning, for example, that the participants with the 3rd and 4th highest scores moved a little closer to one another in their FAI scores. Father routine caregiving involvement at 2 years was only associated with children's functional abilities at 4 years in the model that did not control for prior father involvement. Given that the model with prior father involvement provides a more conservative estimate, it is probably safer to conclude that father routine caregiving involvement at 2 years was not associated with children's functional abilities at 4 years.

Father play involvement. Propensity score analyses indicated that child disability/delay status had a direct, causal effect on father play involvement at 9 months and 2 years, but not at 4 years. Child disability/delay status was linked with lower levels of father play involvement at 9 months and 2 years. Mediation analyses testing whether the effect of child disability/delay status on father play involvement was mediated by earlier father involvement revealed 3 significant indirect effects. Specifically, child disability/delay status was linked to lower levels of father play involvement at 9 months which in turn was linked to father play involvement at 2 years. Thus, child disability/delay status led to lower levels of father play involvement at 2 years through its effect on father play involvement at 9 months. Once father play involvement at 9 months was included in the model, child disability/delay status no longer significantly predicted father play involvement at 2 years. Similarly, father play involvement at 9 months and at 2 years mediated the effect of child disability/delay status on father play involvement at 4 years. In both cases, child disability/delay status led to lower levels of father play involvement at 4 years through play involvement at 9 months and 2 years. In summary, child disability/delay status had a direct effect on father play involvement at 9 months, but only indirect effects on father play involvement at 2 years and 4 years were detected. In all cases, child disability status was linked to lower levels of father play involvement.

Analyses for Question 2 did not indicate that children's functional abilities at 2 years were associated with later father play involvement. Nor did the analyses for Question 3 indicate that children's functional abilities mediated the effect of child disability/delay status on later father play involvement. Analyses for question 4 found no evidence that father play involvement was associated with children's functional abilities or that father play involvement mediated the association between child disability/delay status and children's later functional abilities.

Father responsive caregiving involvement. Few significant findings for father responsive caregiving involvement were identified. There were no significant direct or indirect effects of child disability/delay status on father responsive caregiving involvement. Children's functional abilities were not associated with father responsive caregiving involvement nor did they mediate association between child disability/delay status and father involvement. Although analyses for Question 4 provided no evidence that father responsive caregiving involvement mediated the effect of child disability/delay status on children's functional abilities, father responsive caregiving involvement at 9 months was associated with lower levels of children's functional abilities at 2 years and at 4 years. Analyses controlling for prior levels of children's functional abilities indicated that father responsive caregiving involvement at 9 months was associated with a decrease in children's functional abilities from 9 months to 2 years and from 9 months to 4 years, meaning a decrease in rank order stability.

Review of Hypothesized Associations and Significant Findings

Effect of child disability/delay status on father involvement. Past research has provided weak and inconsistent evidence that fathers of children with disabilities/delays are less optimally involved than fathers of typically developing children (Dyer et al., 2009; MacDonald & Hastings, 2010; Pelham et al., 1997; Ricci & Hodapp, 2003; Roach et al., 1999; Sanders & Morgan, 1997). Past research examining father involvement with children with disabilities/delays has also been limited to observational/correlational methods of research. The present study used a quasi-experimental method known as propensity score analysis to estimate the causal effect of child disability/delay status on father involvement. Because the causal effect of disability/delay status on father involvement cannot be estimated using the observational/correlational methods used in past research, the findings of such studies may be biased. Consequently, they provide little guidance on the direction of the effect of child disability/delay status on father involvement.

Theoretical guidance on the direction of effects is not readily available either. For example, family systems theory suggests that members of a family subsystem, such as the father-child subsystem, mutually influence one another, but does not specify *how* they influence one another. Furthermore, research on families of children with disabilities/delays has historically been guided by theoretical assumptions that the presence of a disability or delay inevitably has detrimental effects on family members; however, more recently, researchers have explored the positive aspects of the presence of a disability/delay as well as family adaptation and resilience (see Risdal & Singer, 2004). This history cautions against hypothesizing negative effects of disability/delay status without strong theoretical justification. As a result, the first research question on the effect of child disability/delay status on father involvement was exploratory.

While the direction of the effect of child disability/delay status was not hypothesized, it is of interest that all significant effects (with one exception)—direct and indirect—were negative. Thus, the presence of a child with a disability/delay does appear to lead to lower levels of father involvement, though the exact mechanism by which this occurs is unknown (functional abilities may be one mechanism as discussed below). However, the size of the effect should be taken into account when considering this finding: standardized regression coefficients never exceeded |.27|, indicating a small effect. The size of the effect is interesting in light of the historical view that the presence of a child with a disability/delay was a family tragedy (Risdal & Singer, 2004). After finding a small effect size (Cohen's d = .21) in their meta-analysis of marital adjustment in families of children with disabilities, Risdal and Singer (2004) noted that the presence of a child with a disability/delay had "a much smaller effect on parents' marital relationships than would be expected under older assumptions about disability and family" (p. 95). A similar statement might be made about the findings of this study: the presence of a child with a

disability/delay had a much smaller effect on father involvement than would be expected under older assumptions about disabilities/delays and fathers. For a discussion of the older, negative views of fathers' experiences with children with disabilities/delays, see Hornby (1995).

Although the effect of disability/delay status on father involvement was small, it still needs to be interpreted. Of particular interest, the direct effects of child disability/delay status were seen (with one exception) exclusively at the earlier time points: 9 months and 2 years. There were direct effects of child disability/delay status on both father routine caregiving involvement and play involvement at both 9 months and 2 years, but not at 4 years. Mediation analyses revealed that the effects of child disability/delay status on father routine caregiving and play involvement were primarily mediated by earlier routine caregiving and play involvement. Although non-significant direct effects must be interpreted with caution (Rucker et al., 2011), there is only evidence that disability/delay status influenced routine caregiving and play involvement at 2 years through routine caregiving and play involvement at 9 months. Similarly, there is only evidence that father routine caregiving and play involvement at 4 years were influenced by child disability/delay status indirectly through prior father involvement. This arrangement of significant effects suggests that child disability/delay status had an early effect on routine caregiving and play involvement and that whatever level of involvement fathers established early on was carried forward. In fact, the level of involvement may have been established prior to 9 months. A significant direct effect of child disability status on father routine caregiving and play involvement at 9 months may have only been detected because it is the earliest time point. Mediation analyses exploring father involvement at 6 months, for example, may lead to a non-significant direct effect on father involvement at 9 months just as occurred with father involvement at 2 years. It should be emphasized again, however, that non-significant direct effects should be interpreted with caution. What can be said with certainty is that child disability/delay status influenced later father involvement through earlier father involvement, though perhaps not exclusively. Given the high stability of father involvement across time in this sample, these early influences on father routine caregiving and play involvement may have considerable impact on what fathers do for years down the road. This finding should not be interpreted to mean that fathers' absolute level of involvement remains the same over time once it is initially established. Rather, child disability/delay status appears to lead fathers of children with disabilities/delays to be less involved relative to fathers of children who are typically developing at 9 months and they continue to be relatively less involved at later time points regardless of whether father involvement for the entire sample increases, decreases, or stays the same on average. In other words, these analyses are looking at between-person differences rather than within-person differences. The term "level of involvement" is used in the between-person sense.

