Research Report

Department of Statistics

No. 2007:3

Sibship Size and Cognitive Ability: Are CognitiveAbilities in Children Affected by the Birth of a Sibling?

Linda Wänström

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

Sibship Size and Cognitive Ability: Are Cognitive Abilities in Children Affected by the Birth of a Sibling?

Linda Wänström
Stockholm University

A consistent negative correlation between sibship size and cognitive ability has been observed in past empirical studies (e.g. Anastasi, 1956; Higgins, Reed, & Reed, 1962; Belmont & Marolla, 1973; Nisbet & Entwistle, 1967; Page & Grandon, 1979; Velandia, Grandon, & Page, 1978; Zajonc & Markus, 1975; Zajonc, Markus, Berbaum, Bargh & Moreland, 1991). If cognitive ability in children is negatively affected by sibship size, there should be a change following the birth of a sibling. Extending previous longitudinal work (e.g. Guo & VanWey, 1999; McCall, 1984), this paper uses longitudinal multilevel modeling to test for effects of sibship size and birth of a sibling in a large American probability sample: the children of the NLSY79. Consistent effects of sibship size on scores from three PIAT subtests measuring verbal and mathematical abilities in children are found. The effects are larger for closely spaced siblings. Results from longitudinal models however suggest that sibling birth has no causal effect on cognitive ability in children over age five.

Key words: sibship size; birth of a sibling; intelligence; cognitive ability; birth order; multilevel models.

Acknowledgement: I would like to express my appreciation to Joseph Lee Rodgers who introduced me to this field of research, and who has provided many helpful comments along the way.

1. Introduction

Are children's cognitive/intellectual abilities affected negatively by having many siblings? Does their intelligence decline after the birth of a sibling? A consistent negative correlation between sibship size and cognitive ability has been observed in past empirical studies. Children with more siblings have on average lower measured intelligence than children with fewer siblings, as measured by IQ tests and other mental aptitude tests (see e.g. Anastasi, 1956; Higgins, Reed, & Reed, 1962; Belmont & Marolla, 1973; Nisbet & Entwistle, 1967; Page & Grandon, 1979; Velandia, Grandon, & Page, 1978; Zajonc & Markus, 1975; Zajonc, Markus, Berbaum, Bargh & Moreland, 1991).

Several theories have been developed to explain this negative correlation. Some of them assume causality, or partial causality, whereas others attribute the association to extraneous factors. Studies based on cross-sectional designs have consistently found a negative association between sibship size and cognitive ability, although it gets weaker after controlling for confounding variables. Guo and VanWey (1999) examined the causality of this association on a large national longitudinal sample: the children of The National Longitudinal Survey of Youth (NLSY-children). They regressed changes in cognitive ability scores onto changes in sibship size and concluded that the cross-sectional effect was spurious. They were criticized for not allowing enough time for large additions to the family, as well as for only looking at widely spaced siblings (Downey, Powell, Steelman & Pribesh, 1999).

The purpose of this paper is to expand both the database and the analytic approach by testing the effect of the birth of a sibling on a larger part of the NLSY-children data, thus allowing for more children to be examined and for larger additions to the family, in the context of longitudinal multilevel models that have not been applied to this problem in the past. We will model the cognitive growth of children from ages five through 14, a period in which up to four siblings have been born into the families in which these children grow up.

2. Theoretical framework

Social scientists have long been interested in the relationship between family structure and cognitive development (see, among many others, Anastasi, 1956; Belmont & Marolla, 1973; Downey, 2001; Guo & Van Wey, 1999; Rodgers, Cleveland, van den Oord & Rowe, 2000; Steelman & Mercy, 1980). Higgins et al. (1962) reported that correlations between sibship size and intelligence typically lie between -.20 and -.30. Scholars have carefully examined this relationship, within various datasets and using different statistical techniques, in order to determine whether or not it is causal, partly causal, or spurious. Three major theories have emerged from these attempts: the confluence model, the resource dilution theory, and the admixture hypothesis. We note that these theories, and most empirical studies, combine family (or sibship) size with birth order in building models of intellectual development. This is, in a sense, an unfortunate pairing. The two measures are necessarily correlated, but derive from different theoretical sources. Birth order is a measure that accounts for within-family variance; that is, it measures and relates to differences between siblings in the same family. On the other hand, sibship size measures differences between families, and is shared by siblings in the family once the family is complete (barring divorce or other disruption). We review each of these theories as they account for these processes in explaining cognitive/intellectual development within the context of the family environment.

2.1 The confluence model

The confluence model (Zajonc & Markus, 1975; Zajonc, 2001) is a mathematical model that predicts absolute intellectual level from sibship size, birth order, and sibling tutoring. The latter was added to account for a lastborn deficit in empirical aggregate data. Children from smaller families and children born earlier, i.e. children with lower birth orders, are predicted to have higher IQs than children in larger families and/or with higher birth orders.

Confluence theorists propose that the intellectual environment is important for every family member's absolute intellectual level, and that each child and parent contributes to this environment. For example, each parent may be given the arbitrary absolute intellectual level of 30, while the child is given an intellectual level measure of approximately zero at his/her birth (see e.g. Zajonc, 2001). The intellectual environment for the first child, as well as for the parents, is then considered to be (30+30+0)/3 = 20 at the time of the birth. The child's intellectual level increases as he/she grows older, and the confluence algebra implies that the intellectual environment necessarily gets weaker at the birth of the next child. Although the second child is born into a weaker intellectual environment, he or she will pass the first using confluence arithmetic[1]. The first-born benefits from tutoring the second-born, however, and by 11 $\pm$ 2 years the benefit of the tutoring function will outweigh the lower intellectual environment for the first-born (Zajonc and Mullally, 1997). Spacing is also important because smaller birth gaps between children lead to a lower intellectual environment than larger gaps at the time of each birth. The confluence model also implies that children in single-parent households have a disadvantage, as can be seen from the formulation above, whereas children who live in households with many adults are at an advantage.

2.2 The resource dilution theory

The resource dilution theory (Blake, 1981; Downey, 1995; 2001; Armor, 2001) focuses on family structure and how it may benefit or disadvantage children. The theory explains the negative relationship between sibship size and cognitive ability by positing that parental resources are fixed and divided among children. Parents have many different resources that they provide, and Blake (1981) has described three of them: environments or settings, e.g. shelter, food, and cultural objects; opportunities, e.g. travel; and treatments,

---

[1] As an example, say that the second child is born when the first is four. The intellectual environment is then (30+30+4+0)/4=16. The intellectual environment *at birth* is thus lower for the second child. When the first child is eight and the second is four, the intellectual environment is (30+30+8+4)/4=18. Thus, the second child has a higher intellectual environment at the age of four (18) compared to the first child's intellectual environment at his/her age of four (16), in our example.

e.g. attention, teaching, and intervention. The more children there are in a family, the less of these resources will be allocated to each individual child. There is a “dilution” of resources in that children in large families get less attention, instruction, affection etc. from parents than children in small families, and the birth of a sibling will influence the cognitive ability of the children in a family negatively.

Some resources will be diluted more than others however, depending on how well they can be shared. Books and computers may be shared broadly, for example, whereas parental attention is shared only partially, and money may be very easily diluted. Parents may also differ in how they allocate non-shared resources, such as money. Downey (2001) has pointed out that one family may decrease their college savings per child after the birth of an additional child, whereas another family may choose less expensive vacations instead. He also distinguished between base resources and surplus resources and has suggested that an additional child will only affect the amount saved for college (surplus) if all base needs (food, shelter etc.) are met.

Although resource dilution theorists assume that parental resources affect the quality of the environment for each child, they do not assume any feedback from the children to the parents, i.e. family variables such as SES are not changed by the number of children that a couple have (Blake, 1981). However, Downey (2001) suggested that the addition of the first child would dilute resources more than the second. He also suggested that siblings might serve as resources in some ways, and not just dilute resources. In this way, having more siblings may reduce the effects of having “bad” parents.

Just as the confluence model, the resource dilution theory may predict some birth order effects depending on the types of resources, and close spacing has a more negative effect than wide spacing (see e.g. Powell & Steelman, 1990). This theory also extends to the period after the children leave their parental homes, when they may have to compete for gifts, loans etc. (Downey, 2001).

2.3 The admixture hypothesis

The admixture hypothesis (Page & Grandon, 1979; Velandia, Grandon, & Page, 1978; Rodgers et al., 2000; Rodgers, 2001) is a hypothesis about population admixtures that suggests that the relationship between sibship size and cognitive ability is due to certain admixtures. The relationship between family structure and intellectual development is believed to be spurious and caused by the difference between parents of different population strata in how many children they have. Page and Grandon (1979) were among the first to suggest that population admixtures might be the cause of the relationship between sibship size and cognitive ability. They tested the confluence model using the National Longitudinal Study of Educational Effects (NLS) data as well as a Colombian dataset and found a strong negative relationship between sibship size and cognitive ability when they analyzed the data through multiple regressions on cell means. However, when they analyzed the data individually, the relationship got weaker, and when they further included race and SES into the analyses, they concluded that sibship size had a small influence on ability compared to race and SES. When they further looked at the relationship between sibship size and cognitive ability within different races and SES groups they found that it was different for different groups. They further noted that there was a strong association between the SES admixtures and the abilities of different size families. They thus concluded that different fertility patterns in different SES groups could explain a large amount of the relationship between sibship size and cognitive ability.

Rodgers et al. (2000) suggested that the negative family structure-IQ correlations were caused by parents with lower IQs having more children than parents with higher IQs, so that parental IQ is correlated both with sibship size and child IQ, causing a spurious relationship between these variables. They presented tables, which are reproduced in Table 1, with mean Peabody Individual Achievement Test scores for children from intact families[2], showing that children of different birth orders have statistically the same mean scores, whereas means of children from different family sizes differed. Higgins et al. (1962) found earlier empirical support for this hypothesis when they studied a cross-section of 1,026 families and found a correlation of -.30 between sibship size and children's IQ, their mean IQs being similar for family sizes up to five, followed by a drop. They also found correlations between -.08 and -.11 between parental IQ and sibship size, with parental mean IQs following the same pattern as child IQs, indicating parental IQ as a possible confounding variable in studies about sibship size and cognitive ability. Thus, the admixture hypothesis, in contrast to the confluence model and the resource dilution theory, does not assume a relationship between sibship size and cognitive ability, but instead attributes the empirical correlation to confounding variables. Page and Grandon (1979) have focused on SES and race as possible confounds, whereas Higgins et al. (1962) and Rodgers et al. (2000) suggested parental IQ.

[2] i.e. only families in which all children had valid scores were included.

