Intelligence and Age at First Intercourse: Cause or Confound?

S. Mason Garrison and Joseph Lee Rodgers

Vanderbilt University

Author Note

This material is based upon work that has been supported by National Institute of Health under Grant No. (R01-HD065865) and the National Science Foundation Graduate Research Fellowship Program under Grant No. (DGE-1445197) and various means of institutional support from the following universities: University of British Columbia, University of Oklahoma & Vanderbilt University.

Abstract

Your abstract here.

2015/10/15 at 12:21:25

Intelligence and Age at First Intercourse: Cause or Confound?

Anecdotal evidence from the popular media, such as MTV's reality television franchise, *16 and Pregnant*, suggests that teenage promiscuity is on the rise. Academic evidence confirms such anecdotes; age at first intercourse (AFI) is indeed declining and has so for some time (Bozon, 2003; Finer, 2007). Early age at first intercourse is associated with a plethora of negative downstream consequences, including lower education attainment (Harden, 2012; Spriggs & Halpern, 2008; Wellings et al., 2001), failure to meet education and career goals (Halpern, Joyner, Udry, & Suchindran, 2000), increased risk of teenage pregnancy (Leitenberg & Saltzman, 2000; Wellings et al., 2001), and increased rates of sexually transmitted infections (Kaestle, Halpern, Miller, & Ford, 2005). Moreover, beyond the obvious benefit of avoiding the aforementioned negative outcomes, delaying AFI is associated with greater relationship satisfaction, perception of increased attractiveness, and higher household income (Harden, 2012). Because the aforementioned consequences are severe and long-reaching, psychology has begun to explore potential causal mechanisms of early AFI. Indeed, the field has found a consistent correlate in the literature – intelligence.

in fact, it looks like shows such as this one reduce rates of teen pregnancy...(Kearney & Levine, 2014)

Higher levels of intelligence are positively associated with delaying first intercourse (Woodward, Fergusson, & Horwood, 2001; Paul, Fitzjohn, Herbison, & Dickson, 2000; Halpern et al., 2000; Mott, 1983). Specifically, it seems that intelligent individuals delay intercourse to "safeguard" their futures (Kirby, 2002b; Manlove, 1998; Raffaelli & Crockett, 2003). They perceive the risks associated with early intercourse, (e.g., pregnancy, STIs) to have career–shattering outcomes (Halpern et al., 2000; **?**). Although this correlate holds promise – much of the field has treated this finding as causal and non-spurious. Yet, there is a fundamental confound in the existing literature that makes it impossible to infer causality.

Practically, all of the AFI-intelligence literature has used between family analyses. In all such analyses, gene and environmental influences, such as education and maternal intelligence are hopelessly confounded (See Harden, 2014). By ignoring such confounds, results are uninterpretable and risk misattributions of causality (Rowe & Rodgers, 1997;

Rutter, 2007). Indeed, both intelligence and AFI are highly heritable and have sizable shared environmental variances (**?**Harden, 2014; Plomin & Spinath, 2004). Thus, we need to critically evaluate whether intelligence is a cause of AFI or merely a theoretically attractive confound.

Cause or Confound?

There are numerous theories on the motivations behind adolescents initiation of first intercourse (See Rodgers, 1996 or Buhi & Goodson, 2007 for a review), and even more specific antecedents (Buhi & Goodson, 2007; Kirby, 2002a; B. C. Miller et al., 1997; Santelli & Beilenson, 1992). Many of these theories either emphasize biology/genetics, where typical adolescent development through puberty (and various hormone changes) drives the interest in sexual behavior (W. B. Miller et al., 1999; Udry, 1979), or social/environmental processes, such as Social Learning (DiBlasio & Benda, 1990; Hogben & Byrne, 1998), where social norms alter the likelihood of early sexual behavior; or Social Control theory (Hirschi, 2002), where societal convention reduces the likelihood that individuals will act on their naturally deviant behavior. Under these environmentally centric theories the underlying biology is either ignored or actively resisted (in the case of Social Control theory, while under many of the genetic centric theories, the environmental components are ignored.

Recently, there have been numerous articles advocating integrative models (See Harden, Mendle, Hill, Turkheimer, & Emery, 2008 and Harden, 2014). The integrative Biopsychosocial Model acknowledges both genetic and environmental contributions to human behavior (Engel, 1977; Petersen, 1987; Rodgers, Rowe, & Buster, 1999). Indeed, biology, psychology, and society jointly influence adolescents' decisions to engage in sexual intercourse (Meschke, Zweig, Barber, & Eccles, 2000; Zimmer-Gembeck & Helfand, 2008). Even though this paper focuses on a single predictor – intelligence, we are doing so within the broader context.

Intelligence as the Cause

We've previously mentioned that the short-term risks of early AFI are overwhelming negative, whereas the rewards for delay are equally positive. These consequences extend into adulthood – early AFI is associated with adult delinquency (Harden et al., 2008), anti-social behavior, and substance abuse (Boislard & Poulin, 2011), while those who delayed had higher household incomes in adulthood (Harden, 2012). It is intuitively appealing to believe that intelligent individuals are more likely to observe this high risk, low reward trade off, and act upon such observations by delaying intercourse. Accordingly, intelligent individuals perceive the consequences of early AFI to have career–shattering outcomes (Halpern et al., 2000; Harden & Mendle, 2011).

Indeed, the literature is consistent with this theory. Those with higher educational goals delay their first intercourse (Boislard & Poulin, 2011; Schvaneveldt, Miller, Berry, & Lee, 2001), while those who had previously reported higher goals, but engaged in early sexual intercourse reduced their goals (Schvaneveldt et al., 2001). Beyond academic goals, those with a greater affinity for risk and those who perceive benefits from teen-pregnancy are more likely to engage in risky sexual activities (Raffaelli & Crockett, 2003). A greater understanding of the risks associated with sexual intercourse, such as HIV transmission, is also associated with delayed AFI (C. Mathews et al., 2009).

Smarter adolescents are more likely to report delayed intercourse (Halpern et al., 2000; Mott, 1983; Paul et al., 2000; Woodward et al., 2001). Beyond intercourse, smarter individuals appear to postpone all sexual/romantic activity (Halpern et al., 2000). Such blanket delays may be a proactive attempt to avoid first intercourse precursors. Thus, many researchers have concluded that "[h]igher intelligence operates as a protective factor against early sexual activity during adolescence, and lower intelligence, to a point, is a risk factor." (Halpern et al., 2000)[pg., 213].

However, Halpern et al. (2000) and many of the other studies we have referenced above(e.g., C. Mathews et al., 2009; B. C. Miller et al., 1997; Paul et al., 2000) have used between family designs, typically cross-sectional in nature. Such designs cannot distinguish between processes that act to create differences between families and

processes that create differences among family members(Lahey & D'Onofrio, 2010). Thus the previous studies do no provide conclusive evidence that intelligence is the causal influence behind the AFI-intelligence relationship.

Intelligence as a Confound

A equally valid family of explanations exist in which intelligence is not the driver of the AFI-intelligence relationship, merely a theoretically attractive confound. Instead, various confounds including family level selection effects, or third variables at the individual level could be causing the relationship. Indeed many such findings that link intelligence with various outcomes are the product of misattributing between family confounds to individual level causes.

The relationship between birth order and intelligence is a classic example of this misattribution (See Rodgers, Cleveland, van den Oord, & Rowe, 2000, Rodgers, 2014, or Damian & Roberts, 2015). Between family studies have consistently found that first born children have higher IQs than later born children (Belmont & Marolla, 1973; Zajonc, 1976), and that first borns are higher achievers (Clark & Rice, 1982; Galton, 1875). Yet within family studies have just as consistently found zero relationship (Berbaum & Moreland, 1980; Retherford & Sewell, 1991; Rodgers et al., 2000). Moreover, when within and between analyses are simultaneously conducted, the methodological source of the IQ-birth order effect are clearly revealed – between family differences in family size (Black, Devereux, & Salvanes, 2011; Rodgers et al., 2000; Wichman, Rodgers, & Maccallum, 2006, 2007). Potential causes of this confound include parental IQ (Rodgers et al., 2000) and SES[1](Page & Grandon, 1979).(See Anastasi, 1956, for an insightful overview, written prior to the IQ-birth order debate).[2].