In contrast to father routine caregiving and play involvement, literacy involvement was not affected by child disability/delay status at 9 months, but was at 2 years. What might explain the absence of a 9-month effect? One possibility is that father literacy involvement at 9 months can vary considerably across children—even children who are typically developing—because of variation in children's interest in and readiness for literacy activities at that young age. Because of this variation, it may be more difficult to detect group differences. Another possibility is that some of the items assessing literacy involvement tap into behaviors that can be done without fully engaging the child. For example, stories can be told to a child or songs sung without the child having to respond in any significant way. If the child does not have to be as fully engaged, then the child's disability/delay may not interfere with these literacy activities. If this is true, then why was there an effect of disability/delay status on later literacy involvement? Perhaps the nature of literacy involvement changed to focus more on book reading at older ages which can be hindered by impairments associated with a disability/delay. Yet the measurement invariance analyses indicated that literacy involvement was the same construct across time. Alternatively, perhaps as children get older, impairments associated with a disability/delay do inhibit singing songs and telling stories in addition to book reading.

One interesting finding related to father literacy involvement was that child disability/delay status had both an indirect negative effect and a (suppressed) positive effect on father literacy involvement at 4 years. This pattern is different from the pattern of findings for routine caregiving and play involvement. Furthermore, father literacy involvement was the only measure of father involvement at 4 years to be influenced directly by child disability/delay status. This pattern suggests that child disability/delay status led to lower levels of father literacy involvement at 2 years which were carried forward and continued at 4 years, but that fathers of children with disabilities/delays were also influenced directly by their child's disability/delay to engage in higher levels of literacy involvement at 4 years. What might explain why fathers of children with disabilities/delays become relatively more engaged in literacy involvement at 4 years? There are several possibilities. Children may have received a diagnosis of their disability/delay leading fathers to become more involved in order to address the disability/delay. Alternatively, many children may have received an IEP and/or early intervention services that might encourage fathers to become more involved in literacy activities. A third possibility is that fathers may have been concerned about their child's readiness to enter kindergarten. McBride, Dyer, Liu, Brown, & Hong (2009) suggested that fathers may become more involved in their child's education when the child is struggling. Others have also found that parents become more involved in their children's education when children are struggling (e.g., Pomerantz, Grolnick, & Price, 2005). For preschool children who have disabilities/delays and are struggling with kindergarten readiness, perhaps father involvement in their education may take the form of greater literacy involvement. This is an important question for future research.

There were no direct or indirect effects of child disability/delay status on father responsive caregiving involvement at 9 months or 2 years. Father responsive caregiving involvement is unique relative to the other forms of involvement in that it is a relative measure of father involvement. Specifically, fathers reported how often they were the one to meet a child's need. Previous research, though dated, has indicated that mothers tend to take on more of the child care responsibilities in families of children with disabilities/delays than in families of children who are typically developing while fathers take on the role of financial provider (see MacDonald & Hastings, 2010). It would seem that a measure of relative involvement would be well-suited to detect such gender role polarization. The absence of a significant effect of child disability/delay status on responsive caregiving involvement suggests that perhaps parents are not differentiating their roles as much as in the past, at least not in terms of the items used to assess responsive caregiving involvement: soothing an upset child, staying home with an ill child, getting up with a child at night, and taking an ill child to the doctor.

An additional characteristic of responsive caregiving involvement is that the needs it meets tend to be more acute (though not necessarily more urgent) than the needs met through routine caregiving involvement. For example, calming a distressed child meets a more acute need than feeding or bathing him or her. One reason that there were no significant effects of child disability/delay status on father responsive caregiving involvement but there were significant effects on routine caregiving involvement is that the needs being met by responsive caregiving involvement are more acute. Children's demands for responsive caregiving may not be something fathers can ignore or pass on to mothers as easily as routine caregiving involvement—even if aspects of their child's disability/delay incline them to do so.

Children's functional abilities and father involvement: Mediation. Research Question 2 addressed the association between children's functional abilities and father involvement. Just as past research and theory gave little guidance on the expected direction of the effect of disability/delay status on father involvement, little guidance is available on the expected direction of the association between children's functional abilities and father involvement. Studies examining the effect of children's functional abilities on father involvement using various approaches have found no difference (Ricci & Hodapp, 2003), a positive association (Bristol, Gallagher, & Schopler, 1988; McBride et al., 2011), and a negative association (i.e., the overall health of children with an intellectual disability was inversely related to father involvement; Heller, Hsieh, & Rowitz, 1997). Thus, children may experience lower levels of father involvement if fathers withdraw in the face of the challenges associated with lower functional abilities or may experience higher levels of father involvement if demands associated with lower functional abilities do not allow fathers to withdraw. In the present study, children's functional abilities were associated only with father routine caregiving involvement at 4 years. Specifically, fathers were more involved with children who had lower functional abilities. This association may be the result of

children with lower functional abilities needing more help with caring for themselves physically. Thus, fathers may have engaged in greater routine caregiving tasks to support their children with lower functional abilities.

In addition to examining the effect of children's functional abilities on father involvement, mediation analyses tested whether children's functional abilities mediated the effect of child disability/delay status on father involvement. In this study, children's functional abilities were conceptualized as manifestations of the child's disability/delay status. The underlying health condition causing a child's disability/delay is conceptualized as being present from birth and continuing to be present. This is not true of all disabilities/delays, but the study sample was limited to disabilities/delays for which this assumption was reasonable. While a child's disability/delay status does not change over time, limitations in his or her functional abilities may change. It may also be the case that the underlying health condition is not detectable at birth because there are no apparent limitations in functional abilities and no direct test, such a genetic test, for identifying the underlying health condition. For example, no genetic test exists for identifying children with ASD at birth. Nor is there a test based on children's limitations in functional abilities at birth. In contrast, a child may be diagnosed with Down syndrome through genetic testing even if there are no clear limitations in functional abilities. In other cases, limitations in functional abilities are what lead to diagnosis. Furthermore, regardless of whether a child's underlying health condition is detectable at birth or there are any limitations in functional abilities at birth, a child's underlying health condition may become apparent over time as the child grows and fails to meet developmental milestones and exhibits limitations in functional abilities.

These examples demonstrate that children's functional abilities and disability/delay status (i.e., underlying health condition) are separate phenomena. They also demonstrate the diagnosis is a separate phenomenon. Each of disability/delay status, limitations in functional abilities, and diagnosis may have an effect on father involvement. However, the effect of disability/delay status on father involvement must be mediated by children's functional abilities or by diagnosis. To elaborate, an underlying health condition—the genetic or biological cause of the disability/delay—cannot be directly observed and therefore cannot directly influence father involvement. However, it can be indirectly observed through diagnosis and or limitations in functional abilities, which can act as mediators of the effect of child disability/delay status on father involvement. As indicated in the examples given above, either or both limitations in functional abilities and diagnosis may be the mediator for a given child. In other words, father involvement may be influenced by disability/delay status—the underlying health condition—as fathers react to children's limitations in functional abilities, the diagnosis of the disability/delay, or both.