Table 1a. PIAT-Composite Scores by Birth Order and Family Size for the 1990/1992 NLSY-Children Sample

Family size and type of score	First sibling	Second sibling	Third sibling	Fourth sibling	Fifth sibling
One child					
Mean	102.3				
SD	11.7				
N	443				
Two children					
Mean	101.2	101.2			
SD	11.4	10.8			
N	565	565			
Three children					
Mean	98.6	98.0	98.4		
SD	12.1	11.9	11.3		
N	233	233	233		
Four children					
Mean	92.7	94.8	92.9	95.9	
SD	11.7	11.0	10.1	10.6	
N	56	56	56	56	
Five children					
Mean	86.4	91.1	91.4	91.2	98.0
SD	13.7	11.6	14.4	13.4	20.1
N	14	14	14	14	14

Table 1b. PIAT-Composite Scores by Birth Order and Family Size for the 1994/1996 NLSY-Children Sample

Family size and type of score	First sibling	Second sibling	Third sibling	Fourth sibling
One child				
Mean	103.5			
SD	11.4			
N	417			
Two children				
Mean	102.5	102.9		
SD	11.4	11.1		
N	597	597		
Three children				
Mean	100.5	99.5	100.8	
SD	12.1	11.4	11.0	
N	242	242	242	
Four children				
Mean	99.0	98.6	98.1	99.3
SD	12.5	11.6	12.2	11.0
N	51	51	51	51

Note. PIAT = Peabody Individual Achievement Test; NLSY = National Longitudinal Survey of Youth. Reproduced with permission from Rodgers, Cleveland, van den Oord, and Rowe (2000).

3. Previous studies

3.1 Cross-sectional studies

As noted above, the confluence model and the resource dilution theory claim a causal, or partly causal, relationship between sibship size and cognitive ability, whereas the admixture hypothesis attributes the relationship to other factors. Scholars have reached very different conclusions in this debate, sometimes by analyzing the same data but using different methods. Scholars using cross-sectional data, i.e. a cross-section of data from a population collected at approximately one point in time, have mostly used different types of regression techniques on aggregate- or individual-, mostly between-family, data. They have typically found negative correlations between sibship size and cognitive ability (e.g. Higgins et al., 1962; Nisbet and Entwistle, 1967; Belmont & Marolla, 1973; Zajonc and Markus, 1975; Velandia, Grandon, and Page, 1978; Page and Grandon, 1979; Zajonc et al., 1991) or related outcomes, such as educational attainment (e.g. Blake, 1981; Kuo and Hauser, 1997; Pong, 1997), and they have dealt with possible confounding variables by controlling for them in various ways, for example by including covariates into the models, or by holding variable levels constant. Belmont and Marolla (1973) and Zajonc and colleagues (e.g. Zajonc et al., 1991), for example, controlled for SES and birth order and found that the negative relationship between sibship size and cognitive ability persisted. Nisbet and Entwistle (1967) and Steelman and Mercy (1980) also found sibship size effects controlling for SES and other background factors, although Steelman et al. found that they were weaker for economically advantaged children. Others have found that correlations weakened or disappeared after controlling for SES (Velandia, Grandon, and Page, 1978; Page and Grandon, 1979; Mascie-Taylor, 1980) and race (Page and Grandon, 1979). Rankin, Gaite, and Heiry (1979) found that for American Samoan families, where the mean sibship size is much larger than for American families, the relationship was non-linear: children in sibship sizes close to the mean had higher cognitive abilities than both smaller and larger sizes, indicating a possible confound of parental child-planning or value systems (Steelman, 1985). Downey (2001) also suggested the possibility that there may be more adults that play an important role in the lives of children where larger families are the norm, resulting in lower correlations between sibship size and cognitive or educational outcomes.

Other techniques, such as structural equation modeling (SEM), have also been used. Mercy and Steelman (1982) used SEM to investigate possible mediating effects of the number of older and younger siblings on the relationship between socioeconomic status and cognitive ability in a sample of six to eleven year old children. They studied White children from unbroken homes and found that, controlling for SES, the number of older siblings had a negative effect on a vocabulary subtest, whereas the number of younger siblings was negatively related to both the vocabulary and a block design subtest.

Kuo and Hauser (1997) studied sibship size, several family background variables, and educational attainment in a cohort of siblings from the 1975 survey of the Wisconsin Longitudinal Study through SEM. They found that siblings from smaller families had more years of education, and that the effects of the measured family background variables did not vary for different sibship sizes, although the effects of the unmeasured between-family variance did. They also found that gender mattered within families and that smaller families were more heterogeneous in educational attainment than larger families.

Many scholars provide causal conclusions regarding the relationship between sibship size and cognitive ability. There may, however, be differences between families that, for one reason or another, have one child, two children etc. and the factor or factors causing these differences may be unknown or for other reasons unmeasured. As described above, researchers using cross-sectional data often control for various variables thought to confound the relationship. However, as Guo and VanWey (1999) have pointed out, some of these variables are very hard to measure, and it is also impossible to know all of them in order to measure them. Rodgers et al. (2000) and Higgins et al. (1962) found that parents that have many children had lower IQs compared to those that have fewer children. Due to possible confounding variables, most cross-sectional designs are not well suited for longitudinal inferences. McCall (1985) and Guo and VanWey (1999), for example, suggested examining cognitive development in children, with longitudinal data, after the birth of an additional sibling.

3.2 Longitudinal studies

Scholars who have used longitudinal designs have generally used regression techniques on some type of change scores, or multiple regressions covarying out prior ability. Some have found significant effects of the birth of a sibling on cognitive ability (Baydar, Greek & Brooks-Gunn, 1997; Baydar, Hyle, and Brooks-Gunn, 1997; McCall, 1984). McCall (1984) examined 45 children born between 1930 and 1938 from the Fels Longitudinal Study. He tested the effect of the birth of a sibling by comparing Stanford-Binet scores of children who did not have any older siblings, but had a younger sibling born before their 7th birthday, to children with older, younger or no siblings at all while covarying out their Stanford-Binet scores at the first assessment. He also controlled for gender, sibship size, and age of assessment. He found a negative effect of the birth of a sibling, however this effect had decreased and was no longer significant after 17 years.

Baydar and colleagues investigated the effects of the birth of a sibling during the first six years of life (Baydar, Greek and Brooks-Gunn, 1997) as well as during preschool and early grade school years (Baydar, Hyle and Brooks-Gunn, 1997) on several developmental outcomes, including cognitive ability, on a European-American subsample of the NLSY-children. They used regression models with residualized change scores as the outcome and found that younger children who had at least one sibling born between 1986 and 1988 had significantly lower scores on the Peabody Picture Vocabulary subtest than those who did not have an additional sibling, controlling for motor and social development scores at baseline. After four years, these children also had significantly lower scores on the PIAT reading recognition and mathematics subtests than children who had not had a sibling. They also found a negative effect of the birth of a sibling on reading achievement for early elementary school aged economically disadvantaged children (this, however, had disappeared after four years) and a positive effect for children who were not economically disadvantaged.

Guo and VanWey (1999) examined the effects of the birth of a sibling on cognitive ability in the NLSY-children as well. They used both conventional regression analyses and change score analyses on siblings and repeated measures and hypothesized that the negative relationship between sibship size and cognitive ability would get weaker, or disappear, in moving from conventional to change score analyses. This would happen because, as they pointed out, change analyses control for unknown time-invariant between-family variation. They measured cognitive ability change by the PPVT and the PIAT Math and Reading Recognition subtests in 1986 and 1992 and used sibship size change between the same years as the factor of interest, and their results agreed with their hypotheses. They concluded that the negative relationship between sibship size and cognitive ability was spurious and possibly due to one or more of the automatically controlled family variables.

Some of the differences between Baydar et al. (1997) and Guo and VanWey's (1999) study were that Baydar and colleagues looked at the first six years of life and preschool and early grade school years separately, and included only Whites in their sample, whereas Guo and VanWey looked at ages three (or five) through fifteen, and examined both majority and minority groups. Baydar and colleagues also compared different tests over time, e.g. they compared scores of the PPVT with PIAT scores at two different time points. Guo and VanWey's (1999) study suffered some limitations, as Downey et al. (1999) and Phillips (1999) pointed out, making it hard to compare their longitudinal analysis with cross-sectional analyses. For example, they did not allow for large additions of children to the family or for closely spaced siblings in the sample. As Downey et al. (1999) noted, if children are first tested at three, four, or five years of age, and the purpose is to measure the cognitive abilities before and after the birth of a sibling, this only allows for comparisons among widely spaced siblings, which may yield biased results. Spacing might be important as mentioned previously (see also Powell and Steelman, 1990, and Zajonc and Mullally, 1997).

It is difficult to draw any general conclusions about the effects of sibship size and birth of a sibling on cognitive ability from the cross-sectional as well as the longitudinal studies that have investigated these effects so far. There is a substantial advantage to using the birth of a sibling to investigate effects of sibship size. This type of design comes much closer to capturing the mechanisms underlying a potential sibship size effect than designs that only measure sibship size. Studies of sibship size, per se, leave open many possible interpretations, including many between-family interpretations. Though it has seldom been used, a design based on the birth of a sibling can be used to account for both within-family and between-family variance, and finding this type of effect would identify more focused and identifiable mechanisms for further investigation. If the consistently found negative effect of sibship size can be attributed to causes within the family and not only to between family differences (e.g. higher IQ parents have fewer children), there should be a change in the intellectual ability following the birth of a sibling.

4. The present study

The purpose of this study is to investigate the effects of sibship size on various cognitive ability tests in the NLSY-children dataset, using both the standard sibship size index, and also indicators of birth of a sibling. Our goal is to study both sibship size effects and to study different method effects. First, negative effects of sibship size will be established for cross-sections of the data. Children from intact families measured during the last survey years will then be studied, and the results will be compared to Rodgers et al.'s (2000) tables. Secondly, the sibship size effect will be investigated using the longitudinal structure of the data. The sibship size measure will be divided into two between-child measures: the number of siblings at birth of the respondent child and the number of additional siblings at age five, and a within-child measure: the number of siblings born since age five at each assessment. These analyses will account for the longitudinal structure of the data by estimating random intercepts and slopes. Third, models with intercepts and slopes that are random between families will be fitted to account for correlations between children with the same mothers, and differences between families (such as maternal IQ) will be accounted for. Finally, the number of older and younger siblings will be partitioned further to examine differential effects of closely and widely spaced older and younger siblings. The analyses in this paper will test for the birth of one (or more) sibling effect. The children in the NLSY have been measured up to five times, between the ages of five and 14. This is a duration of 10 years, and up to four[3] siblings have been born during these times.

Thus, there are several differences between this study and Guo and VanWey's (1999) study. Guo and VanWey only examined children with scores in 1986 and 1992, and this is a limited sample, both because it only spans over six years, but also because it only consists of a small proportion of all assessed children. In the present study, we will include children with at least one valid score between 1986 and 2004. Secondly, Guo and VanWey examined change scores, whereas the present study will model the growth over time by fitting random intercepts and slopes that vary between children. Change scores only capture the difference between scores at two points in time, whereas multiple scores can be used to model the growth curves. Third, several of the children are siblings, and this will be accounted for by also fitting models with random intercepts and slopes for families. Doing this, it is possible to see the proportion of variance that is attributed to variation within children (over time), to variation between children within families, and to variation between families.