Between family influences such as SES and maternal intelligence could drive the

[1]Selection effects based on SES should not to be confused with the confluence/resource dilution model (Blake, 1981; Zajonc & Bargh, 1980)

[2]"Parenthetically, it may be added that studies on the relation of birth order to intellectual and other psychological characteristics have frequently yielded ambiguous and inconsistent results because of the failure to take family size into account."(Anastasi, 1956, pg 201)

relationship. Socioeconomic status is associated with the onset of first intercourse (Lammers, Ireland, Resnick, & Blum, 2000) and correlated with intelligence (Murray, 1998; Neisser et al., 1996; Strenze, 2007). Parental intelligence and parental education are also linked with child intelligence (Bouchard, Jr., 2004; Devlin, Daniels, & Roeder, 1997; Mercy & Steelman, 1982), and pose very viable alternative explanations in which parents are the one dissuading their children from engaging in early intercourse. For example, daughters whose mothers communicated frequently about the risk associated with sexual intercourse were less likely to have unprotected sex and engaged in sex less frequently (Hutchinson, Jemmott, Jemmott, Braverman, & Fong, 2003). Thus it could be that intelligent mothers, not intelligent children, are the ones recognizing the consequences of early intercourse and acting accordingly.in order to truly understand the causal nature of intelligence on Age at First Intercourse, we need to be able to untangle between and within family processes.

Prior Within Family Analyses

The authors are aware of two studies which explicitly untangled between and within family influences on the AFI-intelligence relationship (Harden & Mendle, 2011; Meredith, 2013).[3]

Current Study

To summarize, the current study examines the relationship between intelligence and age at first intercourse, using siblings and their children from a multigenerational nationally representative sample. This examination extends the intelligence literature in several key ways. We

1. tested whether the relationship between intelligence and age at first intercourse was consistent using between and within family analyses;

[3]Technically, three studies – Nedelec, Schwartz, Connolly, & Beaver (2012) conducted an extensive exploratory analysis of MZ twin pairs from the Add Health Study. They used intelligence difference scores to predict various social outcomes. They generally found null results in their small samples (N ranged from 48 to 166 pairs). Their sample is an underpowered subset of the same sample that **?** used.

2. evaluated the alternative explanation that maternal intelligence influences child AFI; and
3. replicated this findings using assessments of intelligence from other ages.

We made the following predictions, based primary upon Harden (2012) and Harden & Mendle (2011):

Between Families,

1. Does Gen2 intelligence predict Gen2 AFI?: We expect intelligence to be associated with age of first intercourse because there is a sizable body of literature reporting that result (Kirby, 2002b; Manlove, 1998; Raffaelli & Crockett, 2003).

2. Does Gen1 intelligence predict Gen2 AFI?: We also expect maternal intelligence to be associated with age of first intercourse because the heritability of intelligence is quite high (Bouchard, Jr., 2004; Devlin et al., 1997). If intelligence does causally influence AFI we would expect that the cross-generational association between AFI and intelligence would be considerably weaker, but existant. However, if the intelligence-AFI relaitonship is the product of between family confounds, then we would expect that the cross-generational associoation between AFI and intelligence would be stronger than the within generation association because maternal intelligence would be more closely linked with household SES and various parental causes. Comparably sized effects would also be consistent with a between family confound. Given that Harden (2012) and Harden & Mendle (2011) found no within family effect for intelligence, we expect that maternal intelligence will have a comparable or larger effect on between family AFI than child intelligence.

Within Families,

4. Does Gen2 intelligence predict Gen2 AFI?:

No, we do not expect to find all within family differences in intelligence and AFI, given that neither Harden (2012), not Harden & Mendle (2011) reported an effect.

5. Does Gen1 intelligence predict Gen2 AFI?:

Unknown: it is possible that maternal intelligence will have an effect, as such a link would explain the between family effects as well as many of the alternative household-level influences.

Is the relationship consistent across methods?:

Doubtful BECAUSE REASONS.

Methods

Design

We adapted Kenny and colleagues (2001; 2006) reciprocal standard dyad model to facilitate sibling comparisons. Sibling-based quasi-experimental models are particularly effective at incorporating genetic and environmental design elements (Lahey & D'Onofrio, 2010; Rutter, 2007).

$$\mathrm{Y_{i\Delta}} = \beta_0 + \beta_1 \bar{\mathrm{Y}}_\mathrm{i} + \beta_2 \bar{\mathrm{X}}_\mathrm{i} + \beta_3 \mathrm{X_{i\Delta}}$$

where

$$Y_{i1} = \max(Y_{ij});\ Y_{i2} = \min(Y_{ij});\ Y_{i\Delta} = Y_{i1} - Y_{i2};\ X_{i\Delta} = X_{i1} - X_{i2}$$

In this model, the relative difference in kin outcomes ($\mathrm{Y_\Delta}$) is predicted from the mean level of the outcome ($\mathrm{Y_{mean}}$), the mean level of the predictor ($\mathrm{X_{mean}}$), and the between-kin predictor difference ($\mathrm{X_\Delta}$). The mean levels support causal inference through at least partial control for genes and shared environment. Therefore, we simultaneously evaluate the individual difference ($\mathrm{X_\Delta}$) and the joint contribution of genes and shared environment ($\mathrm{Y_{mean}}$ & $\mathrm{X_{mean}}$).

More broadly, this model allows us to explicitly untangle between and within family influences. If there is a true causal effect between the individual difference and the outcome (in our case – intelligence and AFI respectively), then we would expect kin differences in intelligence to be significantly associated with kin differences in the outcome. If the effect is a spurious effect – ie the function of between family differences – then we would expect to find no significant relationship between the differences in the outcome with the differences in the predictor.

Sample

The National Longitudinal Survey of Youth 1979 (NLSY79) is a nationally representative household probability sample, jointly sponsored by the US Department of Labor and US Department of Defense. In 1980, 12,686 adolescents were surveyed from 8,770 households on a battery of measures. The initial survey consisted of three subsamples:

- a cross–sectional probability sample of 6,111 non-institutionalized adolescents residing in the United States on December 31^{st} of 1978,
- an over–sampled civilian subsample of 5,295 racial minorities and disadvantaged whites, and
- a representative sample of 1,280 youths serving in the US Military on September 30^{th}, 1978.

In the civilian samples, subjects' birthdates ranged from January 1, 1957 to December 31, 1964, and were between the ages of 14 and 21 on 31^{st} of 1978, whereas military subject's birthdates ranged from January 1, 1957 to December 31, 1961, and were likewise between 17 and 21 years old. Participants were surveyed annually until 1994, and then surveyed biannually to the present day. Two waves of planned attrition occurred. After the 1984 interview, all but 201 randomly selected members of the military sample were dropped. After the 1990 interview, all 1,643 disadvantaged whites from the oversample were dropped. More information, such as the data and documentation can be found on the Bureau of Labor Statistics (BLS) website: `http://www.bls.gov/nls/nlsy79.htm`.

In 1986, the biological children of the female NLSY79 participants were surveyed, resulting in the NLSY79 Children and Young Adults (NLSY79–CYA) survey. These 11,512 participants are also surveyed on a biannual basis. Accordingly, participants in the NLSY79 will be periodically referred to as the Generation 1 (Gen1) sample, whereas the NLSY79–CYA will be referred to as the Generation 2 (Gen2) sample.

Tetrads

Mother-Child-Aunt-Nibling (MCAN) tetrads were created using the NLSY Kinship Links (Rodgers et al., 2015) and supporting R package (Beasley et al., 2015). The oldest two female kin (Mother, Aunt) were selected from each household; additional female Gen1 kin were excluded. Three tetrad designs were employed, in which the genders of Gen2 were the defining feature:

- Mother-Daughter-Aunt-Niece (MDAN) included the oldest female child from each of the sisters,
- Mother-Son-Aunt-Nephew (MSAN) included the oldest male child from each of the sisters, and
- Mother-Child-Aunt-Nibling (MCAN) included the first born child from each of the sisters.

All outcomes were standardized by gender prior to tetrad creation. Table 3 on page 26 reports descriptive statistics for all variables used throughout this paper by whether the respondent has a sibling in the sample.

Age at First Intercourse

Gen1. NLSY-79 subjects were surveyed about their AFI over a maximum of three time-points (1983, 1984, 1985). In theory, subjects were only to be asked in later waves, if subjects had not reported an AFI in the 1983 wave. However, in practice, many subjects were surveyed twice. Female participants were asked additional information (Year of First Intercourse, Month of First Intercourse) in waves 1984 and 1985. Because subjects were surveyed repeatedly, we used this opportunity to estimate the reliability of self-reported AFI as well as the reliability of the AFI difference scores. In Table 6 on page 29, the lower triangle reports the correlations of self-reported AFI across 1983-1985; the diagonal indicates the number of respondents reporting AFI for that year, and upper triangle reveal the number of respondents who reported AFI for both respect years. Stars indicate significant at the .01 level. The test-retest

correlations are high (r > .75) across all viable pairings, suggesting that our subjects are consistently reporting their AFIs.