Diagnosis was not tested as a mediator in this study because of difficulty in measuring disability/delay diagnosis. In this study, timing of diagnosis cannot be fully determined because there

were no direct questions about the timing of diagnosis. Although timing may be approximated based upon when parents first reported a diagnosis or receipt of early intervention services, there is no guarantee that a diagnosis was not obtained earlier for a disability/delay only asked in later time points. For example, mothers did not report on the presence of Type 1 Diabetes until 4 years, but the child could have been diagnosed with it at an earlier age. Furthermore, the diagnosis may not have to be of high "quality" in order to have an effect on father involvement. For example, in this study, mothers reported whether a doctor had told them that their child had a specific condition, such as blindness or a heart defect. While a diagnosis from a doctor would likely be taken seriously and therefore influence father involvement, it is also possible that an unofficial "diagnosis" by a teacher or even a family friend may be sufficient to changes fathers' behaviors. An official diagnosis from a doctor may not have as much of an impact on father involvement if the unofficial "diagnosis" has already led to changes in father involvement. For all of these reasons, the ECLS-B dataset was not well-suited for examining diagnosis as a mediator of the effect of disability/delay status on father involvement. This is an important area for further study. For future research examining the effect of receiving a diagnosis of a disability/delay (independent of the presence of a disability/delay and limitations in functional abilities) on father involvement, it is important to note that there is no theoretical or empirical guidance on the direction the effect would take. For example, as with disability/delay status and functional abilities, fathers could just as easily react to their child receiving a diagnosis of a disability/delay by becoming more involved with their child as withdrawing from their child.

Although I could not examine the effect of disability/delay diagnosis, I was able to explore the mediating role of children's functional abilities on the link between disability/delay status and father involvement. As part of the mediation model, the causal effect of disability/delay status on children's functional abilities was estimated. Child disability/delay status had a significant, negative effect on children's functional abilities at every time point. These results are consistent with the idea that children's functional abilities are a manifestation of the underlying health condition, or disability/delay status. In terms of significant mediation effects, only the effect of child disability/delay status on routine caregiving involvement at 4 years was mediated by children's functional abilities (at 2 years). Specifically, child disability/delay status was negatively associated with functional abilities at 2 years, which in turn were negatively associated with father routine caregiving involvement at 4 years. The net indirect effect is positive in that child disability/delay status leads to higher levels of routine caregiving involvement. As previously discussed, there may be a negative direct effect, but this was only present if prior father involvement was not included in the model. That the direct and indirect effects have opposite signs underscores the importance of considering both functional abilities and disability/delay status in research on children with disabilities.

Children's functional abilities did not mediate the effect of disability/delay status for any other type of father involvement or for routine caregiving involvement at 2 years. Why did children's functional abilities only mediate the effect of disability/delay status on routine caregiving involvement at 4 years? One possibility is that mediation effects were difficult to detect because of co-occurence of limitations in functional abilities at birth for some children but not for others. To elaborate, for some children the underlying health condition causing their disability/delay may have had limitations in functional abilities associated with it from birth, such as blindness. For other children, there may have been no apparent limitation in functional abilities at birth. Children's functional abilities were assessed at later time points sometime after birth. For some children, their FAI score at a later time point reflected new limitations in functional abilities that were not present at birth. For other children, their FAI score reflected the same limitations that were present at birth and still were present at the later time point. For yet others, their FAI score reflected a combination of new limitations and limitations that were present at birth. Perhaps because the FAI score did not reflect the same thing for all children, it was difficult to detect mediation effects. Furthermore, analyses of the effect of disability/delay status on father involvement indicated that disability/delay status may influence primarily early father involvement and that the levels of father involvement established early on are carried forward. Thus, mediation may not be detected because children's functional abilities may have been assessed after disability/delay status had already influenced father involvement. More precisely, children's functional abilities may have been assessed after disability/delay status had already had an effect on father involvement, *through earlier limitations in functional abilities*. These limitations may have been present at birth or appeared some time prior to the 9 month assessment. Future research may address these possibilities by measuring children's functional abilities at or closer to birth.

There is another possible reason why the mediation of the effect of child disability/delay status on father involvement by children's functional abilities was only found for routine caregiving involvement. Specifically, routine caregiving involvement may be the most easily influenced by children's functional abilities. For example, if a child is limited in her motor ability, she may need additional assistance with routine caregiving tasks. As another example, if a child is limited in his cognitive ability, he may likewise need assistance with routine caregiving tasks. In contrast, fathers may be able to easily adapt literacy involvement activities of singing or telling stories to their child to match their child's functional abilities. Likewise, many of the play involvement activities assessed in this study (peek-a-boo, shoulder rides, playing inside with games, or playing outside) can easily be adapted to the child's level of functional ability. To better understand the impact of children's functional abilities on father literacy and play involvement, future studies may want to focus on activities that are more dependent on the ability of the child, such as book reading, putting together puzzles, and building things (as opposed to just playing with

building blocks). Furthermore, researchers can examine how fathers adapt their involvement to their child's functional abilities and thus are able to still engage in similar levels of involvement as fathers of children who are typically developing.

Father involvement and children's functional abilities: Mediation. The fourth research questions essentially explores whether father involvement influences children's functional abilities directly and or as a mediator of the effect of disability/delay status. Past research provides some indication that father involvement may be beneficial to children with disabilities/delays by improving their functional abilities. For example, Bronte-Tinkew and her colleagues (Bronte-Tinkew et al., 2008) reported that father involvement was related to greater reductions in cognitive delays for infants with disabilities than for typically developing infants. Thus, father involvement was hypothesized to be positively associated with children's functional abilities. Furthermore, because both father involvement and children's functional abilities are affected by disability/delay status, it is possible that father involvement might mediate the effect of disability/delay status on children's functional abilities. No evidence of mediation was found in the present study. However, father routine and responsive caregiving involvement at 9 months were associated with lower levels of children's functional abilities at 2 and 4 years. The direction of the association between these forms of father involvement and children's functional abilities is the opposite of what was expected. One possible explanation for this finding draws from research on father involvement and coparenting. Recent research indicated that greater father involvement in routine caregiving involvement was associated with a decrease in supportive and an increase in undermining coparenting behaviors (Jia & Schoppe-Sullivan, 2011). As the authors of the study suggested, increases in some forms of father involvement may not be desired or well-received by mothers, and may ultimately lead to negative outcomes for family functioning. For example, any time two individuals engage in a task together, such as caring for a child, there is an increased risk for conflict over how to perform the task. This may be particularly true for families of children with disabilities in which mothers have historically been the primary caretakers. Furthermore, there is a strong research base linking coparenting quality to children's development (e.g., Abidin & Brunner, 1995; Feinberg, Kan, & Hetherington, 2007; Frosch, Mangelsdorf, & McHale, 2000; McHale & Rasmussen, 1998). Thus, it may be that greater father routine and responsive caregiving involvement may lead to greater coparenting conflict which in turn leads to lower levels of children's functional abilities. Given the speculative nature of this interpretation, additional research is needed to better understand the association of father routine caregiving and responsive caregiving involvement and children's functional abilities.

Study Limitations and Generalization of Findings

Study limitations. Several limitations of this study should be considered when interpreting the findings. First, a wide range of disabilities were represented in the CWD group (e.g., cerebral palsy, spina

bifida, ASD, etc.) used in the analyses. Ideally, one would estimate the effect of having a specific disability/delay on father involvement. This was not possible because (1) no specific disability/delay was experienced by a sufficient number of children to conduct an analysis with just that group and (2) many of the children had more than one disability/delay making it difficult to focus on the effect of a single disability/delay. This limitation was actually the result of a tradeoff made in using ECLS-B data. Since the ECLS-B is nationally representative data set, data were not collected based on children having a specific disability. The advantage of nationally representative data is that the sample is representative of all children born in the US in 2001 and not just children with specific disabilities or delays. Furthermore, because children and their families were not sampled based on having a specific disability, there was a large number of children who were typically developing and could be used as matches for children with disabilities/delays. Furthermore, because the children who were typically developing came from the same sample of the population as the children with disabilities/delays, it was safe to assume that there would be high quality matches. This could not be safely assumed in a study using two different samples if matches were drawn from an unrelated sample to be matched to the disability/delay sample. The disadvantage of the sampling procedure used in the ECLS-B study is that there were not enough children with a specific disability/delay to focus on the effect of a specific disability/delay. The ECLS-B was used for these analyses despite this limitation because the benefits of a nationally representative sample outweigh the disadvantages.