5. Method

5.1 Data

The NLSY79 2004 Child and Young Adult Data (which we will call the NLSY-children) were used for all analyses in this paper. This is an ongoing longitudinal study that contains within-family information on 11,428 children and their mothers from whom many sibling relationships can be defined. It originated as a multi-stage stratified area probability sample. Sponsored by the Bureau of Labor Statistics (BLS), the National Opinion Research Center randomly selected two samples of households from a list of housing units in areas, which included almost all 50 states, and the District of Columbia, in the U.S. in 1978. One of the samples was designed to oversample minority groups and economically disadvantaged groups. Interviewers visited the selected households and collected information such as age, sex, and race on the members, resulting in information on over 155,000 individuals. Using this information, all individuals who were between 14 and 21 years old on December 31st, 1978, were assigned to each sample. They also selected a random sample of military members (from the Department of Defense records). All individuals who completed the first interview when visited a second time were then included in the original NLSY79 sample.

[3] One child actually had five siblings born during this time.

This resulted in 6111 (90%) individuals from sample one, 5295 (89%) individuals from sample two (the minority oversample), and 1280 (72%) individuals from sample three (the military sample). Table 2 shows response rates for the original NLSY respondents between 1979 and 2004. The 6403 males and 6283 females interviewed in 1979 constituted 87% of the total intended sample. A little more than 80% of the original respondents still eligible[4] for interviews were interviewed in 2004. More information about the sampling process and response rates can be found in the NLSY79 users guide (Center for Human Resource Research (CHRR), 2006).

Table 2. Response rates, number of children born to interviewed mothers, and PIAT completion rates for assessment years 1979-2004.

Year		1979	1986	1988	1990	1992	1994	1996	1998	2000	2002	2004
Response rate[a]		100	92.6	91.2	91.1	91.9	91.1	88.88	86.7	83.2	80.3	80.1
Number of children born[b]			5255	6543	6427	7255	7862	8125	8395	8323	8100	8267
							6622[c]	6010	5343	4438	3502	2755
Completion rate[d]	M								83.4		81.4	80.3
	RR								83.3		81.5	80.9
	RC								82.5		80.8	80.8

Note: PIAT=Peabody Individual Achievement Test, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension. This table is based on tables in the NLSY79- and NLSY79 child & young adult users guides (CHRR, 2006).
[a] Percentage interviewed of 1979 respondents remaining eligible and not known to be deceased.
The original NLSY79 sample consisted of 6403 males and 6283 females interviewed in 1979. They constituted 87% of the total intended sample.
[b]The number of children born to interviewed mothers.
[c]Number of children ages 0 – 14 born to interviewed mothers.
[d]Percentage of eligible children born to interviewed mothers with valid test scores for M, RR, and RC.

Starting in 1986, assessments were administered to the biological children of the 6283 females included in the survey. In Table 2 we can see the number of children born to the interviewed females each year. The number of children born to the females interviewed in 2004 was 8267, of which 2755 were between 0 and 14 years old. The NLSY-children represent a cross-section of children born to an approximately representative sample of mothers in the U.S. who were between 14 and 21 years old on December 31st, 1978, although minority groups are overrepresented. More information about the NLSY-children can be found in the NLSY79 child and young adult data users guide (CHRR, 2006).

[4] Much of the military sample as well as economically disadvantaged non-Black non-Hispanics from the minority oversample was dropped because of economic constraints.

The NLSY-children data are particularly suitable for the current analyses because the children have been cognitively assessed biannually starting in 1986 until the present time. The NLSY data also contain background information about the children and their families, such as race, maternal IQ, maternal age at the birth of the first child, and whether the father is present in the household or not. One limitation of the NLSY sample is that all mothers have not finished reproducing. However, in 2004, the mothers were between 39 and 47 years old and they had had over 90% of their children (NLSY79 child and young adult data users guide, CHRR, 2006).

5.2 Cognitive ability assessments

The *Peabody Individual Achievement Test (PIAT)* subtests were administered to children ages five to 14 biannually starting in 1986. They measure academic achievement and have high test-retest reliability and concurrent validity. They have also been found to be predicted by and to predict other assessment tests.

The *Mathematics* (M) subtest consists of 85 multiple-choice questions of increasing difficulty, assessing skills such as numerals recognition, geometry, and trigonometry. The child answers each problem by pointing to or naming one of four options. The *Reading Recognition* (RR) subtest is designed to measure word recognition and pronunciation ability and consists of 84 multiple-choice questions of increasing difficulty where children match letters, name names, and read words out loud. The *Reading Comprehension* (RC) subtest is designed to measure ability to derive meaning from silently read sentences by children who scored at least 19 (15 between 1986 and 1992) on the Reading Recognition subtest. Children who scored lower were assigned their Reading Recognition score. The subtest consists of 66 multiple-choice questions of increasing difficulty.

Table 2 shows completion rates for some years. We can see that the completion rates for M, RR, and RC are over 80% in 2004. More information on the PIAT subtests can be found in the NLSY79 child and young adult data users guide (CHRR, 2006).

5.3 Sibship size measures

We included several sibship size measures in our analyses. Sibship size at a certain assessment year (*SSS*) was measured by the total number of children born to the mother up to that particular assessment. In some analyses we divided *SSS* into the number of siblings born at birth of the respondent child (*OLD*), the number of additional siblings born at age five (*YOUNG*), and the number of additional siblings born at each assessment (*BIRTH*). A child who was first assessed at age five (60 months) will thus have a “0” as a first

BIRTH score, whereas a child first measured after 60 months can have a first score greater than 0. *BIRTH* will then increase by 1 for each additional sibling born since the previous assessment. We chose the age five for *YOUNG* because this is the age that the children could first be assessed with the PIAT subtests. Around 28% of the children with at least one valid PIAT score had at least one sibling born between age five and their last assessment. The percentage for children with five valid scores is higher. Thus, *BIRTH* is expected to capture longitudinal effects of the birth of a sibling that happens after age five, however it will also pick up some between-child effects because some children had one or more siblings born between the age of five and their first assessment.

In some analyses, we further divided *OLD* into the number of siblings at least 24 months older (*OLD2+*), and the number of siblings less than 24 months older than the child (*OLD0-2*). In these analyses, we also divided *YOUNG* into the number of siblings less than 24 months younger than the child (*YOUNG0-2*), and the number of additional siblings born at age five (*YOUNG2+*).

5.4 Family- and other background variables

Because the PIAT scores increase over time, we included *AGE* (age of the child in months - centered at 108 months) in the analyses. Further, *COHORT* (birth year - centered at 1980) was included because the scores of children in later cohorts have been found to be higher than the scores of children in earlier cohorts in the NLSY (Rodgers & Wänström, 2007) referred to as the Flynn effect (e.g. Flynn, 1984; 1987). It is known that the data contain more young mothers than the population in general. We will therefore control for *M.AGE* (maternal age at the birth of the first child - centered at 22), which can also be used as a proxy for socioeconomic status. A closely related variable is maternal age at the birth of the respondent child (the correlation between *M.AGE* and maternal age at birth of child is .70 for children with at least one PIAT M score). Steelman (1985) pointed out the importance of considering this varaible in family structure and intelligence studies. Older mothers may, for example, be better emotionally prepared for children. Also, due to the increasing ages of the mothers in later years, the sample of young children in later years may not be comparable to the sample of young children in earlier years (NLSY79 child and young adult users guide, CHRR, 2006). We will focus on *M.AGE* in this paper, which controls for differences between families, however we included maternal age at the birth of the respondent child in some analyses as well. It was mostly non-significant and it did not change any of the conclusions, and is thus not presented in the result section.

We included *M.IQ* (maternal IQ measured by the Armed Forces Qualification Test in percentiles centered at 35) in the analyses because maternal IQ has been found to negatively relate to sibship size (e.g. Rodgers et al., 2000). Minorities were oversampled in the original NLSY sample, and we included *RACE(1)* (1 = Hispanic mother, 0 = non-Hispanic mother), and *RACE(2)* (1 = Black mother, 0 = non-Black mother) to control for, and investigate, differential levels for different racial groups. Finally, *FATHER* (1 = father present in household, 0 = father non–present in household) was included. Both the dilution theory and the confluence model predict higher intelligence for children in households with both parents present.

5.5 The samples

We constructed three samples from the NLSY children sample: the M-, RR-, and RC- samples. The children who had at least one valid score on the PIAT math (M), reading recognition (RR), or reading comprehension (RC) subtests, who lived most of their times with their mothers, and who were non-twins, were included in each respective sample. We also excluded families in which one or more of the biological children[5] died prior to, or during the span of the study. Because children with at least one score were included, some children contribute with only one assessment score, whereas others contribute with two, three, four, or five assessment scores. The time intervals between consecutive assessments need not be the same for different children. Table 3 shows some background information for the children in the samples[6] as well as for all children in the NLSY eligible for the PIAT assessments. We can see that the proportion of children with Hispanic and Black mothers, as well as average birth orders, maternal ages at birth of first child, and maternal IQ for the children are similar in the samples and for all children in the NLSY. As shown, there are about 50% non-Black non-Hispanic children in the samples, and as mentioned previously, these proportions are not representative of the whole U.S. population.

[5] i.e. children born to the mother.
[6] Means, standard deviations, and proportions were very similar in the three samples, and are here shown only for the M-sample.

Table 3. Proportions, means, standard deviations, and sample sizes for the M-sample as well as the entire NLSY sample

		M	NLSY
Race	Hispanic	P=.21	P=.21
	Black	P=.30	P=.30
	Non-B non-H	P=.48, N_{child} =8135	P=.49, N_{child} =8882
Birth order		M=1.9, SD=1.0, N_{child} =8135	M=1.9, SD=1.1, N_{child} =8882
Maternal age at birth of first child		M=21.7, SD=4.8, N_{child} =8135	M=21.6, SD=4.9, N_{child} =8882
Maternal IQ		M=35.0, SD=26.9, N_{child} =7796	M=34.4, SD=26.8, N_{child} =8483

Note: M = PIAT math, NLSY = NLSY-children sample (all children within age range, who reside most of their times with their mother, and who are non-twins), P = proportion, non-B non-H = non-Black non-Hispanic.

Table 4 shows mean M-, RR-, and RC scores for children at different ages. We can see that all PIAT scores increase for older ages. The table also shows mean sibship sizes and proportions of children who have their fathers present in the household in the M-sample. Five to six year olds have 1.6 siblings on average, whereas 13 to 14 year olds have 1.8 siblings on average. The proportion of children who have their fathers present are higher among younger children, whereas less than 50% of the 13 to 14 year olds have fathers

Table 4. Means, standard deviations, and proportions for selected variables by age in the M-, RR-, and RC-samples

Age	M-Score	RR-Score	RC-Score	Sibship Size[e]	Father Presence[f]
5-6	M=15.6	M=17.4	M=16.7	M=1.6	P=.65
	SD=6.7	SD=7.1	SD=6.3	SD=1.2	N_{child} =6196
	N_{child}=6225	N_{child} =6115	N_{child}=5886	N_{child}=6225	
7-8	M=30.2	M=33.2	M=31.0	M=1.7	P=.60
	SD=10.5	SD=10.8	SD=10.1	SD=1.2	N_{child}=6190
	N_{child} =6208	N_{child} =6193	N_{child} =5954	N_{child} =6208	
9-10	M=43.2	M=45.5	M=41.5	M=1.7	P=.55
	SD=10.4	SD=12.5	SD=10.8	SD=1.2	N_{child} =5879
	N_{child} =5899	N_{child} =5897	N_{child} =5827	N_{child} =5899	
11-12	M=50.2	M=54.3	M=48.4	M=1.8	P=.51
	SD=10.3	SD=13.8	SD=11.5	SD=1.2	N_{child} =5306
	N_{child} =5319	N_{child} =5307	N_{child} =5266	N_{child} =5319	
13-14	M=54.0	M=59.9	M=52.4	M=1.8	P=.47
	SD=11.1	SD=14.1	SD=12.1	SD=1.2	N_{child} =3709
	N_{child} =3718	N_{child} =3724	N_{child} =3699	N_{child} =3718	

Note: M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension, Score = score on PIAT subtest, Father Presence = father presence in household, P = proportion of children with father present in the household.