Gen2. Over the life-time of the NLSY-CYA survey, participants were surveyed about their AFI. The exact phrasing of the question varied by administration. Between 1988 and 2000, subjects were asked for age, year, and month of first intercourse. After 2000, subjects were only asked their age. The reason for this change is unknown. However, the first author suspects that the change had to do with the fact that the modal response for month was consistently: "Don't Know". Indeed, only 1147 subjects reported a viable month of first intercourse.

Regardless, we calculated AFI as follows, using SAS University Edition(citation). First, we transformed year of first intercourse into age. If subjects reported both age and year within the same survey, we averaged the age scores. Across all survey years, we identified the minimum AFI and maximum AFI for each subject. Then we took the average of those two scores. Given that the expected AFI of a subject $\neq$ the reported AFI, we added 1 to the Maximum AFI. Therefore, if the subject only reported one instance of AFI, their AFI would now reflect their expected AFI. For example, a subject who reports AFI at 16 could be 16 years and 1 day old OR 16 years and 364 days old. Thus the expected value for 16 is in fact 16.5. We calculated AFI in this manner because we wished to include the maximum amount of information without ignoring the expected value problem with self-reported age. Using this method, the average Gen2 AFI was 16.01 years (sd = 2.30; n = 6288).

Taking the average of all AFIs (without addressing expected value), results in 15.49 (sd equal 2.30; n equal 6288). Adding in expected value of .5 changes this value to 15.99. Pretty much indistinguishable.)

After transforming all AFI scores, we recoded all impossible AFIs as missing. We considered a score to be impossible if the reported AFI that exceeded participant's age at time of survey ($\overline{\text{AFI}}$ = 15.99, sd = 2.30, n = 6235). Next, we excluded all AFIs below age 12 (16.14, sd = 2.10, n = 6087). Finally we excluded subjects who reported AFI prior to menstruation (16.16, sd = 2.09, n = 6047). We excluded those below age 12 because those responses likely are the result of misunderstanding or non-consensual sexual activity, while we excluded those with premenstrual AFI because of we were only interested in post-pubescent sexual activity. AFI varied by gender and race. Most

notably, women reported AFIs that were 6 months later than men, and black men reported the lowest AFI (15 yrs). See Table 3 on page 26 for a complete breakdown.

Measures

Generation 1

The Armed Services Vocational Aptitude Battery (ASVAB; Form 8A; Palmer, Hartke, Ree, Welsh, & Valentine, Jr., 1988) was administered to Gen1 participants during the summer and fall of 1980 (U.S. Department of Defense, 1982), and was used to establish national norms for the Department of Defense (Waters, Laurence, Camara, & Green, 1987). The Armed Forces Qualification Test (AFQT) is derived from the ASVAB, and used as a measure of general trainability (Maier & Sims, 1986). It is a composite of four subscales: Arithmetic Reasoning (AR; 30 items), Math Knowledge (MK; 25 items), Paragraph Comprehension (PC; 15 items), and Word Knowledge (WK; 35 items). Arithmetic Reasoning targets the ability to solve word problems. Math Knowledge also tests quantitative ability, by assessing knowledge of high school level mathematics, with special emphasis on algebra, fractions, and geometry. The remaining subscales focus on verbal ability, and are sometimes referred to as the Verbal Composite (VE). Specifically, Word Knowledge tests the subjects' knowledge of the meaning of words within a given context, whereas Paragraph Comprehension targets a subject's ability to understand the meanings of paragraphs. Other administrations of the pencil and paper ASVAB reveal that all the AFQT subscales have high internal consistency ($\alpha_{AR} = .91$; $\alpha_{WK} = .92$; $\alpha_{PC} = .81$; $\alpha_{MK} = .87$; Kass, Mitchell, Grafton, & Wing, 1982). Reported reliability of the AFQT (8A) ranges from .87 to .93 (Palmer et al., 1988).

Methods of calculating the AFQT have varied throughout the ASVAB's administrative lifetime (Mayberry & Hiatt, 1992). For pencil and paper administrations, standard scores were created for each of the subscale scores ($\bar{x} = 50$, sd = 10), and then combined into a standard score. Then, the AFQT standard score is

derived from the following formula:

$$\text{AFQT} = \text{AR} + \text{MK} + 2\text{VE}, \tag{1}$$

$$\text{where VE} = \text{PC} + \text{WK}. \tag{2}$$

This score is then converted into a percentile, which determines an applicant's basic qualification for enlistment. All applicants must earn a score at or above the 10^{th}percentile(Defense Manpower Data Center, 2012). Each branch has its own minimum score, ranging from 31 to 36 (U.S. Department of the Army, 2013; U.S. Coast Guard, 2004), and each branch uses different linear combinations of these subtests to determine an applicant's eligibility for specialty positions. Additionally, multiple researchers have used the AFQT standard score as a proxy for general intelligence (g) (Herrnstein & Murray, 1994; Der, Batty, & Deary, 2009). Indeed, the military has found that the AFQT correlated 0.8 with the Wechsler Adult Intelligence Scale (WAIS; McGrevy, Knouse, & Thompson, 1974). Moreover, the AFQT consistently predicts outcomes traditionally associated with intelligence(Welsh, Kucinkas, & Curran, 1990), including grades (Wilbourn, Valentine, Jr., & Ree, 1984; J. J. Mathews, 1977).

Generation 2

Administration of ability measures has varied considerably across the lifecourse of the NLSY-CYA survey (See Table 2.12 from Center for Human Resources Research, 2009 for a summary). However, the vast majority of subjects have completed the following test batteries:

- Peabody Individual Achievement Test (PIAT; Dunn & Markwardt, 1970):
  - Math Subtest (84 items),
  - Reading Recognition Subtest (84 items),
  - Reading Comprehension Subtest (84 items),
- The Peabody Picture Vocabulary Test-Revised (PPVT-R; Form L; Dunn & Dunn, 1981; 175 items), and

- Wechsler Intelligence Scales for Children–Revised (WISC-R; Wechsler, 1974) Digit Span Subscale (28 items).

Although individual item level data was available for all of the aforementioned tests, conducting a unidimensional 2-PL is not a viable means of estimating general ability because of the nature of test construction and administration. The PIATs and PPVT-R were administered to subjects in an adaptive manner. The starting items on the PIAT Math and PVVT-R were determined by age, whereas the starting items for the remaining PIAT subtests were determined based on PIAT Math performance. Moreover, administration of a given test were terminated when a subject reached a "ceiling." For example, testing was terminated for the PIAT Math if a subject incorrectly answered 5 of the most recent 7 questions (See Baker, Keck, Mott, & Quinlan, 1993 for a thorough overview of NLSY-CYA test administration protocols). In essence, this administration procedure results in a tremendous amount of non-randomly missing data.

Although the administration created non-randomly missing data, the standard scores of the PVVT-R, PIATs, and WISC-R Digit Span themselves are valid and very reliable assessments of cognitive ability (Mott & Baker, 1995). Accordingly, we elected to use the standard scores of all the Gen2 ability measures already mentioned. However, subjects were surveyed on a biannual basis. Thus we could not use cognitive tests at a given age. Instead, we aggregated scores across a 4 year window, and targeted ages 9 and 10. We targeted 9.5 because all cognitive tests were administered within the 8–11 age window, we wanted to maximize the number of subjects with viable ability scores, and we wanted to ensure temporal precedence with respect to AFI. In the case of missing subtests, we allowed age 11 scores to replace age 9 scores, and age 8 scores to replace age 10 scores. By employing a 4 year window, all subjects had an equal chance of replacing the primary test administration. Our replacement strategy ensured that the average age of testing matched the average of our targeted ages.

Measurement. A unidimensional confirmatory factor analytic model was run in M*plus* 7.31 (Muthén & Muthén, 2014), and used a robust maximum likelihood estimator (MLR). There were 8,254 useable observations in 3,742 clusters. A single

factor solution fit the model decently (RMSEA = .101, p(RMESA < .05) = 0; CFI = .973; TLI = .946, SRMR = .027). Table 4 on page 27 contains a full summary of the model fit statistics, and Table 5 on page 28 contains the factor loadings.