A second limitation was that the sample was restricted to children who lived with their biological parents from birth. One reason for this restriction is that it was necessary in order to obtain pre-birth data reported retrospectively at 9 months. Another reason for this restriction was that it was necessary to examine the causal effect of disability/delay status on father involvement. To elaborate, it could only be assumed that analyses were providing the estimated effect of disability/delay status on father involvement if the person reporting father involvement at a later time point was the same person that had experienced and been influenced by the child's disability/delay since his or her birth. A consequence of this restriction was that several important parenting contexts for children with disabilities/delays were excluded (e.g., non-residential fathers, father figures, etc.). Even though this restriction was necessary to accomplish one of the primary goals of this study (estimating the causal effect of disability/delay status), there is a great need for future research to consider father involvement with children with disabilities by non-resident fathers, father figures, and other men.

An additional restriction of the sample was that children were from singleton births. A limitation of the ECLS-B twin data is that father involvement was not reported separately for each twin. Thus, comparisons between twins if one had a disability/delay and the other was typically developing would not be possible. Furthermore, given that most of the predictors of disability/delay status would be identical for

each twin, it is very likely that the typically developing twin would be a match of the twin with a disability/delay. As a result, comparison of the group of children with disabilities/delays to the group of children who were typically developing would include comparisons of unrelated children as well as comparisons of related and, in many cases, genetically identical children. Restriction of the sample to singleton children provided a cleaner comparison.

A third limitation is that the measures of father involvement were limited to the items available in the ECLS-B. As noted previously, many of the items assessing father involvement were somewhat broad. Specifically, some of the items assessed father involvement in such a way that the father of a child who adapted his involvement to his child's disability/delay or functional abilities would have the same score as a father who did not need to adapt his involvement for his typically developing child. Although the nature of the involvement task changed, its frequency did not. Questions that focused on specific activities that are more dependent on the child's ability, such as putting together puzzles or shared book reading may reveal more differences between fathers of children who are typically developing and fathers of children with disabilities/delays. Furthermore, observational coding of father-child interactions in families of children with disabilities/delays would provide important insight into how fathers adapt their parenting behaviors to their child's disability/delay and how they interact differently with their children than fathers of typically developing children.

Finally, a potential limitation of this study is that not all predictors of child disability/delay status may have been included in estimating the propensity score. Any excluded predictors that are also related to father involvement or children's functional abilities could bias the estimated effect of disability/delay status on those variables, respectively. There is no way to know for sure that all predictors have been included, particularly given how relatively little is known about the causes of some disabilities/delays. That being said, the ECLS-B dataset was thoroughly searched for any even *weakly* relevant predictors resulting in the extensive list of predictors used in this study. It would seem reasonable to assume that at least all of the critical predictors have been included. In considering this limitation, it is important to recognize that it is not a limitation of only the propensity score analysis method—observational/correlational studies using the traditional regression/covariate adjustment approach can also be biased by omitted predictors. Although both approaches may be limited in this way, propensity score analysis still has the multiple advantages over covariate adjustment previously discussed.

Generalization of findings. In this section I offer some cautionary notes and guidelines for generalizing and interpreting the findings of this study. First, the sample in this study was restricted by design to children with a disability/delay that could be assumed to be present from birth, though the manifestation of the disability/delay in the form of limitations in functional abilities may not appear until later. This restriction was necessary in order to eliminate any environmental influences biasing the

estimated effect of child disability/delay status on father involvement. In other words, an unbiased estimate of the effect of child disability/delay status on father involvement could only be obtained for disabilities/delays that are present from birth. The implication of this restriction is that this study's findings can only be safely generalized to children with the disabilities/delays used in this study (see Table 6). Results may not be applicable to families of children with other disabilities. One specific finding (that disability/delay status influenced father involvement early on and later involvement may be primarily just a continuation of the level of involvement established early on) may not apply to families of children with disabilities/delays that are not necessarily present from birth.

On a related note, the weighting procedure used in this analysis was used to estimate the average treatment effect on the treated (ATT). In other words, this estimated causal effects and the associations reported in this study can only be appropriately applied to families of children with disabilities/delays. For example, it may not be appropriate to apply the results of the analyses examining how children's functional abilities were related to father involvement in families of children without disabilities/delays.

Another caution in interpreting and applying these findings relates to the range of the propensity score. As noted previously, few children who were typically developing had propensity scores greater than 0.75. Indeed, given the high value of these propensity scores of these children, these children may actually be children with disabilities/delays who had not yet been diagnosed. Furthermore, in the group of children with propensity scores greater than 0.75, 10 children with disabilities/delays did not have adequate matches among the typically developing group. Given these gaps and the low number of children who were typically developing in the upper quartile of the propensity score, the findings of this study may not be representative of their experiences. The findings of this study are more appropriately applied to those children with disabilities/delays and their families whose propensity scores are in the "region of common support," or who have approximately a 0-75% probably of having a disability.

A final cautionary note on applying these findings was mentioned previously: the findings of this study only apply to singleton children. In conclusion, the estimated effects and associations reported in this study can be assumed to be representative of the experiences of all singleton children born in 2001 who lived with their biological parents from birth to 4 years of age, had a disability/delay ("underlying health condition") present from birth, and had a 0 – 75% chance of having a disability/delay.

Implications of Findings for Intervention and Conclusion

There were three major implications of this study's findings for intervention. The first implication is related to the pattern that the effect of child disability/delay status on father involvement appears early on. There was only one direct effect of child disability/delay status on later father involvement once the mediating role of earlier father involvement was included in the model. As noted previously, suppression effect from unobserved variables or other mediators may mask significant direct effects, so we cannot

concluded that there are not hidden, significant direct effects of disability/delay status on later father involvement. What we do have evidence for, however, suggests that whatever pattern or level of involvement is established early on is carried forward. Furthermore, if mediation analyses had not been conducted, the indirect associations between disability/delay status and father routine caregiving, play, and literacy involvement would not have been detected. Thus, we would have erroneously concluded that disability/delay status was unrelated to father involvement at 4 years.

Given that almost all of the direct effects of child disability/delay status on father involvement were negative, it would be important for practitioners to engage fathers as early as possible to support them and encourage them to remain involved with their child. By providing support early on, practitioners can help fathers establish a higher level of initial involvement that could be carried forward and continued for years to come. Because many children with disabilities/delays will not have received a diagnosis by 9 months, it would be advisable for policy makers to support interventions not only for family of children diagnosed with disabilities/delays, but also for families of children who are *at risk* for having a disability/delay. Because at-risk children would most likely be identified through limitations in their functional abilities, support is needed to further develop and validate screening tools for limitations in functional abilities at earlier ages.