[e]Means and standard deviations were very similar for all samples, and descriptive statistics are therefore only shown for the M-sample.

[f] Proportions were very similar for all samples, and are therefore only shown for the M-sample.

present. More children, in general, have valid scores on M, followed by RR, and RC. Some discrepancies can be explained by children being assessed with one test first, and then getting tired and not being able to complete a second or third test. The generally lower sample sizes for RC can be explained, at least partly, by younger children not being able to read (RC), although they may still be able to understand and recognize some words (RR).

Table 5 shows, by race, the number of children with 1, 2, 3, 4, and 5 valid scores in the three samples respectively. As shown, the average number of valid scores are fairly similar across races, although Black children have been assessed more times, on average.

Table 5. The number of children with 1, 2, 3, 4 and 5 valid PIAT scores in the M-, RR-, and RC-samples, by race.

Score	M				RR				RC			
Race	H	B	n-B n-H	Sum	H	B	n-B n-H	Sum	H	B	n-B n-H	Sum
1	240	226	633	1099	238	221	642	1101	236	235	671	1142
2	219	361	747	1327	222	364	745	1331	234	378	745	1357
3	298	485	625	1408	296	498	617	1411	329	620	655	1604
4	525	678	910	2113	533	683	939	2155	530	585	966	2081
5	414	664	1110	2188	401	644	1079	2124	352	584	972	1908
Sum	1696	2414	4025	8135	1690	2410	4022	8122	1681	2402	4009	8092
Mean[g]	3.4	3.5	3.3	3.4	3.0	3.5	3.3	3.4	3.3	3.4	3.2	3.3

Note: PIAT = Peabody Individual Achievement Test, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension, H = Hispanic mother, B = Black mother, n-B n-H = non-Black non-Hispanic mother.
[g]The mean number of scores.

6. Statistical Models

We used multilevel models (see e.g. Goldstein, 2003) for our cross-sectional and longitudinal analyses[7]. Multilevel models can be used when data are structured in different levels. In our case, measurements are nested within children which are nested within families. We can therefore use a three level model with the repeated measures, children, and families at our first, second, and third levels. Correlations among repeated measures within children as well as among siblings within families can be handled by specifying a random intercept and time slope that vary between children and between families. We use four models with increasing complexity.

[7] see Wichman, Rodgers & MacCallum (2006) for an application of multilevel modeling to the NLSY-children data, in the context of a different design than the current study.

6.1 Cross-sectional models

First, let $SCORE_{ij}$ be the PIAT test score from the math (M)-, reading recognition (RR)-, or reading comprehension (RC) subtests, for the *i*th child in the *j*th family for a cross-section of the data. We can specify the following two-level model:

$$SCORE_{ij} = \gamma_{00j} + \gamma_{01}AGE_{ij} + \gamma_{02}AGE_{ij}^2 + \varepsilon_{ij}, \qquad \text{(model I)}$$

$$\varepsilon_{ij} \in iidN(0, \sigma_\varepsilon^2),$$

$$\gamma_{00j} = \beta_{000} + \beta_{001}SSS_j + u_{00j},$$

$$u_{00j} \in iidN(0, \sigma_{u_0}^2)$$

where AGE^2 is *AGE* squared and the other variables are as defined in the method section. We will use subscripts (e.g. *ij*) in the equations, however we will drop them, for simplicity, when we discuss variables in the text. As seen in model I, the intercept is random with variance $\sigma_{u_0}^2$ to account for correlations among children within the same families.

6.2 Longitudinal models

In order to investigate the effect of sibship size longitudinally, we used three models. The first and simplest model (see model II below) included sibship size at the time of each assessment (*SSS*), *AGE*, AGE^2, and *COHORT*. Let $SCORE_{ti}$ be the PIAT M-, RR-, or RC- score for the *t*th observation on the *i*th child. Model II is specified as follows

$$SCORE_{ti} = \pi_{0i} + \pi_{1i}AGE_{ti} + \pi_2 AGE_{ti}^2 + \pi_3 SSS_{ti} + \varepsilon_{ti}, \quad \varepsilon_{ti} \in iidN(0, \sigma_\varepsilon^2)$$

$$\pi_{0i} = \gamma_{00} + \gamma_{01}COHORT_i + r_{0i}$$

(model II)

$$\pi_{1i} = \gamma_{10} + r_{1i}$$

$$\begin{bmatrix} r_{0i} \\ r_{1i} \end{bmatrix} \in iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} \sigma_{r_0}^2 & \sigma_{r_0r_1} \\ \sigma_{r_0r_1} & \sigma_{r_1}^2 \end{pmatrix}\right]$$

The first row in model II specifies the first level, with time varying variables (*AGE*, *AGE*^{2}, and *SSS)*, and the second and third rows specify the second level with child varying variables (*COHORT*). Note that the sibship size (*SSS*) variable in the first row is assumed to have identical effects within and across children. Because longitudinal effects are often not the same as cohort effects (e.g. Diggle, Heagerty, Liang, & Zeger, 2002), we tested a third model in which we divided *SSS* into variables that vary between children, *OLD* and *YOUNG*, and within children, *BIRTH*. Model III is specified as follows

$$SCORE_{ti} = \pi_{0i} + \pi_{1i} AGE_{ti} + \pi_2 AGE_{ti}^2 + \pi_3 BIRTH_{ti} + \varepsilon_{ti}, \quad \varepsilon_{ti} \in iidN(0, \sigma_\varepsilon^2)$$

$$\pi_{0i} = \gamma_{00} + \gamma_{01} COHORT_i + \gamma_{02} OLD_i + \gamma_{03} YOUNG_i + r_{0i}$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11} OLD_i + \gamma_{12} YOUNG_i + r_{1i} \qquad \text{(model III)}$$

$$\begin{bmatrix} r_{0i} \\ r_{1i} \end{bmatrix} \in iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} \sigma_{r_0}^2 & \sigma_{r_0 r_1} \\ \sigma_{r_0 r_1} & \sigma_{r_1}^2 \end{pmatrix}\right].$$

As noted, we have included *OLD* and *YOUNG* in both the equations for the intercept (row 2) and the slope (row 3). We can thus investigate whether the age slope is different for children with different number of older and younger siblings.

We also introduced a fourth model in which we included a third level with family variables. Let $SCORE_{tij}$ be the PIAT M-, RR-, or RC- score for the *t*th observation on the *i*th child in the *j*th family. Model IV is specified as follows

$$SCORE_{tij} = \pi_{0ij} + \pi_{1ij} AGE_{tij} + \pi_2 AGE_{tij}^2 + \pi_3 BIRTH_{tij} + \pi_4 FATHER_{tij} + \varepsilon_{tij},$$

$$\varepsilon_{tij} \in iidN(0, \sigma_\varepsilon^2)$$

$$\pi_{0ij} = \gamma_{00j} + \gamma_{01} COHORT_{ij} + \gamma_{02} OLD_{ij} + \gamma_{03} YOUNG_{ij} + r_{0ij}$$

$$\pi_{1ij} = \gamma_{10j} + \gamma_{11} OLD_{ij} + \gamma_{12} YOUNG_{ij} + r_{1ij}$$

$$\begin{bmatrix} r_{0ij} \\ r_{1ij} \end{bmatrix} \in iidN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} \sigma_{r_0}^2 & \sigma_{r_0 r_1} \\ \sigma_{r_0 r_1} & \sigma_{r_1}^2 \end{pmatrix}\right] \qquad \text{(model IV)}$$

$$\gamma_{00j} = \beta_{000} + \beta_{001} M.AGE_j + \beta_{002} M.IQ_j + \beta_{003} RACE(1)_j + \\ + \beta_{004} RACE(2)_j + u_{00j}$$

$$\gamma_{10j} = \beta_{100} + \beta_{101} M.AGE_j + \beta_{102} M.IQ_j + \beta_{103} RACE(1)_j + \\ + \beta_{104} RACE(2)_j + u_{10j}$$

$$\begin{bmatrix} u_{00j} \\ u_{10j} \end{bmatrix} \in iidN \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \begin{pmatrix} \sigma^2_{u_0} & \sigma_{u_0 u_1} \\ \sigma_{u_0 u_1} & \sigma^2_{u_1} \end{pmatrix} \right].$$

Presence of the father (*FATHER*) is included at the first level because this variable can change over time. Maternal age at the birth of the first child (*M.AGE*), maternal IQ (*M.IQ*), and race (*RACE(1)*, *RACE(2)*) are included as family varying variables at the third level, and the intercept and *AGE* slope are random across children and families.

All multilevel analyses were conducted in SAS PROC MIXED (Littell, Milliken, Stroup, Wolfinger, Schabenberger, 2006). Introductory descriptions of how to use SAS PROC MIXED for multilevel models can be found, for example, in Singer (1998).

7. Results

7.1 Cross-sectional analyses

To replicate previous cross-sectional analyses, and to show that there is a consistent cross-sectional negative effect of sibship size on test score in the NLSY-children data, we plotted cross-sectional means for each subtest for sibship sizes one to five for ages 5-6, 7-8, 9-10, 11-12, and 13-14 in Figures 1 - 3. We can see that children with no siblings or one sibling scored highest, on average, for all age groups, followed by those that had two, three, four, or five siblings at the time of the assessment. We can also see that there appears to be a quadratic as well as a linear effect of age. These results replicate when test score is regressed onto age and sibship size as in model I – see Table 6. We can see that sibship size has a significant negative effect on all PIAT subtests across all years.

Figure 1. Cross-sectional PIAT Math means for different ages and sibship sizes.

Figure 2. Cross-sectional PIAT Reading Recognition means for different ages and sibship sizes.

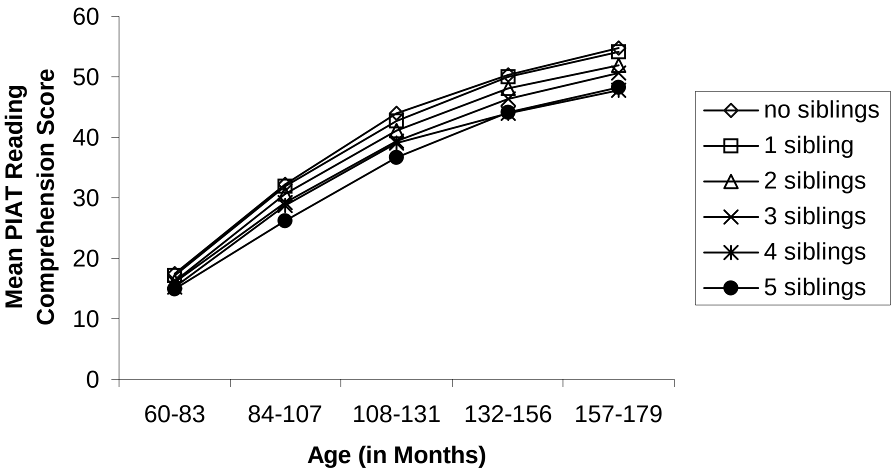

Figure 3. Cross-sectional PIAT Reading Comprehension means for different ages and sibship sizes.