Replicability & Reliability. Given the recent concerns about replicatability in psychology (Open Science Collaboration, 2015), we repeated our aggregates of Gen2 intelligence, centered at ages 10.5 and 11.5, and replicated all of our analyses. These replications can be found in the Appendices A and B, respectively. Appendix A begins on page 41 and appendix B begins on page 49. The test-retest reliabilities of Gen2 intelligence across our three aggregations is reported in the lower triangle of Table 8 on page 30. The diagonal indicates the number of respondents with intelligence aggregations for that year, and upper triangle reveal the number of respondents with viable scores for both respective ages. Stars indicate significant at the .01 level. The test-retest correlations are very high (r > .90) across all pairings, suggesting that our method captures consistent (but not identical) measures of intelligence across ages. Additional analyses examining the reliability of intelligence difference scores are reported in a later section.

Reliability of Difference Scores

Our design assumes that the difference scores of our measures are reliable. INSERT MORE ON THIS. We've reported the test-retest reliability of Gen2 intelligence and Gen1 AFI in earlier sections. Here, we report the test-retest reliability of the differences of those measures across kin.

Estimated Reliabilities. Sibling differences in AFI as reported in 1983, 1984, and 1985 were strongly correlated with each other (See Table 7 on page 29). Comparing sibling differences in AFI as reported in 1983 and 1984 (n = 783 pairs) we found a strong correlation (r = .76). The sample of sibling pairs with complete information in 1985 was too small (n = 12) to compare to the other two years. Regardless, sibling differences in self-reported AFI appear reliable. Although we could not calculate test-retest reliabilities for Gen2, we have no reason to believe that those differences

would fundamentally differ from Gen1's reports.

Cousin differences in intelligence as assessed at ages 9.5, 10.5, and 11.5 were correlated using three different linking methods (Mixed, Daughers, Sons). Table 9 on page 31 reports the correlated differences of the first borns of each sister, Table 10 on page 32 reports the correlated differences of the first born girls, and Table 11 on page 33 reports the correlated differences of the first born sons. Reliabilities across linking methods was consistent and high (min r = .86; max r = .95). However, again, we were unable to calculate the test-retest difference score reliablities for Gen1.

Calculated Reliabilities. Nonetheless, we were able to confirm that difference scores for both generations were reliable for the measures we could estimate. For the remainder, we have calculated the reliability of the difference scores using the following equation (Lord, 1963):

$$\rho_{dd'} = \frac{\sigma_x^2 \rho_{xx'} + \sigma_y^2 \rho_{yy'} - 2\rho_{xy}\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 - 2\rho_{xy}\sigma_x\sigma_y}$$

where,

- $\rho_{dd'}$ is the reliability of the difference score,
- σ_x^2 is the variance of kin_1's score,
- $\rho_{xx'}$ is the reliability of kin_1's score,
- σ_y^2 is the variance of kin_2's score,
- $\rho_{yy'}$ is the reliability of kin_2's score,
- ρ_{xy} is the correlation between kin_1's and kin_2's scores.

Accordingly, we can substitute the following values into the above equation to calcualte the difference score reliability for Gen1 intelligence. where,

- σ_x^2 and σ_y^2 are 1.1881 (from Table 1 on page 25)
- $\rho_{xx'}$ and $\rho_{yy'}$ are .87 (the low end of AFQT reliability reported in Palmer et al., 1988),
- ρ_{xy} is the correlation between sisters is .67.

```
sigmax = sqrt(1.1881)

sigmay = sigmax

rxx = .87

ryy = rxx

rxy=.67

(sigmax*sigmax*rxx +sigmay*sigmay*ryy -2*rxy*sigmax*sigmay) /

  (sigmax*sigmax + sigmay*sigmay -2*rxy*sigmax*sigmay)


## [1] 0.606
```

Accordingly, we can substitute the following values into the above equation to calcualte the difference score reliability for Gen2 AFI. where,

- σ_x^2 and σ_y^2 are 4.41 (from Table 2 on page 25)
- $\rho_{xx'}$ and $\rho_{yy'}$ are .76 (from Table 6 on page 29),
- ρ_{xy} is the AFI correlation between first born cousins is .099.

```
sigmax = sqrt(4.41)

sigmay = sigmax

rxx = .76

ryy = rxx

rxy=.099

(sigmax*sigmax*rxx +sigmay*sigmay*ryy -2*rxy*sigmax*sigmay) /

  (sigmax*sigmax + sigmay*sigmay -2*rxy*sigmax*sigmay)


## [1] 0.734
```

The calculated reliability of Generation 1's differences in AFQT was .606, acceptable, but lower than the empirical correlation we derived for cousin differences. Gen2 AFI difference scores were also reliable (r = .73) and comparable to Gen1 sibling differences.

Results

We examined the relationship between AFI and intelligence using two designs: between and within families. The between family analyses report the relationships between the average AFI and various measures of ability. The within family analyses

attempt to replicate the between family findings by testing whether differences in AFI can be explained by differences in various measures of ability. If there is a causal relationship between intelligence and AFI then differences in AFI will be significantly associated with differences in ability. If the relationship is the result of between family confounds, such as shared environmental influences, then differences in AFI will not be significantly associated with differences in ability, and accordingly, AFI cannot be caused by intelligence.

Between Family Anlayses

First, we examined the between family results. We tested whether the family average of Gen2 AFI could be predicted by the family averages of Gen1 ability and of Gen2 ability. We evaluated the influences both independently and simultaneously. All ability scores have been standardized by generation ($\overline{g} = 0$, sd $= 1$), prior to averaging by household. AFI scores have been standardized by gender ($\overline{\text{AFI}} = 0$, sd $= 1$), prior to averaging by household.

Mean Gen1 Intelligence → Mean Gen2 AFI. Gen1 sister averages of standardized AFQT scores were used to predict Gen2 averages of gender standardized AFI. Table 12 on page 34 displays the results by Gen2 linking. The Mixed model reports the averages of the first borns of each sister (n = 342), the Daughters model reports the averages of the first born girls (n = 264), and the Sons model reports the averages of the first born sons (n = 282). All three models reveal similar results. A one unit increase in the average standardized intelligence of the children's mothers predicted $\approx .013$ increase in average Gen2 AFI. The adjusted R^2 varied slightly by Gen2 linking (Mixed = .087, Daughters = .097, Sons = .103).

Mean Gen2 Intelligence → Mean Gen2 AFI. Gen2 cousin averages of standardized ability scores were used to predict Gen2 averages of gender standardized AFI. Table 13 on page 35 displays the results by Gen2 linking. The Mixed model reports the averages of the first borns of each sister (n = 344), the Daughters model reports the averages of the first born girls (n = 267), and the Sons model reports the

averages of the first born sons (n = 283). All three models reveal similar results. A one unit increase in the average standardized intelligence of the children predicted ≈ .075 increase in average Gen2 AFI. The adjusted R^2 varied slightly by Gen2 linking (Mixed = .014, Daughters = .016, Sons = .009).

Mean Joint Intelligence → Mean Gen2 AFI. Results from the Gen1 sister averages of standardized AFQT scores and Gen2 cousin averages of standardized ability scores predictions of Gen2 averages of gender standardized AFI are displayed in Table 14 on page 36. Again, three models based on Gen2 linking are displayed: Mixed (n = 337), Daughters(n = 260), and the Sons(n = 278). All three models reveal similar results. Gen1 intelligence was significantly associated with Gen2 AFI ($p < .01$), while Gen2 intelligence was not significantly associated with Gen2 AFI. A one unit increase in the average standardized intelligence of the children's mothers predicted ≈ .014 increase in average Gen2 AFI. The adjusted R^2 varied slightly by Gen2 linking (Mixed = .086, Daughters = .097, Sons = .100), but each were practically identical to the Mean Gen1 Intelligence models.

Within Family Analyses

We attempted to replicate the between family analyses reported in the previous subsection, using within family analyses. Using the discordant sibling model, we predicted the differences in Gen2 AFI as a function of differences in intelligence, controlling for means of the outcomes and predictors. We ran three series of models, where we examined the individual and then joint influence of Gen1 intelligence and Gen2 intelligence. Moreover, within each series we included three Gen2 linking variants, just as we did in the between family analyses: Mixed model reports the differences of the first borns of each sister, the Daughters model reports the differences of the first born girls, and the Sons model reports the differences of the first born sons.

Dif Gen1 Intelligence → Dif Gen2 AFI. Gen1 sister differences in standardized AFQT scores were used to predict Gen2 differences of gender standardized AFI, controlling for Gen1 sister averages of standardized AFQT scores and Gen2

averages of gender standardized AFI. Table 15 on page 37 displays the results by Gen2 linking. The Mixed model reports the averages and differences of the first borns of each sister (n = 336), the Daughters model reports the averages and differences of the first born girls (n = 258), and the Sons model reports the averages and differences of the first born sons (n = 278). All three models reveal similar results. Gen2 averages of gender standardized AFI were significant predictors of Gen2 differences in gender standardized AFI ($p < .01$), across all three linking methods. A one unit increase in the average gender standardized AFI predicted ≈ 0.34 increase in average Gen2 AFI difference, controlling for all over variables in the model.