A second important implication for intervention providers is derived from the negative association between father routine and responsive caregiving involvement and children's functional abilities. If, as previously suggested, this negative association is the result of increased coparenting conflict when fathers become more involved, then early intervention providers need to act carefully when working with families of children with disabilities/delays. Family systems theory posits that what occurs in one family subsystem can influence what occurs in another family subsystem. In designing interventions to promote father involvement with children with disabilities/delays, interventionists should consider not only how to support what occurs in the father-child subsystem, but also how changes in the father-child subsystem might affect other subsystems, such as the coparenting or marital subsystems, or even the mother-child subsystem. Interventions aimed at promoting father-child interaction should also elicit support from the mother and address any barriers to successfully increasing father involvement while maintaining family harmony.

One final implication of these findings relates to the size of the effect of disability/delay status on father involvement. While fathers of children with disabilities/delays tended to be less involved than fathers of typically developing children, the size of the effect was small. Paraphrasing Risdal and Singer (2004), the presence of a child with a disability/delay had a much smaller effect on father involvement than would be expected under older assumptions about disabilities/delays and fathers. In contrast to the idea that fathers of children with disabilities/delays are not involved or interested in their child with a

disability/delay, the findings of this study suggest that these fathers are generally involved, but at a slightly lower level than fathers of typically developing children. False beliefs based on old assumptions cannot be used as an excuse for excluding fathers from participation in early intervention or other important activities with their child.

In conclusion, the findings in this study have important implications for researchers seeking to understand how father involvement is influenced by child disability/delay status and children's functional abilities. The findings also have implications for policy makers, practitioners, and other interventionists. Although the findings of this study do not fully address the complex relationship between child disability/delay status, functional abilities, and father involvement, this study does begin to provide a conceptual framework and empirical data for use by those seeking to promote father involvement with children with disabilities and support their families.

References

Abadie, A. & Imbens, G. W. (2002, October). *Simple and bias-corrected matching estimators for average treatment effects* (Technical Working Paper No. 283). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://core.kmi.open.ac.uk/download/pdf/6919604.pdf

Abidin, R.R. & Brunner, J.F. (1995). Development of a Parenting Alliance Inventory. *Journal of Clinical Child Psychology, 24*, 31-40.

Acock, A. C. (2005). Working with missing values. *Journal of Marriage and Family, 67*, 1012-1028.

Acock, A. C. (2012). What to do about missing values. In H. M. Cooper & P. M. Camic (Eds.), *APA handbook of research methods in psychology, Vol. 3,* (1st ed., pp. 27-50). Washington, DC: American Psychological Association.

Aldous, J. & Mulligan, G. M. (2002). Fathers' child care and children's behavior problems. *Journal of Family Issues, 23*, 624– 647.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. Arlington, VA: American Psychiatric Association.

Andreassen, C. & Fletcher, P. (2007). Early Childhood Longitudinal Study, Birth Cohort (ECLS–B) Psychometric Report for the 2-Year Data Collection (NCES 2007–084). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Austin, P. C., Grootendorsi, P., & Anderson, G. M. (2007). A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study. *Statistics in Medicine, 26,* 734-753. doi: 10.1002/sim.2580

Bailey, D. B., Simeonsson, R. J., Buysse, V., & Smith, T. (1993). Reliability of an index of child characteristics. *Developmental Medicine and Child Neurology, 35*, 806-815. doi: 10.1111/j.1469-8749.1993.tb11732.x

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182. doi: 10.1037/0022-3514.51.6.1173

Bayley, N. (1993). *Bayley Scales of Infant Development, Second Edition Manual*. San Antonio, TX: The Psychological Corporation.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development, 55*, 83-96. doi: 10.2307/1129836

Bolisetty, S., Bajuk, B., Me, A-L., Vincent, T., Sutton, L., & Lui, K. (2006). Preterm outcome table (POT): A simple tool to aid counselling parents of very preterm infants. *Australian and New Zealand Journal of Obstetrics and Gynaecology, 46,*189-192. doi: 10.1111/j.1479-828X.2006.00564.x

Bollen, K. A. (1998). *Structural equation models*. New York, NY: Wiley.

Bradley, R. & Corwyn, R. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53,* 371-399. doi: 10.1146/annurev.psych.53.100901.135233

Brandon, P. D. (2011). Time for work and work timing among married couples raising children with disabilities. *International Journal of Sociology, 41*, 6-26. doi 10.2753/IJS0020-7659410101.

Bristol, M. M., Gallagher, J. J., & Schopler, E. (1988). Mothers and fathers of young developmentally disabled and nondisabled boys: Adaptation and spousal support. *Developmental Psychology, 24*, 441–451. doi:10.1037/0012-1649.24.3.441

Bronte-Tinkew, J., Carrano, J., Horowitz, A., & Kinukawa, A. (2008). Involvement among resident fathers and links to cognitive outcomes. *Journal of Family Issues, 29,* 1211-1244. 10.1177/0192513X08318145

Buysse, V., Smith, T. M., Bailey, D. B., & Simeonsson, R. J. (1993). Consumer validation of an index characterizing the functional abilities of young children with disabilities. *Journal of Early Intervention, 17*, 224-238. doi: 10.1177/105381519301700302

Chambers, J. G., Perez, M., Socias, M., Shkolnik, J. Esra, P., & Brown, S. C. (2004). *Education students with disabilities: Comparing methods for explaining expenditure variation*. Palo Alto, CA: American Institutes for Research.

Cox, M. J. & Paley, B. (1997). Families as systems. *Annual Review of Psychology, 48,* 243-267. doi: 10.1146/annurev.psych.48.1.243

Croen, L. A., Grether, J. K., & Selvin, S. (2001). The epidemiology of mental retardation of unknown cause. *Pediatrics, 107*, e86. doi: 10.1542/peds.107.6.e86

Croen, L. A., Grether, J. K., & Selvin, S. (2002). Descriptive epidemiology of autism in a California population: Who is at risk? *Journal of Autism and Developmental Disorders, 32*, 217-224. 10.1023/A:1015405914950

Dalby, D. M., Hirdes, J. P., Stolee, P., Strong, J., Poss, J., Tjam, E. Y., & ... Ashworth, M. (2009). Characteristics of individuals with congenital and acquired deaf-blindness. *Journal of Visual Impairment & Blindness, 103*, 93-102.

Daley, T. C., Simeonsson, R. J., & Carlson, E. (2009). Constructing and testing a disability index in a US sample of preschoolers with disabilities. *Disability and Rehabilitation, 31*, 538-552. doi:10.1080/09638280802214352

DeGangi, G.A., Poisson, S., Sickel, R.Z., and Wiener, A.S. (1995). *Infant/Toddler Symptom Checklist*. San Antonio, TX: Therapy Skill Builders, The Psychological Corporation.

Dehejia, R. H. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84*, 151-161. doi:http://www.mitpressjournals.org/loi/rest

Duncan, S.E. & De Avila, E.A. (1998). *PreLAS 2000*. Monterey, CA: CTB/McGraw-Hill.

Durkin, M. S., Maenner, M. J., Newschaffer, C. J., Lee, L. C., Cunniff, C. M., Daniels, J. L., ... & Schieve, L. A. (2008). Advanced parental age and the risk of autism spectrum disorder. *American Journal of Epidemiology, 168*, 1268-1276. doi: 10.1093/aje/kwn250

Dyer, W. J., McBride, B. A., Santos, R. M., & Jeans, L. M. (2009). A longitudinal examination of fathers' involvement with children with developmental delays: Does timing of diagnosis matter? *Journal of Early Intervention, 31*, 265-281. doi: 10.1177/0192513X09340386

Dyson, L. L. (1997). Fathers and mothers of school-age children with developmental disabilities: Parental stress, family functioning, and social support. *American Journal on Mental Retardation, 102,* 267-279.