Table 6. Model I: *SSS* regression slopes (β) from multilevel regression models controlling for *AGE* and AGE^2 by PIAT score and cross-section.

Cross section	M	RR	RC
1986	β = -1.21(0.19)***	β = -1.43(0.21)***	β = -1.41(0.20)***
	N_{family}= 1208, N_{child}= 1671	N_{family}= 1202, N_{child}= 1662	N_{family}= 1120, N_{child}= 1530
1988	β = -1.11(0.15)***	β = -1.42(0.18)***	β = -1.25(0.17)***
	N_{family}= 1861, N_{child}= 2824	N_{family}= 1850, N_{child}= 2803	N_{family}= 1805, N_{child}= 2705
1990	β = -1.05(0.16)***	β = -1.60(0.19)***	β = -152(0.18)***
	N_{family}= 1845, N_{child}= 2968	N_{family}= 1828, N_{child}= 2919	N_{family}= 1807, N_{child}= 2865
1992	β = -0.96(0.15)***	β = -1.41(0.19)***	β = -1.25(0.17)***
	N_{family}= 2052, N_{child}= 3382	N_{family}= 2030, N_{child}= 3327	N_{family}= 1984, N_{child}= 3187
1994	β = -1.04(0.15)***	β = -1.76(0.19)***	β = -1.52(0.16)***
	N_{family}= 2207, N_{child}= 3606	N_{family}= 2202, N_{child}= 3597	N_{family}= 2176, N_{child}= 3525
1996	β = -1.02(0.16)***	β = -1.32(0.20)***	β = -1.36(0.17)***
	N_{family}= 2087, N_{child}= 3358	N_{family}= 2079, N_{child}= 3349	N_{family}= 2067, N_{child}= 3307
1998	β = -1.15(0.17)***	β = -1.57(0.20)***	β = -1.48(0.17)***
	N_{family}= 1984, N_{child}= 3108	N_{family}= 1983, N_{child}= 3107	N_{family}= 1972, N_{child}= 3077
2000	β = -1.45(0.21)***	β = -1.45(0.25)***	β = -1.46(0.22)***
	N_{family}= 1488, N_{child}= 2289	N_{family}= 1489, N_{child}= 2290	N_{family}= 1487, N_{child}= 2278
2002	β = -1.53(0.19)***	β = -1.43(0.23)***	β = -1.62(0.19)***
	N_{family}= 1532, N_{child}= 2307	N_{family}= 1533, N_{child}= 2310	N_{family}= 1522, N_{child}= 2289
2004	β = -1.65(0.22)***	β = -1.55(0.26)***	β = -1.71(0.22)***
	N_{family}= 1283, N_{child}= 1856	N_{family}= 1291, N_{child}= 1872	N_{family}= 1290, N_{child}= 1869

Note. ***p<.001, ** p<.01, * p<.05
SSS = sibship size, PIAT = Peabody Individual Achievement Test, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension. Standard errors are shown in parentheses.

Rodgers et al. (2000) presented mean composite age normed PIAT scores (average scores across M, RR, and RC) for subsamples of the NLSY-children for different sibship sizes and birth orders and showed that the mean scores of children from intact families (i.e. families in which all siblings have PIAT scores) are statistically the same for different birth orders. The means differed, on the other hand, between children from different size families. They presented tables for two subsamples: children from intact families in 1990/1992 and children from intact families in 1994/1996. Their tables were reproduced in Table 1. As mentioned previously, the females in the original NLSY sample have had most of their children as of 2004, and we replicated Rodgers et al.'s analyses for 2002/2004, however with separate estimates for the subtests, by computing the average age-normed score[8] across 2002 and 2004 for each subtest. The means, standard deviations, and sample sizes are shown in Table 7. As shown, the means do not decrease for increasing birth orders. These results, together with the results from Rodgers et al. (2000) in Table 1, indicate that cross-sections of children from intact families do not differ in their age-normed PIAT scores across birth orders.

Table 7. PIAT Scores by Birth Order, Family Size, and Test Type for the 2002/2004 NLSY-Children Sample

Family size and type of score	First sibling			Second sibling			Third sibling			Fourth sibling		
One child	M	RR	RC	M	RR	RC	M	RR	RC	M	RR	RC
Mean	106.1	110.0	104.6									
SD	13.8	14.2	13.2									
N	236	238	225									
Two children												
Mean	107.7	110.0	103.9	109.3	110.3	106.8						
SD	13.5	13.9	12.1	12.0	13.1	12.2						
N	315	316	289	315	316	289						
Three children												
Mean	107.6	113.3	105.5	107.2	110.1	105.5	107.2	110.3	107.1			
SD	12.4	11.9	13.0	12.5	13.8	12.5	13.4	12.1	13.2			
N	80	80	72	80	80	72	80	80	72			
Four children												
Mean	111.3	112.2	103.0	113.0	113.2	105.6	109.9	107.6	106.1	107.3	108.9	109.1
SD	16.6	14.3	13.1	12.2	12.3	11.2	10.0	10.0	10.7	14.1	14.1	12.6
N	23	23	15	23	23	15	23	23	15	23	23	15

Note. PIAT = Peabody Individual Achievement Test, NLSY = National Longitudinal Survey of Youth, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension.

[8] There are age-normed PIAT scores available in the NLSY-children dataset. These scores have been normed against an external norming sample. For more information on the norming process, see the NLSY79 child and young adult users guide (CHRR, 2006).

Table 7 also indicates, however, that children from larger sibship sizes do not have lower scores. This is in contrast to what Rodgers et al. (2000) found, although it follows a trend noted in their tables. Mean differences for different sibship sizes are larger for the 1990/1992 sample than for the 1994/1996 sample. Maternal age at birth of child is necessarily higher for children who are five to 14 in 2002/2004 compared to earlier years. For example, children between five and 14 in 1992 were born to mothers who were at most 30 years old, whereas children between five and 14 in 2004 were born to mothers at least 32 years old. These differences can thus be due to differences in the samples. It is noteworthy, however, that the 2002 and 2004 cross-sections (see Table 6) show significant effects of sibship size when all children, and not just children of intact families and sibship sizes one to four, are included.

Most results from the cross-sectional analyses presented here indicate that sibship size is negatively related to PIAT scores. A possible causal interpretation is that a child's intelligence is affected negatively by having many siblings, and in particular that his or her intelligence declines, compared to what it would otherwise be, after an additional child is born into the family. We will now investigate this latter statement by conducting longitudinal analyses.

7.2 Longitudinal analyses

We first tested model II and the results are shown in Table 8. We can see that M-, RR-, and RC scores increase over time and sibship size (*SSS*) shows negative and significant effects for all three subtests. The intercepts and age slopes vary significantly among children[9]. Table 9 shows results from model III. When *SSS* is divided into between- and within child effects, the between child effects of the number of older siblings (*OLD*) are large, negative, and significant for all tests. The number of younger siblings at age five (*YOUNG*), as well as the number of additional siblings born at each assessment (*BIRTH*) also show negative and significant effects, although smaller than *OLD*. There are also two interactions included in Table 9 – the interaction between *OLD* and *AGE* (*OLD.AGE*) and between *YOUNG* and *AGE* (*YOUNG.AGE*). The logic of including these interactions into the model can be understood by substituting the second and third rows in model III into the first row (see e.g. Singer, 1998). These interactions are negative and signifi-

[9] The random components are significantly different from zero as seen in Table 8. However a better test is to examine the difference in deviance (-2lnLikelihood) between nested models, which is asymptotically distributed as χ^2 where the degrees of freedom are the number of additional estimated variance components. For M, the difference in deviance between a model with no variance components for the intercept and slope and model II is 198762.4 – 187472.79 = 11289.61 which indicates that the variance components are significant. For RR, the difference is 18149.7 and for RC, it is 10636, and these differences are also significant.

cant, with the exception of *YOUNG.AGE* for M. Children with more older and younger siblings thus have smaller *AGE* slopes on average for most PIAT subtests. The part of the explainable variance that is attributed to differences between children is 53.8%, 68.8%, and 49.0% for the M, RR, and RC tests respectively[10].

Table 8. Model II: *SSS* regression slopes from longitudinal multilevel regression models controlling for *AGE*, AGE^2, and *COHORT* by PIAT score.

Effects	M	RR	RC
	N_{family}=3770 N_{child}=8135	N_{family}=3767 N_{child}=8122	N_{family}=3759 N_{child}=8092
CONS	37.54 (0.15)***	40.19 (0.17)***	37.50 (0.15)***
AGE	0.49 (0.001)***	0.52 (0.002)***	0.44 (0.002)***
AGE^2	-0.003 (0.00004)***	-0.003 (0.00004)***	-0.003 (0.00004)***
COHORT	0.29 (0.01)***	0.23 (0.01)***	0.18 (0.01)***
Sibship size			
SSS	-1.03 (0.06)***	-0.95 (0.06)***	-1.07 (0.06)***
Residuals			
$\sigma^2_{r_0}$	38.90 (0.72)***	69.40 (1.32)***	36.85 (0.70)***
$\sigma_{r_0 r_1}$	0.32[h]	0.89 (0.02)***	0.39[h]
$\sigma^2_{r_1}$	0.003 (0.0002)***	0.01 (0.0004)***	0.004 (0.0002)***
σ^2_{ε}	33.49 (0.40)***	31.46 (0.39)***	37.96 (0.46)***
-2lnL	187472.79	189753.47	185474.15

Note. ***p<.001, ** p<.01, * p<.05
PIAT = Peabody Individual Achievement Test, CONS=constant, SSS = sibship size, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension. Standard errors are shown in parentheses.
[h]The correlation between the intercept and slope was estimated to be over one and was fixed to one.

[10] e.g. for M: $(38.44 + 2\times 0.31 + 0.003)/(38.44 + 2\times 0.31 + 0.003 + 33.49) = 0.538$.

Table 9. Model III: Sibship size regression slopes from longitudinal multilevel regression models controlling for *AGE*, *AGE^2*, and *COHORT* by PIAT score.