In the Sons model, the Gen1 sister average of standardized AFQT scores was a significant predictor of differences in Gen2 AFI ($p < .01$). A one unit increase in the average standardized intelligence of the children's mothers predicted $\approx .0083$ decrease in the AFI difference between siblings. All other variables were not significant, including all kin difference variables. The adjusted R^2 varied slightly by Gen2 linking (Mixed = .066, Daughters = .072, Sons = .106).

Dif Gen2 Intelligence → Dif Gen2 AFI. Gen2 cousin differences in standardized ability scores were used to predict Gen2 differences of gender standardized AFI, controlling for Gen2 cousin averages of standardized ability scores and gender standardized AFI. Table **??** on page ?? displays the results by Gen2 linking. The Mixed model reports the averages and differences of the first borns of each sister (n = 291), the Daughters model reports the averages and differences of the first born girls (n = 223), and the Sons model reports the averages and differences of the first born sons (n = 238). All three models reveal similar results. Gen2 averages of gender standardized AFI were significant predictors of Gen2 differences in gender standardized AFI ($p < .01$), across all three linking methods. A one unit increase in the average gender standardized AFI predicted ≈ 0.38 increase in average Gen2 AFI difference, controlling for all over variables in the model.

In the Sons model, the Gen2 cousin average of standardized ability scores was a significant predictor of differences in Gen2 AFI ($p < .05$). A one unit increase in the

average standardized intelligence of the children predicted ≈ .107 decrease in the AFI difference between siblings. All other variables were not significant, including all kin difference variables. The adjusted R^2 varied slightly by Gen2 linking (Mixed = .103, Daughters = .121, Sons = .132).

Dif Joint Intelligence → Dif Gen2 AFI. Gen1 sister differences in standardized AFQT scores and Gen2 cousin differences in standardized ability scores were used to predict Gen2 differences of gender standardized AFI, controlling for Gen1 sister averages of standardized AFQT scores, Gen2 cousin averages of standardized ability scores, and Gen2 cousin averages of gender standardized AFI. Table 17 on page 39 displays the results by Gen2 linking. The Mixed model reports the averages and differences of the first borns of each sister (n = 285), the Daughters model reports the averages and differences of the first born girls (n = 217), and the Sons model reports the averages and differences of the first born sons (n = 235). All three models reveal similar results. Gen2 averages of gender standardized AFI were significant predictors of Gen2 differences in gender standardized AFI ($p < .01$), across all three linking methods. A one unit increase in the average gender standardized AFI predicted ≈ 0.38 increase in average Gen2 AFI difference, controlling for all over variables in the model.

All other variables were not significant, including all kin difference variables. The adjusted R^2 varied slightly by Gen2 linking (Mixed = .090, Daughters = .105, Sons = .131).

Discussion

This article presents the relationship between AFI and intelligence using two difference designs: between and within families. The between-family design allowed us to replicate previous researchers who used a cross-sectional sample. The within-family design allowed us to evaluate intelligence differences within the family, to address issues of causality. The results revealed a stark contrast between the two methods. Notably, the between family analyses show a relationship between intellgience and AFI, however the relationship is stronger between mother intelligence and child AFI than between

child intelligence and child AFI, which suggests that family-level rathar than individual level intelligence is source. In the within family analyses, the effect vanishes for both maternal intelligence and child intelligence.

Intrepretation

Rodgers et al. (2008) looked at the relationship between IQ and education as they influenced age at first birth in Danish twin data. Their conclusion:

> variance in AFB emerges from [IQ and education] differences between families, not differences between sisters within the same family.

We have exactly the same result in the current study Note that the IQ differences between siblings are relatively small; is this important? Under a purely genetic model, Rodgers Rowe, Intelligence, 1987, estimated the average absolute deviation in IQ among random pairs to be 17.1 IQ points, compared to 12.1 for full siblings. But cousins, who also are in our study, deviate on average by 16.1 IQ points, nearly as much as random pairs We believe that differences within the family simply are not important for defining AFI outcome differences This finding matches the Harden Mendle biometrical analysis of Add Health, where they found that only shared environmental influences mattered – those would manifest in between-family differences, but not in within-family differences Further, we're not convinced that in these between-family analyses, intelligence is the actual cause of AFI differences – if so, we think they would perhaps diffuse a bit, but would still show up in within-family analyses Rather, we think maternal and child IQ are indirect measures of many other household features, any one of which may be more proximal as the causal explanation – income, parental education, family interaction, etc. Or, the whole package of these features may stand in for a general environmental factor, a "little e," which indexes the quality of home environment – a composite of parental income, intelligence, education, family interaction, etc.

Reliability of Difference Scores

Limitations

Tables

Descriptive Statistics

Table 1

Gen1 Summary Statistics by Sibling Status.

		Has Sibling in NLSY?				
		Mean_3			Sd_2	
		0	1		0	1
Mother Age at Birth	23.31	22.595	23.949	5.49	5.16	5.69
Standardized AFQT	0.00	0.108	− 0.088	1.09	1.07	1.10
AFI	17.84	17.796	17.870	2.33	2.27	2.39

Table 2

Gen2 Summary Statistics by Sibling Status.

		Has Sibling in NLSY?				
		Mean_3			Sd_2	
		0	1		0	1
Year Born	1985.73	1986.160	1985.679	26.99	6.62	28.39
Standardized Ability Score	0.00	0.275	− 0.023	1.00	0.89	1.01
AFI	16.16	16.052	16.173	2.10	2.09	2.10

Table 3

Gen2 AFI by Gender, Race, and Gender by Race.

			AFI	
			$Mean_3$	Sd_2
			16.16	2.10
		Male	15.88	2.15
		Female	16.47	1.99
		Hispanic	16.22	2.14
		Black	15.66	2.01
		Non-Black, Non-Hispanic	16.57	2.05
Male	RACE	Hispanic	15.92	2.16
		Black	15.04	1.96
		Non-Black, Non-Hispanic	16.54	2.06
Female		Hispanic	16.60	2.05
		Black	16.26	1.88
		Non-Black, Non-Hispanic	16.61	2.05

Measurement

Table 4

Gen2 Measurement Model.

	g at Age 9.5
Estimator	MLR
Observations	8254
Parameters	15
ChiSqM_Value	423
ChiSqM_DF	5
ChiSqM_PValue	0
ChiSqBaseline_PValue	0
LL	-142896
UnrestrictedLL	-142653
LLCorrectionFactor	1.44
UnrestrictedLLCorrectionFactor	1.37
CFI	0.973
TLI	0.946
AIC	285821
BIC	285927
aBIC	285879
RMSEA_Estimate	0.101
RMSEA_90CI_LB	0.093
RMSEA_90CI_UB	0.109
RMSEA_pLT05	0
SRMR	0.027
AICC	285821

Table 5

Gen2 Factor Loadings.

	Test	Estimate	S.E.	Est./.S.E.	P.Value
1	DIGIT	0.530	0.010	52.400	0
2	MATH	0.720	0.007	98.100	0
3	RECOG	0.871	0.005	184.000	0
4	COMP	0.855	0.006	154.000	0
5	PPVT	0.695	0.010	68.800	0

Difference Scores

Table 6

Correlation of Gen1 Self-Reported AFI Across 1983-1985

	1983	1984	1985
1983	8432	3765	88
1984	0.86*	4516	0
1985	0.76*	NANA	424

Table 7

Correlation of Gen1 Sibling Differences in Self-Reported AFI across 1983-1985 (Standardized by Gender)

	1983	1984	1985
1983	3517	783	0
1984	0.76*	1058	0
1985	NANA	NANA	12

Table 8

Correlation of Gen2 Aggregated Intelligence across Ages 9.5, 10.5, 11.5

	G_G2_9.5	G_G2_10.5	G_G2_11.5
G_G2_9.5	8254	7974	7669
G_G2_10.5	0.95*	8143	7838
G_G2_11.5	0.90*	0.96*	7970

Table 9

Correlation of Gen2 First Born Cousin Differences in Intelligence across Ages 9.5-11.5

	Age 9.5	Age 10.5	Age 11.5
Age 9.5	291	281	270
Age 10.5	0.91*	287	275
Age 11.5	0.87*	0.95*	286