Ender, P. (2006). sgmediation: Program to compute Sobel–Goodman mediation tests. *Los Angeles, CA: Academic Technology Services, Statistical Consulting Group.*

Fauth, R. C., Brady-Smith, C., & Brooks-Gunn, J. (2003). *Parent-Child Interaction Rating Scales for the Play Doh® Task and Father-Child Interaction Rating Scales for the Three-Bag Task*. New York, NY: National Center for Children and Families (NCCF), Teachers College, Columbia University.

Feinberg, M. E., Kan, M. L., & Hetherington, E. M. (2007). The longitudinal influence of coparenting conflict on parental negativity and adolescent maladjustment. *Journal of Marriage and Family, 69*, 687-702. doi: 10.1111/j.1741-3737.2007.00400.x

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242), 1-185. doi: 10.2307/1166093

Finkel, S. E. (1995). *Causal analysis with panel data.* Thousand Oaks, CA: Sage.

Flores, G. & The Committe on Pediatric Research. (2010). Technical report—racial and ethnic disparities in the health and health care of children. *Pediatrics, 125*, e979-e1020. doi:10.1542/peds.2010-0188

Flouri, E. & Buchanan, A. (2004). Early father's and mother's involvement and child's later educational outcomes. *British Journal of Educational Psychology, 74*, 141-153. doi: 10.1348/000709904773839806

Forna, F., Jamieson, D. J., Sanders, D., & Lindsay, M. K. (2003). Pregnancy outcomes in foreign-born and US-born women. *International Journal of Gynecology and Obstetrics, 83*, 257–265. doi: 10.1016/S0020-7292(03)00307-2

Frosch, C. A., Cox, M. J., & Goldman, B. D. (2001). Infant-parent attachment and parental and child behavior during parent-toddler storybook interaction. *Merrill-Palmer Quarterly, 47*, 445-474. doi: 10.1353/mpq.2001.0022

Frosch, C.A., Mangelsdorf, S.C., & McHale, J.L. (2000). Marital behavior and the security of preschooler-parent attachment relationships. *Journal of Family Psychology, 14*, 144-161. doi: 10.1037/0893-3200.14.1.144

Glasson, E. J., Bower, C., Petterson, B., de Klerk, N., Chaney, G., & Hallmayer, J. F. (2004). Perinatal factors and the development of autism: A population study. *Archive of General Psychiatry, 61,* 618-627. doi:10.1001/archpsyc.61.6.618

Gresham, F. M. & Elliott, S. N. (1990*). Social skills rating system manual*. Circle Pines, MN: American Guidance Service.

Gueorguieva, R. V., Carter, R. L., Ariet, M., Roth, J., Mahan, C. S., & Resnick, M. B. (2001). Effect of teenage pregnancy on educational disabilities in kindergarten. *American Journal of Epidemiology, 154,* 212-220. doi: 10.1093/aje/154.3.212

Guo, S. & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications.* Washington, DC: Sage.

Hastings, R. P., Kovshoff, H., Ward, N. J., Espinosa, F. D., Brown, T., & Remington, B. (2005). Systems analysis of stress and positive perceptions in mothers and fathers of pre-school children with autism. *Journal of Autism and Developmental Disorders, 35,* 635-644. doi: 10.1007/s10803-005-0007-8

Heamann, M. I., Newburn-Cook, C. V., Green, C. G., Elliott, L. J., & Helewa, M. E. (2008). Inadequate prenatal care and its association with adverse pregnancy outcomes: A comparison of indices. *BMC Pregnancy and Childbirth, 8,* 15. doi:10.1186/1471-2393-8-15

Hediger, M. L., Overpeck, M. D., Ruan, W. J., & Troendle, J. F. (2002). Birthweight and gestational age effects on motor and social development. *Paediatric and Perinatal Epidemiology, 16,* 33-46. doi: 10.1046/j.1365-3016.2002.00393.x

Heiskanen, N., Raatikainen, K., & Heinonen, S. (2006). Fetal macrosomia – a continuing obstetric challenge. *Biology of the Neonate, 90,* 98-103. doi: 10.1159/000092042

Heller, T., Hsieh, K., & Rowitz, L. (1997). Maternal and paternal caregiving of persons with mental retardation across the lifespan. *Family Relations, 46,* 407-415. doi: 10.2307/585100

Herring, S., Gray, K., Taffe, J., Tonge, B., Sweeney, D., & Einfeld, S. (2006). Behaviour and emotional problems in toddlers with pervasive developmental disorders and developmental delay: Associations with parental mental health and family functioning. *Journal of Intellectual Disability Research, 50,* 874-882. doi: 10.1111/j.1365-2788.2006.00904.x

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81,* 945-970. doi: 1080/01621459.1986.10478354

Hornby, G. (1995). Fathers' views of the effects on their families of children with Down syndrome. *Journal of Child and Family Studies, 4,* 103-117. doi: 10.1007/BF02233957

Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression (3rd ed).* Hoboken, New Jersey: Wiley.

Hueston, W. J., Geesey, M. E., & Diaz, V. (2008). Prenatal care initiation among pregnant teens in the United States: An analysis over 25 years. *Journal of Adolescent Health, 42,* 243-248. doi: 10.1016/j.jadohealth.2007.08.027

James, L. R. & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology, 69,* 307-321. doi: 10.1037/0021-9010.69.2.307

Jeans, L. M., Santos, R. M., Laxman, D. J., McBride, B. A., & Dyer, W. J. (2014). Early Predictors of ASD in Young Children Using a Nationally Representative Data Set. *Journal of Early Intervention*. Advance online publication. doi: 10.1177/1053815114523319

Jia, R. & Schoppe-Sullivan, S. J. (2011). Relations between coparenting and father involvement in families with preschool-age children. *Developmental Psychology, 47,* 106-18. doi: 10.1037/a0020802

Jo, B., Stuart, E. A., MacKinnon, D. P., & Vinokur, A. D. (2011). The use of propensity scores in mediation analysis, *Multivariate Behavioral Research, 46,* 425-452. doi: 10.1080/00273171.2011.576624

Judd, C. M. & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5,* 602-619. doi: 10.1177/0193841X8100500502

Kelaher, M. & Jessop, D. J. (2002). Differences in low-birthweight among documented and undocumented foreign-born and US born Latinas. *Social Science and Medicine, 55,* 2171-2175. doi: 10.1016/S0277-9536(01)00360-4

Kotelchuck, M. (1994a). An evaluation of the Kessner Adequacy of Prenatal Care Index and a proposed Adequacy of Prenatal Care Utilization Index. *American Journal of Public Health, 84,* 1414-1420. doi: 10.2105/AJPH.84.9.1414

Kotelchuck, M. (1994b). The Adequacy of Prenatal Care Utilization Index: Its US distribution and association with low birthweight. *American Journal of Public Health, 84,* 1486-1489. doi: 10.2105/AJPH.84.9.1486

Lamb, M. E. (2010). How do fathers influence children's development? Let me count the ways. In M.E. Lamb (Ed.), *The role of the father in child development* (5th ed., pp. 1-26). New York, NY: Wiley.