Effects	M	RR	RC
	N_{family}=3770 N_{child}=8135	N_{family}=3767 N_{child}=8122	N_{family}=3759 N_{child}=8092
CONS	37.44 (0.16)***	40.41 (0.19)***	37.52 (0.16)***
AGE	0.50 (0.002)***	0.54 (0.003)***	0.46 (0.003)***
AGE^2	-0.003 (0.00004)***	-0.003 (0.00004)***	-0.003 (0.00004)***
COHORT	0.33 (0.01)***	0.29 (0.02)***	0.25 (0.01)***
Sibship size			
OLD	-1.58 (0.08)***	-1.97 (0.10)***	-2.08 (0.08)***
YOUNG	-0.60 (0.12)***	-0.60 (0.15)***	-0.53 (0.12)***
BIRTH	-0.72 (0.12)***	-0.47 (0.14)**	-0.50 (0.13)***
Interactions			
OLD.AGE	-0.008 (0.001)***	-0.01 (0.002)***	-0.02 (0.002)***
YOUNG.AGE	-0.002 (0.002)	-0.01 (0.003)***	-0.006 (0.002)**
Residuals			
$\sigma^2_{r_0}$	38.44 (0.71)***	67.78 (1.29)***	35.63 (0.69)***
$\sigma_{r_0 r_1}$	0.31[h]	0.87 (0.02)***	0.37[h]
$\sigma^2_{r_1}$	0.003 (0.0002)***	0.01 (0.0004)***	0.004 (0.0002)***
σ^2_{ε}	33.49 (0.40)***	31.46 (0.39)***	37.90 (0.46)***
-2lnL	187419.62	189608.24	185245.27

Note. ***p<.001, ** p<.01, * p<.05
PIAT = Peabody Individual Achievement Test, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension, CONS=constant, OLD = number of siblings at birth, YOUNG = number of additional siblings at age 5, BIRTH = number of children born since age 5. Standard errors are shown in parentheses.
[h]The correlation between the intercept and slope was estimated to be over one and was fixed to one.

When variance attributed to differences between families is introduced in model IV (see Table 10), the *BIRTH* effect decreases, but is still significant for M. Effects of the number of older (*OLD*) and younger (*YOUNG*) siblings also decrease but are still significant for all subtests. We can also see that maternal age at birth of first child (*M.AGE*) has positive effects on M and RR, and maternal IQ (*M.IQ*) has positive effects on all subtests. Children with Hispanic mothers have on average lower M scores, but higher RR scores, compared to children with non-Black non-Hispanic mothers (see *RACE(1)*), whereas children with Black mothers have lower scores on M and RC (*see RACE(2)*) when all variables are included. Presence of the father (*FATHER*) is positive and significant for RR and RC, however not for M. The variances due to differences between children within families are now 27.5%, 35.8%, 23.7%, and the variances due to differences between families

Table 10. Model IV: Sibship size regression slopes from longitudinal multilevel regression models controlling for *AGE*, *AGE*^{2}, *COHORT*, and family variables by PIAT score.

Effects	M	RR	RC
	N_{family}=3616 N_{child}=7787	N_{family}=3612 N_{child}=7776	N_{family}= 3604 N_{child}=7747
CONS	38.49 (0.20)***	40.03 (0.25)***	37.66 (0.21)***
AGE	0.50 (0.003)***	0.54 (0.004)***	0.47 (0.003)***
AGE^2	-0.003 (0.00004)***	-0.003 (0.00004)***	-0.003 (0.00004)***
COHORT	0.11 (0.02)***	0.13 (0.03)***	0.06 (0.02)**
Sibship size			
OLD	-0.57 (0.11)***	-1.05 (0.13)***	-1.24 (0.11)***
YOUNG	-0.56 (0.11)***	-0.36 (0.14)**	-0.49 (0.12)***
BIRTH	-0.32 (0.12)**	-0.16 (0.14)	-0.23 (0.13)
Family variables			
M.AGE	0.13 (0.03)***	0.10 (0.04)**	0.05 (0.03)
M.IQ	0.11 (0.004)***	0.12 (0.005)***	0.11 (0.004)***
RACE(1)	-1.08 (0.25)***	0.65 (0.32)*	0.12 (0.26)
RACE(2)	-1.84 (0.23)***	-0.19 (0.30)	-0.48 (0.24)*
FATHER	0.14 (0.14)	0.60 (0.14)***	0.54 (0.14)***
Interactions			
OLD.AGE	0.00005 (0.002)	-0.005 (0.002)*	-0.01 (0.002)***
YOUNG.AGE	0.002 (0.002)	-0.004 (0.003)	-0.003 (0.002)
M.AGE.AGE	0.002 (0.0004)***	0.001 (0.0005)**	-0.0004 (0.0004)
M.IQ.AGE	0.0008 (0.00007)***	0.001 (0.00009)***	0.001 (0.00008)***
RACE(1).AGE	-0.005 (0.004)	0.02 (0.006)***	0.004 (0.005)
RACE(2).AGE	-0.02 (0.004)***	-0.04 (0.005)***	-0.05 (0.004)***
Residuals			
$\sigma^2_{r_0}$	17.20 (0.60)***	29.79 (0.73)***	15.80 (0.62)***
$\sigma_{r_0 r_1}$	0.11[h]	0.37[h]	0.13[h]
$\sigma^2_{r_1}$	0.0007 (0.0002)***	0.005 (0.0003)***	0.001 (0.0003)***
$\sigma^2_{u_0}$	12.47 (0.68)***	22.85 (1.10)***	13.62 (0.72)***
$\sigma_{u_0 u_1}$	0.13 (0.008)***	0.28 (0.01)***	0.17 (0.009)***
$\sigma^2_{u_1}$	0.002 (0.0002)***	0.005 (0.0003)***	0.002 (0.0002)***
σ^2_{ε}	33.15 (0.40)***	31.29 (0.39)***	37.82 (0.46)***
-2lnL	177670.90	180027.07	175818.57

Note. ***p<.001, ** p<.01, * p<.05
PIAT = Peabody Individual Achievement Test, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension, CONS=constant, OLD = number of siblings at birth, YOUNG = number of additional siblings at age 5, BIRTH = number of children born since age 5, M.AGE = maternal age at birth of first child, M.IQ = maternal IQ, RACE(1) = mother Hispanic (1) or not (0), RACE (2) = mother Black (1) or not (0), FATHER = father present in household. Standard errors are shown in parentheses.
[h]The correlation between the intercept and slope was estimated to be over one and was fixed to one.

are 20.1%, 27.5%, and 20.6% respectively for the three tests[11]. Because not all children in our samples had measurements on all background variables, the sample sizes are smaller for model IV compared to model III, however.

Results from model IV show that there are still some significant birth-of-a-sibling effects after age five when family variables are included in the model. As mentioned in the method section, however, this variable (*BIRTH*) picks up some between child effects because several children had siblings born between age five and their first assessment. We thus reanalyzed model IV excluding children who had siblings born between age five and their first assessment. This left 3532 families with 6774 children for M, 3527 families with 6758 children for RR, and 3519 families with 6708 children for RC respectively. The regression coefficients are still significant for *OLD* (for M: -0.57, $p < .001$, for RR: -1.10, $p < .001$, for RC: -1.28, $p < .001$) and *YOUNG* (for M: -0.53, $p < .001$, for RR: -0.33, $p < .05$, for RC: -0.40, $p < .01$) but they decrease and are no longer significant for either subtest for *BIRTH* (for M: -0.17, $p = .289$, for RR: -0.02, $p = .900$, for RC: -0.15, $p = .384$). We also ran analyses in which we defined *YOUNG* as the number of younger siblings born to the mother (as answered by the mother) at the assessment year that the child was five or six. *BIRTH* is thus the additional number of siblings born since the assessment year that the child was five or six years old[12]. These analyses also showed negative and significant effects of the number of older and younger children, however even smaller, non-significant, effects of the birth of a sibling.

As mentioned above, effects of the number of siblings at birth (*OLD*), as well as the additional number of siblings at five (*YOUNG*) were negative and significant in the above analyses. Because children were assessed no earlier than age five with the PIAT subtests, we cannot investigate birth of a sibling effects longitudinally prior to age five for these tests. We instead tested a new model, model V, where we divided *OLD* into the number of siblings at most 23 months older (*OLD0-2*) and the number of siblings 24 months or older (*OLD2+*) than the child. We also divided *YOUNG* into the number of siblings at most 23 months younger (*YOUNG0-2*) and the number of additional siblings born at age five (*YOUNG2+*). The results are shown in Table 11. We can see that the number of older siblings have negative and significant effects on all subtests, as has the number of closely spaced younger siblings (*YOUNG0-2*). In addition, the slopes for the number of closely

[11] Because the original NLSY sample consisted of households, some children in the M-, RR-, and RC-samples are cousins. To account for this, we also tested a model in which we added a fourth level accounting for different levels across original households. The coefficients did not change much, however, and the conclusions remain the same. We therefore do not present these results here.

[12] These ages were chosen because assessments took place every second year, and children were thus 5 or 6 years old at their first possible assessment.

spaced older siblings (*OLD0-2*) are larger than the slopes for the number of widely spaced older siblings (*OLD2+*). The part of the explainable variance that is attributed to differences between children is 53.7%, 68.8%, and 48.9% for the M, RR, and RC tests respectively.

Table 11. Model V: Sibship size regression slopes from longitudinal multilevel regression models controlling for *AGE*, AGE^2, and *COHORT* by PIAT score.

Effects	M	RR	RC
	N_{family}=3770 N_{child}=8135	N_{family}=3767 N_{child}=8122	N_{family}=3759 N_{child}=8092
CONS	37.41 (0.16)***	40.38 (0.19)***	37.49 (0.16)***
AGE	0.50 (0.002)***	0.54 (0.003)***	0.46 (0.003)***
AGE^2	-0.003 (0.0004)***	-0.003 (0.0004)***	-0.003 (0.0004)***
COHORT	0.33 (0.01)***	0.29 (0.02)***	0.25 (0.01)***
Sibshipsize			
OLD(2+)	-1.51 (0.09)***	-1.84 (0.11)***	-2.03 (0.09)***
OLD(0-2)	-1.78 (0.22)***	-2.54 (0.27)***	-2.19 (0.21)***
YOUNG (0-2)	-1.47 (0.21)***	-1.56 (0.27)***	-1.35 (0.21)***
YOUNG (2+)	-0.19 (0.14)	-0.12 (0.18)	-0.15 (0.14)
BIRTH	-0.72 (0.12)***	-0.46 (0.14)**	-0.50 (0.13)***
Interactions			
OLD(2+).AGE	-0.006 (0.002)***	-0.01 (0.002)***	-0.02 (0.002)***
OLD(0-2).AGE	-0.02 (0.004)***	-0.03 (0.005)***	-0.03 (0.004)***
YOUNG(0-2).AGE	-0.005 (0.004)	-0.02 (0.005)***	-0.01 (0.004)***
YOUNG(2+).AGE	0.00009 (0.002)	-0.003 (0.003)***	-0.002 (0.003)
Residuals			
$\sigma^2_{r_0}$	38.29 (0.71)***	67.56 (1.29)***	35.51 (0.69)***
$\sigma_{r_0 r_1}$	0.31[h]	0.87 (0.02)***	0.37[h]
$\sigma^2_{r_1}$	0.003 (0.0002)***	0.01 (0.0004)***	0.004 (0.0002)***
σ^2_{ε}	33.49 (0.40)***	31.46 (0.39)***	37.90 (0.46)***
-2lnL	187403.51	189587.24	185244.08

Note. ***p<.001, ** p<.01, * p<.05
PIAT = Peabody Individual Achievement Test, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension, CONS=constant, OLD(2+) = number of siblings at least 24 months older, OLD(0-2) = number of siblings at most 23 months older, YOUNG(0-2) = number of siblings at most 23 months younger, YOUNG(2+)= number of additonal children born at age 5, BIRTH = number of children born since age 5. Standard errors are shown in parentheses.
[h]The correlation between the intercept and slope was estimated to be over one and was fixed to one.