Table 10

Correlation of Gen2 First Female Cousin Differences in Intelligence across Ages 9.5-11.5

	Age 9.5	Age 10.5	Age 11.5
Age 9.5	223	216	211
Age 10.5	0.90*	219	213
Age 11.5	0.86*	0.95*	223

Table 11

Correlation of Gen2 First Male Cousin Differences in Intelligence across Ages 9.5-11.5

	Age 9.5	Age 10.5	Age 11.5
Age 9.5	238	228	218
Age 10.5	0.92*	234	223
Age 11.5	0.87*	0.95*	230

Between Family Analyses

Table 12

Mean Gen1 Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean Gen2 AFI		
	Mixed	Daughers	Sons
Gen1 Intel	0.012^{***}	0.013^{***}	0.014^{***}
	(0.008, 0.017)	(0.009, 0.018)	(0.009, 0.019)
Constant	-0.754^{***}	-0.801^{***}	-0.856^{***}
	(−1.010, −0.500)	(−1.090, −0.513)	(−1.140, −0.569)
Observations	342	264	282
R^2	0.090	0.101	0.106
Adjusted R^2	0.087	0.097	0.103
Residual Std. Error	0.700 (df = 340)	0.703 (df = 262)	0.691 (df = 280)
F Statistic	33.500^{***} (df = 1; 340)	29.400^{***} (df = 1; 262)	33.200^{***} (df = 1; 280)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 13

Mean Gen2 Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean Gen2 AFI		
	Mixed	Daughers	Sons
Gen2 Intel	0.077**	0.086**	0.066*
	(0.014, 0.140)	(0.013, 0.159)	(−0.004, 0.135)
Constant	−0.021	−0.020	−0.039
	(−0.100, 0.057)	(−0.109, 0.070)	(−0.125, 0.047)
Observations	344	267	283
R^2	0.016	0.020	0.012
Adjusted R^2	0.014	0.016	0.009
Residual Std. Error	0.739 (df = 342)	0.744 (df = 265)	0.732 (df = 281)
F Statistic	5.700** (df = 1; 342)	5.320** (df = 1; 265)	3.430* (df = 1; 281)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 14

Mean Joint Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean Gen2 AFI		
	Mixed	Daughers	Sons
Gen1 Intel	0.013***	0.014***	0.015***
	(0.008, 0.017)	(0.008, 0.019)	(0.009, 0.020)
Gen2 Intel	−0.002	−0.0002	−0.015
	(−0.070, 0.066)	(−0.077, 0.077)	(−0.089, 0.060)
Constant	−0.766***	−0.814***	−0.892***
	(−1.050, −0.482)	(−1.130, −0.497)	(−1.220, −0.568)
Observations	337	260	278
R^2	0.091	0.104	0.107
Adjusted R^2	0.086	0.097	0.100
Residual Std. Error	0.704 (df = 334)	0.706 (df = 257)	0.695 (df = 275)
F Statistic	16.700*** (df = 2; 334)	14.900*** (df = 2; 257)	16.500*** (df = 2; 275)

Note: *p<0.1; **p<0.05; ***p<0.01

Within Family Analyses

Table 15

Dif Gen1 Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.303***	0.328***	0.376***
	(0.187, 0.419)	(0.192, 0.465)	(0.251, 0.501)
INTMOMmean	−0.003	−0.002	−0.008***
	(−0.007, 0.002)	(−0.008, 0.004)	(−0.013, −0.003)
INTMOMdiff	0.001	0.001	0.0003
	(−0.003, 0.006)	(−0.005, 0.006)	(−0.005, 0.005)
Constant	1.180***	1.180***	1.500***
	(0.893, 1.470)	(0.838, 1.520)	(1.180, 1.830)
Observations	336	258	278
R^2	0.074	0.083	0.116
Adjusted R^2	0.066	0.072	0.106
Residual Std. Error	0.757 (df = 332)	0.780 (df = 254)	0.734 (df = 274)
F Statistic	8.830*** (df = 3; 332)	7.640*** (df = 3; 254)	12.000*** (df = 3; 274)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 16

Dif Gen2 Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.357***	0.395***	0.398***
	(0.239, 0.475)	(0.259, 0.531)	(0.269, 0.527)
Int_Mean	−0.064	−0.069	−0.107**
	(−0.142, 0.014)	(−0.160, 0.021)	(−0.193, −0.021)
Intdiff	0.029	0.035	0.009
	(−0.018, 0.075)	(−0.019, 0.089)	(−0.041, 0.059)
Constant	1.050***	1.070***	1.070***
	(0.962, 1.140)	(0.970, 1.180)	(0.970, 1.160)
Observations	291	223	238
R^2	0.112	0.133	0.143
Adjusted R^2	0.103	0.121	0.132
Residual Std. Error	0.769 (df = 287)	0.787 (df = 219)	0.753 (df = 234)
F Statistic	12.100*** (df = 3; 287)	11.200*** (df = 3; 219)	13.000*** (df = 3; 234)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 17

Dif Joint Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.345***	0.379***	0.428***
	(0.221, 0.469)	(0.234, 0.524)	(0.293, 0.562)
Int_Mean	−0.064	−0.071	−0.068
	(−0.153, 0.024)	(−0.174, 0.032)	(−0.164, 0.028)
INTMOMmean	0.001	0.001	−0.005
	(−0.005, 0.007)	(−0.006, 0.008)	(−0.012, 0.001)
Intdiff	0.025	0.034	0.007
	(−0.023, 0.073)	(−0.022, 0.090)	(−0.044, 0.058)
INTMOMdiff	0.001	−0.001	0.001
	(−0.004, 0.006)	(−0.007, 0.005)	(−0.005, 0.006)
Constant	0.999***	1.010***	1.370***
	(0.642, 1.360)	(0.591, 1.430)	(0.975, 1.770)
Observations	285	217	235
R^2	0.106	0.125	0.150
Adjusted R^2	0.090	0.105	0.131
Residual Std. Error	0.774 (df = 279)	0.795 (df = 211)	0.754 (df = 229)
F Statistic	6.610*** (df = 5; 279)	6.040*** (df = 5; 211)	8.080*** (df = 5; 229)

Note: *p<0.1; **p<0.05; ***p<0.01

Figures

Figure 1. Comparable Between Family and Within Family Effects

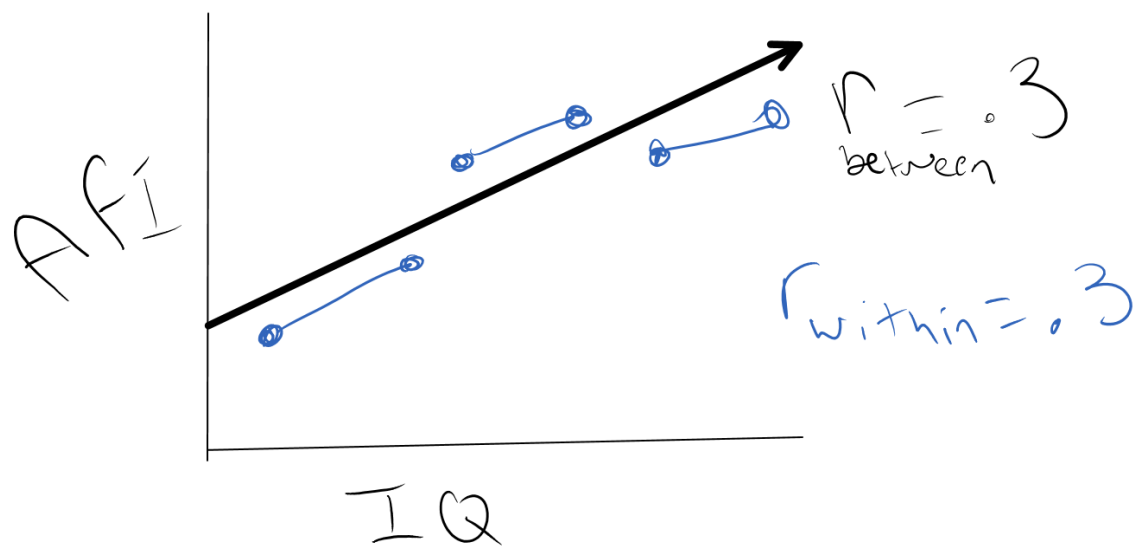
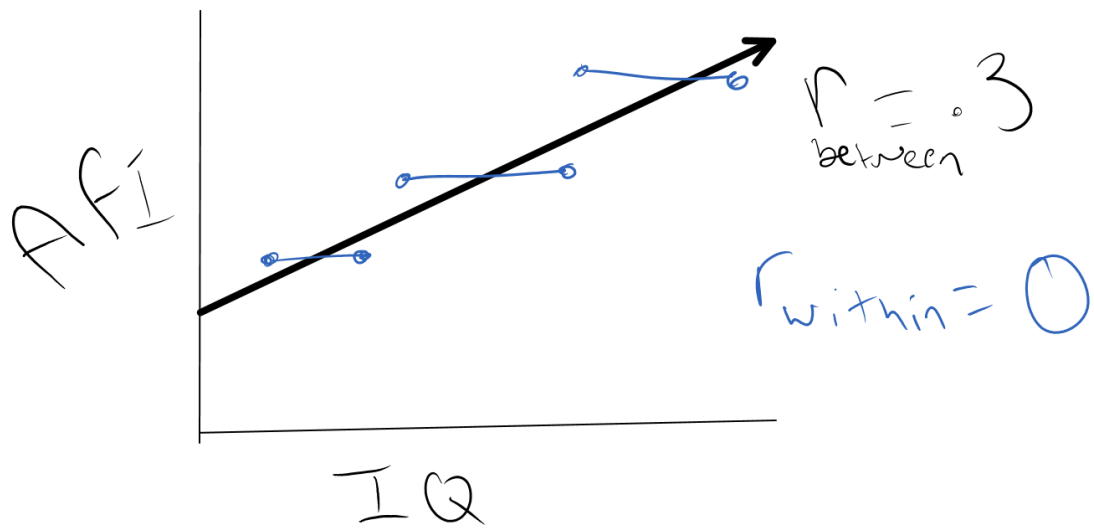