Lamb, M. E., Pleck, J. H., Charnov, E. L., & Levine, J. A. (1985). Paternal behavior in humans. *American Zoologist, 25,* 883-894.

Leventhal, F. (1998). *The need for early intervention: Development of a language skills behavioral checklist for preschool children* (Unpublished doctoral thesis), Columbia University, New York.

Levy, A., Fraser, D., Katz, M., Mazor, M., & Sheiner, E. (2005). Maternal anemia during pregnancy is an independent risk factor for low birthweight and preterm delivery. *European Journal of Obstetrics & Gynecology and Reproductive Biology, 122,* 182-186. doi: 10.1016/j.ejogrb.2005.02.015

Lillie, T. (1993). A harder thing than triumph: Roles of fathers of children with disabilities. *Mental Retardation, 31,* 438-443.

MacDonald, E. E. & Hastings, R. P. (2010). Fathers of children with developmental disabilities. In M.E. Lamb (Ed.), *The role of the father in child development* (5th ed., pp. 486-516). New York, NY: Wiley.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis.* New York, NY: Erlbaum.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science, 1*, 173-181. doi: 10.1023/A:1026595011371

Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuiseppi, C., Durkin, M. S., … Kirby, R. S. (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health*, *99*, 493–498. doi:10.2105/AJPH.2007.131243

Maupin, R., Lyman, R., Fatsis, J., Prystowiski, E., Nguyen, A., Wright, C., … Miller, J. (2004). Characteristics of women who deliver with no prenatal care. *The Journal of Maternal-Fetal & Neonatal Medicine, 16*, 45-50. doi:10.1080/14767050412331283913

McBride, B. A., Dyer, W. J., Liu, Y., Brown, G. L., & Hong, S. (2009). The differential impact of early father and mother involvement on later student achievement. *Journal of Educational Psychology, 101*, 498-508. doi: 10.1037/a0014238

McBride, B. A., Dyer, W. J., Santos, R. M., Laxman, D. J., & Jeans, L. M. (2011, April). *The impact of child disability status on father involvement in childrearing*. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Quebec.

McBride, B., Laxman, D. J., Dyer, W. J., Kern, J. L., Santos, R. M., & Jeans, L. M. (2013, April). *Disability status versus functional abilities in predicting child outcomes.* Poster presentation at the annual meeting of the American Education Research Association (AERA), San Francisco, CA.

McHale, J. P. & Rasmussen, J. L. (1998). Coparental and family-level dynamics during infancy: Early family precursors of child and family functioning during preschool. *Development and Psychopathology, 10*, 39-59.

Menard, S. (2002). *Longitudinal research (2nd ed.).* Thousand Oaks, CA: Sage.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543. doi:10.1007/BF02294825

Merrell, K. M. (2003). *Preschool and Kindergarten Behavior Scales (PKBS-2).* Austin, TX: Pro-Ed, Inc.

Mezulis, A. H., Hyde, J. S., & Clark, R. (2004). Father involvement moderates the effect of maternal depression during a child's infancy on child behavior problems in kindergarten. *Journal of Family Psychology, 18*, 575-588. doi: 10.1037/0893-3200.18.4.575

Milne, D. & Glasier, A. (2008). Preventing repeat pregnancy in adolescents. *Current Opinion in Obstetrics and Gynecology, 20,* 442-446. doi:10.1097/GCO.0b013e3283086708

Moos, M. K., Dunlop, A. L., Jack, B. W., Nelson, L., Coonrod, D. V., Long, R., ... & Gardiner, P. M. (2008). Healthier women, healthier reproductive outcomes: Recommendations for the routine care of all women of reproductive age. *American Journal of Obstetrics and Gynecology, 199*, S280-S289. doi: 10.1016/j.ajog.2008.08.060

Morris, J. K., Mutton, D. E., & Alberman, E. (2002). Revised estimates of the maternal age specific live birth prevalence of Down's syndrome. *Journal of Medical Screening, 9*, 2-6. doi: 10.1136/jms.9.1.2

Morse, S. B., Zheng, H., Tang, Y., & Roth, J. (2009). Early school-age outcomes of late preterm infants. *Pediatrics, 123*, e622-e629. doi: 10.1542/peds.2008-1405

Murnane, R. J. & Willett, J. B. (2010). *Methods matter: Improving causal inference in educational and social science research.* New York, NY: Oxford University Press.

Najarian, M., Snow, K., Lennon, J., & Kinsey, S. (2010). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Preschool–Kindergarten 2007 Psychometric Report* (NCES 2010-009). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Newsom, J. T. (2012). Basic longitudinal analysis approaches for continuous and categorical variables. In J. T. Newsom, R. N. Jones, & S. M. Hofer (Eds), *Longitudinal data analysis: A practical guide for researchers in aging, health, and social sciences* (pp. 143-180). New York, NY: Routledge.

Nord, C., Edwards, B., Andreassen, C., Green, J. L., & Wallner-Allen, K. (2006). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), User's Manual for the ECLS-B Longitudinal 9-Month–2-Year Data File and Electronic Codebook* (NCES 2006–046). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Olsson, M. B. & Hwang, C. P. (2006). Well-being, involvement in paid work and division of child-care in parents of children with intellectual disabilities in Sweden. *Journal of Intellectual Disability Research, 50*, 963-969. doi: 10.1111/j.1365-2788.2006.00930.x

Pelham, W. E., Lang, A. R., Atkeson, B., Murphy, D. A., Gnagy, E. M., Greiner, A. R.,…Greenslade, K. E. (1997). Effects of deviant child behavior on parental distress and alcohol consumption in laboratory interactions. *Journal of Abnormal Child Psychology, 25*, 413–424. doi: 10.1023/A:1025789108958

Pettit, G. S., Brown, E. G., Mize, J., & Lindsey, E. (1998). Mothers' and fathers' socializing behaviors in three contexts: Links with children's peer competence. *Merrill-Palmer Quarterly, 44*, 173-193.

Pleck, J. H. (2007). Why could father involvement benefit children? Theoretical perspectives. *Applied Developmental Science, 11*, 196-202. doi: 10.1080/10888690701762068

Pleck, J. H. (2010). Paternal involvement: Revised conceptualization and theoretical linkages with child outcomes. In M. E. Lamb (Ed.), *The role of the father in child development*, (5th ed., pp. 58-93). New York, NY: Wiley.

Pleck, J. H. & Hofferth, S. L. (2008). Mother involvement as an influence on father involvement with early adolescents. *Fathering, 6*, 267-286. doi:10.3149/fth.0603.267

Pleck, J. H. & Masciadrelli, B. P. (2004). Paternal involvement by U.S. resident fathers: Levels, sources, and consequences. In M. E. Lamb (Ed.), *The role of the father in child development* (4th ed., pp. 222-271). New York, NY: Wiley.

Pomerantz, E. M., Grolnick, W. S., & Price, C. E. (2005). The role of parents in how children approach school: A dynamic perspective. In A. J. Elliot & C. S. Dweck (Eds.), *The handbook of competence and motivation* (pp. 259-278). New York, NY: Guilford Press.

Preacher, K. J. & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717-731. doi: 10.3758/BF03206553

Rasmussen, K. M., Catalano, P. M., & Yaktine, A. L. (2009). New guidelines for weight gain during pregnancy: What obstetrician/gynecologists should know. *Current Opinion in Obstetrics & Gynecology, 21,* 521-526. doi: 10.1097%2FGCO.0b013e328332d24e

Rasmussen, K. M. & Yaktine, A. L. (Eds.). (2009). *Weight gain during pregnancy: Reexamining the guidelines*. Washington, DC: National Academies Press.