Table 12 shows results for an additional model, model VI, where we also entered family variables. We can now see that all sibship size slopes have decreased, and the *BIRTH* effects are no longer significant for RR and RC. The variances due to differences between children within families are now 27.6%, 35.0%, 23.6%, and the variances due to differences between families are 20.1%, 27.8%, and 20.6%.

Table 12. Model VI: Sibship size regression slopes from longitudinal multilevel regression models controlling for *AGE*, *AGE*^{2}, *COHORT*, and family variables by PIAT score.

Effects	M N_{family}=3616 N_{child}=7787	RR N_{family}=3612 N_{child}=7776	RC N_{family}=3604 N_{child}=7747
CONS	38.52 (0.21)***	40.06 (0.25)***	37.69 (0.21)***
AGE	0.50 (0.003)***	0.54 (0.004)***	0.47 (0.003)***
AGE^2	-0.003 (0.0004)***	-0.003 (0.0004)***	-0.003 (0.0004)***
COHORT	0.10 (0.02)***	0.12 (0.03)***	0.05 (0.02)*
Sibshipsize			
OLD(2+)	-0.45 (0.12)***	-0.90 (0.14)***	-1.16 (0.12)***
OLD(0-2)	-1.05 (0.19)***	-1.82 (0.24)***	-1.57 (0.20)***
YOUNG (0-2)	-0.91 (0.19)***	-0.94 (0.23)***	-0.85 (0.19)***
YOUNG (2+)	-0.42 (0.13)**	-0.13 (0.16)	-0.34 (0.13)*
BIRTH	-0.29 (0.12)*	-0.12 (0.14)	-0.21 (0.13)
Family variables			
M.AGE	0.14 (0.03)***	0.11 (0.04)**	0.06 (0.03)
M.IQ	0.11 (0.004)***	0.12 (0.005)***	0.11 (0.004)***
RACE(1)	-1.07 (0.25)***	0.66 (0.32)*	0.13 (0.26)
RACE(2)	-1.83 (0.23)***	-0.16 (0.30)	-0.46 (0.24)
FATHER	0.14 (0.14)	0.60 (0.14)***	0.54 (0.14)***
Interactions			
OLD(2+).AGE	0.002 (0.002)	-0.002 (0.002)	-0.01 (0.002)***
OLD(0-2).AGE	-0.01 (0.004)**	-0.02 (0.004)***	-0.02 (0.004)***
YOUNG(0-2).AGE	0.004 (0.004)	-0.01 (0.004)*	-0.006 (0.004)
YOUNG(2+).AGE	0.0008 (0.002)	-0.001 (0.003)	-0.002 (0.003)
M.AGE.AGE	0.002 (0.0004)***	0.002 (0.0005)**	-0.0003 (0.0004)
M.IQ.AGE	0.0008 (0.00007)***	0.001 (0.00009)***	0.001 (0.00008)***
RACE(1).AGE	-0.003 (0.004)	0.02 (0.006)***	0.004 (0.005)
RACE(2).AGE	-0.02 (0.004)***	-0.04 (0.005)***	-0.05 (0.004)***
Residuals			
$\sigma^2_{r_0}$	17.21 (0.60)***	28.75 (0.73)***	15.77 (0.62)***
$\sigma_{r_0 r_1}$	0.11[h]	0.37[h]	0.13[h]
$\sigma^2_{r_1}$	0.0007 (0.0002)***	0.004 (0.0003)***	0.001 (0.0003)***
$\sigma^2_{u_0}$	12.43 (0.68)***	22.83 (1.10)***	13.62 (0.72)***
$\sigma_{u_0 u_1}$	0.13 (0.008)***	0.28 (0.01)***	0.17 (0.009)***
$\sigma^2_{u_1}$	0.002 (0.0002)***	0.005 (0.0003)***	0.002 (0.0002)***
σ^2_{ε}	33.14 (0.40)***	31.29 (0.39)***	37.81 (0.46)***
-2lnL	177674.28	180019.68	175831.05

Note. ***p<.001, ** p<.01, * p<.05
PIAT = Peabody Individual Achievement Test, M = PIAT math, RR = PIAT reading recognition, RC = PIAT reading comprehension, CONS=constant, OLD(2+) = number of siblings at least 24 months older, OLD(0-2) = number of siblings at most 23 months older, YOUNG(0-2) = number of siblings at most 23 months younger, YOUNG(2+)= number of additonal children born at age 5, BIRTH = number of children born since age 5, M.AGE = maternal age at birth of first child, M.IQ = maternal IQ, RACE(1) = mother Hispanic (1) or not (0), RACE (2) = mother Black (1) or not (0), FATHER = father present in household. Standard errors are shown in parentheses.
[h]The correlation between the intercept and slope was estimated to be over one and was fixed to one.

To control for some differential parental values, we also included desired family size as answered by the mother in 1979 in models IV and VI. This variable showed no significant effects, however, and is therefore not reported here. The females were between 14 and 21 in 1979 and their responses to this question may therefore not be very indicative of future desires or actions. It is also likely that desired family size responses of 14 year olds have different meanings than responses of 21 year olds.

8. Discussion

This study extended previous longitudinal work (e.g. Baydar, Greek, & Brooks-Gunn, 1997; McCall, 1984; Guo & VanWey, 1999) that has tried to examine whether the consistent negative correlation between sibship size and cognitive ability found cross-sectionally still holds when looked at longitudinally, that is whether the birth of a sibling negatively affects children's cognitive abilities. We used multilevel models to account for the longitudinal structure of the data as well as for children nested within families. There were up to four siblings born during the span of the current study. We accounted for fairly large additions to family size, and larger additions compared to Guo and VanWey (1999). We also included all children with at least one valid PIAT score between the years 1986 and 2004. We found a consistent negative relationship between sibship size and cognitive ability across different years for three cognitive ability tests: PIAT math, PIAT reading recognition, and PIAT reading comprehension. When the sibship size measure was partitioned into between- and within child variables, however, the between child variables, such as the number of older and younger siblings, still had significant effects on the subtests, whereas the within variable, birth-of-a-sibling[13], had no such effects once differences between families were controlled.

8.1 Conclusions

In general, the results of the longitudinal analyses support Guo and VanWey's (1999) findings of no negative birth effects, providing further support to the notion that the observed cross-sectional effects may be due to uncontrolled differences between large and small families. This, in turn, supports the admixture hypothesis (Page & Grandon, 1979; Velandia, Grandon, & Page, 1978; Rodgers et al., 2000; Rodgers, 2001).

Downey et al. (1999) suggested that the sibship size effect might only be present for closely spaced siblings, and should thus be examined for young

[13] When we excluded children with siblings born after age 5 but before their first assessment, no birth-of-a-sibling effects were significant.

children. As they pointed out, closely spaced siblings are incorporated into cross-sectional analyses, but sometimes not in longitudinal analyses if the children had to be several years old at their first cognitive assessment. It is difficult to conduct longitudinal analyses testing the birth of a sibling on cognitive ability for closely spaced siblings because the child of interest then must be cognitively assessed at a very young age. Cognitive measures at very young ages are not very reliable, making it hard to use them as cognitive ability indicators in longitudinal studies. This study tried to account for some of this close spacing problem cross-sectionally by dividing the number of older and younger siblings into closely spaced siblings (less than two years older or younger) and widely spaced siblings. We could see that the magnitudes of the slopes for the number of closely spaced siblings were larger than for the number of widely spaced siblings.

The mean tables reproduced from Rodgers et al. (2000) and replicated here several years later show, however, that effects of older and younger siblings do not lie within the family, at least not within the intact families assessed in these tables. There may thus be other reasons for the above results. Families that have closely spaced siblings may differ from other families, for example by differential planning, delaying of education for the mother etc. There was an interesting pattern in the tables, however. Differences in means of different size families are larger for earlier years, and appear to have vanished in 2002/2004. In contrast, the sibship size slopes in Table 6 show negative effects of sibship size for all cross-sections, and the magnitude of the slopes does not appear to decrease for later years. Tables 1 and 7 only include children from intact families, i.e. children from families in which all children were cognitively assessed in certain years. Families with very widely spaced siblings are therefore not included, and family sizes larger than four or five are also not included (because of sample sizes). In addition, maternal ages at birth of the respondent child, as well as at the birth of the first child, are higher for later cross-sections because the ages of the children in the families are restricted to be between five and 14 years. A possible conclusion may thus be that family size (0 to 4) and cognitive ability are not related for children born to slightly older mothers, and whose mothers also were at least 23 years old when they had their first child. This should be investigated further, however.

Baydar, Greek, and Brooks-Gunn (1997) found that young children who had at least one sibling born during an early age had lower scores on a cognitive test compared to those who did not have an additional sibling, controlling for another test, completed by the mother at baseline, designed to measure social, motor, and cognitive development in young children. They concluded that these lower scores were associated with a decline in positive interactions with the mother, and an increase in a punishing parenting style. They also found that these children had lower scores on the PIAT math and reading recognition tests after four years. Our results also showed lower scores for children who had a sibling born during the first couple of years in life, although we did not control for motor and social development scores. The results in the present paper however agree with their findings that children who have closely spaced younger siblings have lower scores on average.

As expected, we found effects of several family variables on PIAT scores. Children whose mothers were older when they had their first child had higher scores on average, as did children whose mothers had higher IQ scores. Children with Black mothers had in general lower scores, whereas children with Hispanic mothers had higher scores on average on the PIAT reading recognition subtest, after controlling for other family variables. We investigated this further, and found that the previously positive and significant slope for *RACE(1)* (see Table 10) became negative and significant when maternal IQ was dropped from the model. Children with fathers living in their households also had significantly higher scores except for the math subtest. The proportion of variance attributed to differences between children was reduced when these family variables were accounted for, and all sibship size effects decreased. Although we acknowledge that there may be some real changes in cognitive ability following the birth of a sibling, especially if that sibling is born early in life, large proportions of the relationship between sibship size and cognitive ability can be attributed to differences between families.

8.2 Limitations of the current study and suggestions for future work

In generalizing these results to other children, we need to take into consideration that children from minority groups were oversampled. Also, attrition rates in Table 2 show that 20% of the children of interviewed mothers did not obtain valid PIAT scores in 2004. In addition, around 20% of the eligible females were not interviewed in 2004 as well. Although mentally handicapped children were not excluded from PIAT assessments per se, in practice, several of these children may have failed to obtain valid scores either because the mother requested they not be assessed, or because they failed to complete the test. We cannot say how additional siblings may affect mentally handicapped children from the present study.

We can also see, from Table 4, that the proportion of children, in the three samples, with different numbers of valid scores differs slightly across race. The NLSY79 child and young adult users guide (CHRR, 2006) report generally higher completion rates for children with Black mothers and generally lower completion rates for children with Hispanic mothers for the PIAT math, reading recognition, and reading comprehension subtests. We did not include any sampling weights in our analyses, which were necessarily different for children at different assessment years because they represented a different proportion of children depending on response rates. It should therefore be noted that minority groups are overrepresented in our analyses.