Figure 2. Unequal Between Family and Within Family Effects



Appendix A

Age 10.5 Replication

Table A1

Gen2 Measurement Model.

	g at Age 10.5
Estimator	MLR
Observations	8143
Parameters	15
ChiSqM_Value	404
ChiSqM_DF	5
ChiSqM_PValue	0
ChiSqBaseline_PValue	0
LL	-143535
UnrestrictedLL	-143304
LLCorrectionFactor	1.48
UnrestrictedLLCorrectionFactor	1.4
CFI	0.974
TLI	0.947
AIC	287100
BIC	287205
aBIC	287158
RMSEA_Estimate	0.099
RMSEA_90CI_LB	0.091
RMSEA_90CI_UB	0.107
RMSEA_pLT05	0
SRMR	0.024
AICC	287100

Table A2

Gen2 Factor Loadings.

	Test	Estimate	S.E.	Est./.S.E.	P Value
1	DIGIT	0.547	0.010	54.400	0
2	MATH	0.731	0.007	97.900	0
3	RECOG	0.852	0.005	168.000	0
4	COMP	0.848	0.006	149.000	0
5	PPVT	0.743	0.009	81.400	0

Between Family Analyses

Table A3

Mean Gen1 Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean AFI		
	Mixed	Daughers	Sons
Gen1 Intel	0.012***	0.013***	0.014***
	(0.008, 0.017)	(0.009, 0.018)	(0.009, 0.019)
Constant	−0.754***	−0.801***	−0.856***
	(−1.010, −0.500)	(−1.090, −0.513)	(−1.140, −0.569)
Observations	342	264	282
R^2	0.090	0.101	0.106
Adjusted R^2	0.087	0.097	0.103
Residual Std. Error	0.700 (df = 340)	0.703 (df = 262)	0.691 (df = 280)
F Statistic	33.500*** (df = 1; 340)	29.400*** (df = 1; 262)	33.200*** (df = 1; 280)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A4

Mean Gen2 Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean AFI		
	Mixed	Daughers	Sons
Gen2 Intel	0.098***	0.105***	0.088***
	(0.039, 0.157)	(0.036, 0.173)	(0.022, 0.153)
Constant	−0.015	−0.015	−0.031
	(−0.093, 0.062)	(−0.104, 0.074)	(−0.117, 0.054)
Observations	345	267	283
R^2	0.030	0.033	0.024
Adjusted R^2	0.027	0.029	0.021
Residual Std. Error	0.734 (df = 343)	0.739 (df = 265)	0.728 (df = 281)
F Statistic	10.500*** (df = 1; 343)	8.990*** (df = 1; 265)	6.910*** (df = 1; 281)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A5

Mean Joint Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean AFI		
	Mixed	Daughers	Sons
INTMOMmean	0.012***	0.013***	0.014***
	(0.007, 0.017)	(0.007, 0.018)	(0.008, 0.019)
Int_Mean	0.021	0.018	0.008
	(−0.046, 0.088)	(−0.058, 0.094)	(−0.066, 0.082)
Constant	−0.714***	−0.771***	−0.844***
	(−1.010, −0.418)	(−1.100, −0.440)	(−1.180, −0.506)
Observations	338	260	278
R^2	0.092	0.103	0.107
Adjusted R^2	0.086	0.096	0.101
Residual Std. Error	0.703 (df = 335)	0.707 (df = 257)	0.695 (df = 275)
F Statistic	16.900*** (df = 2; 335)	14.700*** (df = 2; 257)	16.500*** (df = 2; 275)

Note: *p<0.1; **p<0.05; ***p<0.01

Within Family Analyses

Table A6

Dif Gen1 Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.303***	0.328***	0.376***
	(0.187, 0.419)	(0.192, 0.465)	(0.251, 0.501)
INTMOMmean	−0.003	−0.002	−0.008***
	(−0.007, 0.002)	(−0.008, 0.004)	(−0.013, −0.003)
INTMOMdiff	0.001	0.001	0.0003
	(−0.003, 0.006)	(−0.005, 0.006)	(−0.005, 0.005)
Constant	1.180***	1.180***	1.500***
	(0.893, 1.470)	(0.838, 1.520)	(1.180, 1.830)
Observations	336	258	278
R^2	0.074	0.083	0.116
Adjusted R^2	0.066	0.072	0.106
Residual Std. Error	0.757 (df = 332)	0.780 (df = 254)	0.734 (df = 274)
F Statistic	8.830*** (df = 3; 332)	7.640*** (df = 3; 254)	12.000*** (df = 3; 274)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A7

Dif Gen2 Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.326***	0.366***	0.357***
	(0.208, 0.444)	(0.228, 0.504)	(0.227, 0.486)
Int_Mean	−0.079**	−0.088**	−0.111***
	(−0.151, −0.008)	(−0.173, −0.003)	(−0.190, −0.032)
Intdiff	0.024	0.023	0.002
	(−0.020, 0.068)	(−0.029, 0.074)	(−0.044, 0.048)
Constant	1.040***	1.060***	1.050***
	(0.951, 1.130)	(0.959, 1.170)	(0.957, 1.150)
Observations	287	219	234
R^2	0.100	0.118	0.125
Adjusted R^2	0.091	0.105	0.114
Residual Std. Error	0.756 (df = 283)	0.778 (df = 215)	0.738 (df = 230)
F Statistic	10.500*** (df = 3; 283)	9.570*** (df = 3; 215)	11.000*** (df = 3; 230)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A8

Dif Joint Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.315^{***}	0.347^{***}	0.388^{***}
	(0.191, 0.438)	(0.202, 0.493)	(0.255, 0.521)
Int_Mean	-0.077^{*}	-0.088^{*}	-0.059
	(−0.161, 0.007)	(−0.187, 0.011)	(−0.150, 0.032)
INTMOMmean	0.0002	0.001	-0.007^{**}
	(−0.006, 0.006)	(−0.007, 0.008)	(−0.013, −0.0001)
Intdiff	0.018	0.021	0.00003
	(−0.028, 0.063)	(−0.033, 0.074)	(−0.047, 0.047)
INTMOMdiff	0.002	−0.001	0.002
	(−0.004, 0.007)	(−0.008, 0.005)	(−0.004, 0.007)
Constant	1.020^{***}	1.020^{***}	1.440^{***}
	(0.656, 1.380)	(0.597, 1.440)	(1.040, 1.840)
Observations	282	214	232
R^2	0.094	0.109	0.138
Adjusted R^2	0.077	0.087	0.119
Residual Std. Error	0.759 (df = 276)	0.783 (df = 208)	0.735 (df = 226)
F Statistic	5.710^{***} (df = 5; 276)	5.080^{***} (df = 5; 208)	7.260^{***} (df = 5; 226)

Note: *p<0.1; **p<0.05; ***p<0.01

Appendix B

Age 11.5 Replication

Table B1

Gen2 Measurement Model.

	g at Age 11.5
Estimator	MLR
Observations	7970
Parameters	15
ChiSqM_Value	273
ChiSqM_DF	5
ChiSqM_PValue	0
ChiSqBaseline_PValue	0
LL	-140683
UnrestrictedLL	-140524
LLCorrectionFactor	1.47
UnrestrictedLLCorrectionFactor	1.39
CFI	0.982
TLI	0.963
AIC	281396
BIC	281501
aBIC	281453
RMSEA_Estimate	0.082
RMSEA_90CI_LB	0.074
RMSEA_90CI_UB	0.09
RMSEA_pLT05	0
SRMR	0.02
AICC	281396

Table B2

Gen2 Factor Loadings.