Reardon, S. F., Cheadle, J. E., & Robinson, J. P. (2009). The effect of Catholic schooling on math and reading development in kindergarten through fifth grade. *Journal of Research on Educational Effectiveness, 2*, 45-87. doi: 10.1080/19345740802539267

Reichenberg, A., Gross, R., Weiser, M., Bresnahan, M., Silverman, J., Harlap, S., …Susser, E. (2006). Advancing Paternal Age and Autism. *Archive of General Psychiatry, 63,* 1026-1032. doi:10.1001/archpsyc.63.9.1026

Ricci, L. A. & Hodapp, R. M. (2003). Fathers of children with Down syndrome versus other types of intellectual disability: Perceptions, stress and involvement. *Journal of Intellectual Disability Research, 47,* 273-284. 10.1046/j.1365-2788.2003.00489.x

Risdal, D. & Singer, G. H. (2004). Marital adjustment in parents of children with disabilities: A historical review and meta-analysis. *Research & Practice for Persons with Severe Disabilities, 29,* 95-103. doi: 10.2511/rpsd.29.2.95

Roach, M. A., Orsmond, G. I., & Barratt, M. S. (1999). Mothers and fathers of children with Down syndrome: Parental stress and involvement in childcare. *American Journal on Mental Retardation, 104,* 422-436.

Roggman, L. A., Boyce, L. K., Cook, G. A., Christiansen, K., & Jones, D. (2004). Playing with daddy: Social toy play, early head start, and developmental outcomes. *Fathering, 2*, 83-108. doi: 10.3149/fth.0201.83

Rosenbaum, P. & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70,* 41-50. doi: 10.1093/biomet/70.1.41

Rosenbaum, P. & Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*, 33-38. doi: 10.1080/00031305.1985.10479383

Rosenberg, S. A., Zhang, D., & Robinson, C. C. (2010). Prevalence of developmental delays and participation in early intervention services for young children. *Pediatrics, 121,* 1503-1509. doi: 10.1542/peds.2007-1680

Royston, P. (2005). Multiple imputation of missing values: Update. *Stata Journal, 5,* 188-201.

Royston, P. (2007). Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. *Stata Journal, 7,* 445-464.

Royston, P. (2009). Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. *Stata Journal, 9,* 466-477.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology, 66*, 688-701. doi: 10.1037/h0037350

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass, 5*, 359-371. doi: 10.1111/j.1751-9004.2011.00355.x

Saloviita, T., Itälinna, M. & Leinonen, E. (2003) Explaining the parental stress of fathers and mothers caring for a child with intellectual disability: A Double ABCX Model. *Journal of Intellectual Disability Research, 47*, 300–312. doi: 10.1046/j.1365-2788.2003.00492.x

Sanders, J. L. & Morgan, S. B. (1997). Family stress and adjustment as perceived by parents of children with autism or Down syndrome: Implications for intervention. *Child and Family Behavior Therapy, 19*, 15-32.

Savitz, D. A., Schwingl, P. J., & Keels, M. A. (1991). Influence of paternal age, smoking, and alcohol consumption on congenital anomalies. *Teratology, 44*, 429-440. 10.1002/tera.1420440409

Selig, J. P. & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. Card (Eds.), *Handbook of developmental research methods*, (pp. 265-278). New York, NY: The Guilford Press.

Seligman, M. & Darling, R.B. (2007) *Ordinary families special children: A systems approach to childhood disability.* New York, NY: The Guilford Press.

Simeonsson, R. J., Leonardi, M., Lollar, D., Bjorck-Akesson, E., Hollenweger, J., & Martinuzzi, A. (2003). Applying the International Classification of Functioning, Disability and Health (ICF) to measure childhood disability. *Disability and Rehabilitation, 25*, 602-610. doi:10.1080/0963828031000137117

Simmerman, S., Blacher, J., & Baker, B. L. (2001). Fathers' and mothers' perceptions of father involvement in families with young children with a disability. *Journal of Intellectual & Developmental Disability, 26*, 325-338. doi:10.1080/13668250120087335

Snow, K., Derecho, A., Wheeless, S., Lennon, J., Rosen, J., Rogers, J., … Einaudi, P. (2009). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Kindergarten 2006 and 2007 Data File User's Manual (2010-010).* Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

StataCorp. (2011). Stata statistical software: Release 12.1. College Station, TX: StataCorp LP.

Stoneman, Z. (2007). Examining the Down syndrome advantage: Mothers and fathers of young children with disabilities. *Journal of Intellectual Disability Research, 51*, 1006-1017. doi: 10.1111/j.1365-2788.2007.01012.x

Stoneman, Z., Brody, G. H., & Abbott, D. (1983). In-home observations of young Down syndrome children with their mothers and fathers. *American Journal of Mental Deficiency, 87,* 591-600.

Sullivan, P. M., & Knutson, J. F. (1998). The association between child maltreatment and disabilities in a hospital-based epidemiological study. *Child Abuse & Neglect, 22,* 271–288. doi: 10.1016/S0145-2134(97)00175-0

Sullivan, P. M. & Knutson, J. F. (2000). Maltreatment and disabilities: A population-based epidemiological study. *Child Abuse & Neglect, 24,* 1257-1273. doi: 10.1016/S0145-2134(00)00190-3

Sumner, G. & Spietz, A. (1994). *NCATS caregiver/parent-child interaction teaching manual*. Seattle, WA: NCAST Publications, University of Washington, School of Nursing.

Tamis-Lemonda, C. S., Shannon, J. D., Cabrera, N., & Lamb, M. E. (2004). Fathers and mothers at play with their 2- and 3-year-olds: Contributions to language and cognitive development. *Child Development, 75,* 1806 - 1820. doi: 10.1111/j.1467-8624.2004.00818.x

The Psychological Corporation (2001). *Bayley Short Form–Research Edition*. San Antonio, TX: Author.

van Buuren S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine, 18,* 681-694. doi: 10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R

Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4-70. doi:10.1177/109442810031002

Walsh, C. A., Mahony, R. T., Foley, M. E., Daly, L., & O'Herlihy, C. (2007). Recurrence of fetal macrosomia in non-diabetic pregnancies. *Journal of Obstetrics & Gynaecology, 27,* 374-378. doi:10.1080/01443610701327545

Whitchurch, G. & Constantine, L. (1993). Systems theories. In P. Boss, W. Doherty, R. LaRossa, W., Schumm, & S. Steinmetz (Eds.), *Sourcebook of family theories and methods: A contextual approach* (pp. 325-352). New York, NY: Plenum Press.

Widaman, K. F. & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance abuse domain. In K. J. Bryant (Ed.), *Alcohol and substance use research* (pp. 281-324). Washington, DC: American Psychological Association.

Wilkinson, L. & the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594-604. doi: 10.1037/0003-066X.54.8.594

World Health Organization (2001). *International classification of functioning, disability and health*. Geneva, Switzerland: World Health Organization.

World Health Organization (2002). *Towards a common language for functioning, disability and health*. Geneva, Switzerland: World Health Organization.

Yeung, W. J., Sandberg, J. F., Davis-Kean, P. E., & Hofferth, S. L. (2001). Children's time with fathers in intact families. *Journal of Marriage and Family, 63*, 136-154. doi:10.1111/j.1741-3737.2001.00136.x