The present study included several family variables that accounted for parts of the family variance, although several other variables could be considered as well. Desired family size was included, although it was not found to affect the cognitive ability measures and therefore not presented. Future studies might instead look at whether or not the children were planned and thus try to control for more possible family differences. Phillips (1999) suggested that parents' mental health or temperament might affect both family size and cognitive ability. She suggested that parents' self-esteem or ability to plan for the future might influence birth control, rather than parental values. In addition, she proposed that children's temperaments might affect parents' choices to have more children. Certain dimensions of temperament have been found to correlate with certain dimensions of cognitive ability (e.g. Strelau, Zawadsky, and Piotrowska, 2001), which could be investigated further in the NLSY because the mothers of the children were given temperament questionnaires about their children during several years.

Guo and VanWey (1999) suggested that home environment may be one of the between family factors that may cause a spurious relationship between sibship size and cognitive ability. This can be investigated further in the NLSY because home environment scales are available for the children. Factors outside of the home, such as the influence of school, friends, and leisure activities on children's cognitive abilities, as well as interactions between home environments, birth of a sibling, and outside environments, can also be examined. It is possible that those that experience "negative" effects at home may be buffered if they have a good school and peer environment and vice versa. Home environments may play a more important role early in life, whereas outside environments may play more important roles as the child grows.

The effects of additional siblings may not be linear. We found cross-sectional effects of closely spaced siblings, in particular. Birth of a sibling effects on cognitive ability may depend both on the child's age, but also on whether it is the first, second, or third etc. younger sibling being born into the family. This can be investigated further by including both the child's age at the birth of each sibling, but also dummy variables indicating whether it is the first, second, third etc. younger sibling.

We looked at three PIAT subtests separately. Intelligence is a construct that preferably can be measured by several indicators. We have conducted some preliminary factor analytic studies on parts of the data, however, which indicated that the three subtests did not show factorial invariance over time (see e.g. Sayer & Cumsille, 2002 for a description of factorial invariance over time). Because of this, and because sibling effects appeared to be different for the subtests, we decided to analyze them separately. Future studies might investigate this further, however, possibly looking at individual items of the subtests that might be used for factor analytic studies[14].

We only included biological (children born to the mother) siblings in our study, partly because exact ages of these siblings could be obtained. However, birth of a sibling effects can also include other children living in the household, such as stepchildren, foster-children etc. Future studies can include all children living in a household, and also investigate possible changes in cognitive abilities as other children (such as step-children) move into the household, move out etc. Because such transitions often occur following marriages or divorces or the like, factors such as these should be considered as well.

Sibling death is another factor that changes the family composition, however this can be more difficult to study. A test assessing the cognitive ability of a child that recently experienced the death of a sibling is probably not very reliable, making it hard to compare the child's score before and after his or her sibling died. Kristensen and Bjerkedal (2007) used a very clever design to study effects of birth order on intelligence. They studied first- second- and third-born Norwegian male conscripts who had no, one, or two older siblings who died in infancy. They found that the birth orders of males who had had older siblings that died behaved as their "social" birth orders, i.e. second-borns with an older sibling who died had average intelligence scores on the same level as first-borns etc. They compared males from different families, however, and families that experience deaths may be different from other families.

The dilution theory suggests that parental resources are diluted as more siblings are born. Children with various handicaps, and who are in need of extra attention, might dilute parental resources more than other siblings. Future research might investigate differential effects of siblings with handicaps. Downey (2001) also suggested that the only birth-of-a-sibling effect might be with additions of five or more siblings. The children in the present study had at most four siblings born during their assessment span, making this hard to study. Downey (2001) also pointed out the necessity to study possible long-term sibling effects. Holmgren, Molander, and Nilsson (2007) investigated long-term (although cross-sectional) effects of sibship size in adulthood, and found that the relationship was affected by the respondent's education, although they noted that cognitive ability could affect educational level as well. As mentioned previously, Kuo and Hauser (1997) found that

[14] Second order latent curve models (Duncan & Duncan, 1996; McArdle, 1988) can, for example, be used to analyze growth in latent constructs.

siblings from smaller families had more years of education. We studied sibship size and birth-of-a-sibling effects on a sample of children with a restricted age range, and effects after 14 years of age, such as effects on educational outcomes (e.g. college attendance) were therefore not studied. It may be possible to look at long-term effects of sibship size in the NLSY-children sample and to include the respondent's educational level by examining the children as they grow older and complete their education.

9. References

Anastasi, A. (1956). Intelligence and family size. *Psychological Bulletin, 53*, 187-209.

Armor, D. J. (2001). On family size and intelligence. *American Psychologist, 55*, 521-522.

Baydar, N., Greek, A. & Brooks-Gunn, J. (1997). A longitudinal study of the effects of the birth of a sibling during the first 6 years of life. *Journal of Marriage and the Family, 59*, 939-956.

Baydar, N., Hyle, P. & Brooks-Gunn, J. (1997). A longitudinal study of the effects of the birth of a sibling during preschool and early grade school years. *Journal of Marriage and the Family, 59*, 957-965.

Belmont, L. & Marolla, F. (1973). Birth order, family size, and intelligence. *Science, 182*, 1096-1101.

Blake, J. (1981). Family size and the quality of children. *Demography*, *18*, 421-442.

Center for Human Resource Research (2006). *NLSY79 user's guide*. Columbus, OH: Ohio State University.

Center for Human Resource Research (2006). *NLSY79 child & young adult data users guide*. Columbus, OH: Ohio State University.

Diggle, P. J., Heagerty, P. J., Liang, K., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford Statistical Science Series. Oxford University Press Inc., NY.

Downey, D. B. (1995). When bigger is not better: Family size, parental resources, and children's educational performance. *American Sociological Review, 60*, 747-761.

Downey, D. B., Powell, B., Steelman, L. C. & Pribesh, S. (1999). Much ado about siblings: change models, sibship size, and intellectual development. Comment on Guo and VanWey. *American Sociological Review, 64*, 193-198.

Downey, D. B. (2001). Number of siblings and intellectual development. The resource dilution explanation. *American Psychologist, 56*, 497-504.

Duncan, S. C., & Duncan, T. E. (1996). A multivariate growth curve analysis of adolescent substance use. *Structural Equation Modeling, 3*, 323-347.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*, 171-191.

Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). New York: Halsted.

Guo, G. & VanWey, L. K. (1999). Sibship size and intellectual development: is the relationship causal? *American Sociological Review, 64,* 169-187.

Higgins, J. V., Reed, E. W. & Reed, S. C. (1962). Intelligence and family size: a paradox resolved. *Eugenics Quarterly, 9*, 84-90.

Holmgren, S., Molander, B., & Nilsson, L (2006). Intelligence and executive functioning in adult age: effects of sibship size and birth order. *European Journal of Cognitive Psychology. Special Issue, 18*, 138-158.

Kristensen, P., & Bjerkedal, T. (2007). Explaining the relation between birth order and intelligence. *Science, 36*, 1717.

Kuo, H. D. & Hauser, R. M. (1997). How does size of sibship matter? Family configuration and family effects on educational attainment. *Social Science Research, 26*, 69-94.

Littell, R., D., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models* (2nd edition). SAS Institute Inc., Cary, NC.

Mascie-Taylor, C. G. N. (1980). Family size, birth order and IQ components: a survey of a Cambridge suburb. *Journal of Biosocial Science, 12*, 309-312.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In R. B. Cattell & J. Nesselroade (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 564-614). New York: Plenum Press.

McCall, R. B. (1984). Developmental changes in mental performance: the effect of the birth of a sibling. *Child Development, 55,* 1317-1321.

McCall, R. B. (1985). The confluence model and theory. *Child Development, 56*, 217-218.

Mercy. J. A. & Steelman, L. C. (1982). Familial influences on the intellectual attainment of children. *American Sociological Review, 47,* 532-542.

Nisbet, J. D. & Entwistle, N, J. (1967). Intelligence and family size, 1949-1965. *British Journal of Educational Psychology, 37*, 188-193.

Page, E. B. & Grandon, G. M. (1979). Family configuration and mental ability: two theories contrasted with U.S. data. *American Educational Research Journal, 16*, 257-272.

Phillips, M. (1999). Sibship size and academic achievement: what we know and what we still need to know. *American Sociological Review, 64,* 188-192.

Pong, S. (1997). Sibship size and educational attainment in peninsular Maylasia: do policies matter? *Sociological Perspectives, 40*, 227-242.

Powell, B. & Steelman, L. C. (1990). Beyond sibship size: sibling density, sex composition, and educational outcomes. *Social Forces, 69*, 181-206.

Rankin, R. J., Gaite, A. J. H. & Heiry, T. (1979). Cultural modification of effect of family size on intelligence. *Psychological Reports, 45,* 391-397.

Rodgers, J. L., Cleveland, H. H., van den Oord, E. & Rowe, D. C. (2000). Resolving the debate over birth order, family size, and intelligence. *American Psychologist, 55,* 599-612.

Rodgers, J. L (2001). What causes birth order – intelligence patterns? The admixture hypothesis, revived. *American Psychologist, 56*, 505-510.

Rodgers, J. L. & Wänstrom, L. (2007). Identification of a Flynn Effect in the NLSY: Moving from the Center to the Boundaries. *Intelligence, 35,* 187-196.

Sayer, A. G. & Cumsille, P. E. (2001). Second-order latent growth models. In: Collins, L. M. & Sayer, A. G. (Eds.). *New methods for the analysis of change* (pp. 177-200). Washington, DC; American Psychological Association.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth curves. *Journal of Educational and Behavioral Statistics, 24,* 323-355.

Steelman, L. C. (1985). A tale of two variables: a review of the intellectual consequences of sibship size and birth order. *Review of Educational Research, 55,* 353-386.

Steelman, L. C. & Mercy, J. A. (1980). Unconfounding the confluence model: a test of sibship size and birth-order effects on intelligence. *American Sociological Review, 45*, 571-582.

Strelau, J., Zawadzki, B. & Piotrowska, A. (2001). Temperament and intelligence: a psychometric approach to the links between both phenomena. In J. M. Collis & S. Messik (Eds.) *Intelligence and personality. Bridging the gap in theory and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Velandia, W., Grandon, G. M. & Page, E. B. (1978). Family size, birth order, and intelligence in a large South American sample. *American Educational Research Journal, 15*, 399-416.

Wichman, A. L., Rodgers, J. L., & MacCallum, R. C. (2006). A multilevel approach to the relationship between birth order and intelligence. *Personality and Social Psychology Bulletin, 32*, 117-127.

Zajonc, R. B. & Markus, G. B. (1975). Birth order and intellectual development. *Psychological Review, 82*, 74-88.

Zajonc, R. B., Markus, G. B., Berbaum, M. L., Bargh, J. A. & Moreland, R. L. (1991). One justified criticism plus three flawed analyses equals two unwarranted conclusions: a reply to Retherford and Sewell. *American sociological Review, 56*, 159-165.

Zajonc, R. B. & Mullally, P. R. (1997). Birth order. Reconciling conflicting effects. *American Psychologist, 52*, 685-699.

Zajonc, R. B. (2001). The family dynamics of intellectual development. *American Psychologist, 56*, 490-496.