	Test	Estimate	S.E.	Est./.S.E.	P.Value
1	DIGIT	0.551	0.010	53.500	0
2	MATH	0.744	0.007	102.000	0
3	RECOG	0.837	0.005	153.000	0
4	COMP	0.843	0.006	148.000	0
5	PPVT	0.751	0.009	83.500	0

Between Family Analyses

Table B3

Mean Gen1 Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean AFI		
	Mixed	Daughers	Sons
Gen1 Intel	0.012***	0.013***	0.014***
	(0.008, 0.017)	(0.009, 0.018)	(0.009, 0.019)
Constant	−0.754***	−0.801***	−0.856***
	(−1.010, −0.500)	(−1.090, −0.513)	(−1.140, −0.569)
Observations	342	264	282
R^2	0.090	0.101	0.106
Adjusted R^2	0.087	0.097	0.103
Residual Std. Error	0.700 (df = 340)	0.703 (df = 262)	0.691 (df = 280)
F Statistic	33.500*** (df = 1; 340)	29.400*** (df = 1; 262)	33.200*** (df = 1; 280)

Note: *p<0.1; **p<0.05; ***p<0.01

Table B4

Mean Gen2 Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean AFI		
	Mixed	Daughers	Sons
Gen2 Intel	0.104^{***}	0.117^{***}	0.091^{***}
	(0.044, 0.163)	(0.047, 0.187)	(0.027, 0.156)
Constant	−0.016	−0.007	−0.035
	(−0.094, 0.062)	(−0.097, 0.083)	(−0.121, 0.050)
Observations	339	262	278
R^2	0.034	0.040	0.027
Adjusted R^2	0.031	0.036	0.023
Residual Std. Error	0.729 (df = 337)	0.738 (df = 260)	0.721 (df = 276)
F Statistic	11.700^{***} (df = 1; 337)	10.700^{***} (df = 1; 260)	7.650^{***} (df = 1; 276)

Note: *p<0.1; **p<0.05; ***p<0.01

Table B5

Mean Joint Intelligence → Mean Gen2 AFI

	Dependent variable:		
	Mean AFI		
	Mixed	Daughers	Sons
INTMOMmean	0.012***	0.013***	0.014***
	(0.007, 0.017)	(0.007, 0.018)	(0.008, 0.020)
Int_Mean	0.010	0.017	−0.006
	(−0.060, 0.079)	(−0.064, 0.099)	(−0.081, 0.070)
Constant	−0.722***	−0.751***	−0.861***
	(−1.030, −0.416)	(−1.100, −0.401)	(−1.210, −0.511)
Observations	333	255	274
R^2	0.087	0.098	0.103
Adjusted R^2	0.082	0.091	0.096
Residual Std. Error	0.705 (df = 330)	0.709 (df = 252)	0.696 (df = 271)
F Statistic	15.800*** (df = 2; 330)	13.800*** (df = 2; 252)	15.600*** (df = 2; 271)

Note: *p<0.1; **p<0.05; ***p<0.01

Within Family Analyses

Table B6

Dif Gen1 Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.303***	0.328***	0.376***
	(0.187, 0.419)	(0.192, 0.465)	(0.251, 0.501)
INTMOMmean	−0.003	−0.002	−0.008***
	(−0.007, 0.002)	(−0.008, 0.004)	(−0.013, −0.003)
INTMOMdiff	0.001	0.001	0.0003
	(−0.003, 0.006)	(−0.005, 0.006)	(−0.005, 0.005)
Constant	1.180***	1.180***	1.500***
	(0.893, 1.470)	(0.838, 1.520)	(1.180, 1.830)
Observations	336	258	278
R^2	0.074	0.083	0.116
Adjusted R^2	0.066	0.072	0.106
Residual Std. Error	0.757 (df = 332)	0.780 (df = 254)	0.734 (df = 274)
F Statistic	8.830*** (df = 3; 332)	7.640*** (df = 3; 254)	12.000*** (df = 3; 274)

Note: *p<0.1; **p<0.05; ***p<0.01

Table B7

Dif Gen2 Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.344***	0.369***	0.382***
	(0.225, 0.463)	(0.232, 0.506)	(0.251, 0.512)
Int_Mean	−0.097***	−0.120***	−0.130***
	(−0.168, −0.026)	(−0.205, −0.035)	(−0.207, −0.052)
Intdiff	0.037	0.034	0.015
	(−0.007, 0.081)	(−0.016, 0.084)	(−0.032, 0.062)
Constant	1.050***	1.070***	1.080***
	(0.962, 1.140)	(0.963, 1.170)	(0.979, 1.170)
Observations	286	223	230
R^2	0.117	0.131	0.149
Adjusted R^2	0.108	0.119	0.138
Residual Std. Error	0.751 (df = 282)	0.773 (df = 219)	0.734 (df = 226)
F Statistic	12.500*** (df = 3; 282)	11.000*** (df = 3; 219)	13.200*** (df = 3; 226)

Note: *p<0.1; **p<0.05; ***p<0.01

Table B8

Dif Joint Intelligence → Dif Gen2 AFI

	Dependent variable:		
	Gen2 AFI Differences		
	Mixed	Daughers	Sons
AFIKIDmean	0.328^{***}	0.339^{***}	0.417^{***}
	(0.203, 0.452)	(0.194, 0.483)	(0.282, 0.552)
Int_Mean	-0.098^{**}	-0.131^{**}	-0.071
	(−0.184, −0.013)	(−0.234, −0.029)	(−0.163, 0.021)
INTMOMmean	0.001	0.002	-0.007^{**}
	(−0.005, 0.007)	(−0.005, 0.010)	(−0.014, −0.0001)
Intdiff	0.034	0.037	0.012
	(−0.011, 0.080)	(−0.016, 0.089)	(−0.036, 0.059)
INTMOMdiff	0.002	−0.002	0.001
	(−0.004, 0.007)	(−0.008, 0.005)	(−0.004, 0.007)
Constant	0.984^{***}	0.925^{***}	1.470^{***}
	(0.615, 1.350)	(0.489, 1.360)	(1.060, 1.890)
Observations	278	215	226
R^2	0.112	0.125	0.162
Adjusted R^2	0.096	0.104	0.143
Residual Std. Error	0.751 (df = 272)	0.775 (df = 209)	0.727 (df = 220)
F Statistic	6.860^{***} (df = 5; 272)	5.970^{***} (df = 5; 209)	8.500^{***} (df = 5; 220)

Note: *p<0.1; **p<0.05; ***p<0.01

References

Anastasi, A. (1956). Intelligence and family size. *Psychological Bulletin*, *53*(3), 187–209.

Baker, P. C., Keck, C. K., Mott, F. L., & Quinlan, S. V. (1993). *NLSY Child Handbook, Revised Edition: A Guide to the 1986-1990 NLSY Child Data.* Columbus, Ohio: The Ohio State University, Center for Human Resource Research.

Beasley, W. H., Rodgers, J. L., Bard, D. E., Hunter, M., Garrison, S. M., & Meredith, K. M. (2015). *NlsyLinks: Utilities and kinship information for research with the NLSY.* Retrieved from `http://liveoak.github.io/NlsyLinks/`

Belmont, L., & Marolla, F. A. (1973). Birth Order, Family Size, and Intelligence A study of a total population of 19-year-old men born in the Netherlands is presented. *Science*, *182*(4117), 1096–1101.

Berbaum, M. L., & Moreland, R. L. (1980). Intellectual development within the family: A new application of the confluence model. *Developmental Psychology*, *16*(5), 506.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Older and wiser? Birth order and IQ of young men. *CESifo Economic Studies*, *57*(1), 103–120.

Blake, J. (1981). Family size and the quality of children. *Demography*, *18*(4), 421–442.

Boislard, P. M. A., & Poulin, F. (2011). Individual, familial, friends-related and contextual predictors of early sexual intercourse. *Journal of adolescence*, *34*(2), 289–300.

Bouchard, Jr., T. J. (2004, aug). Genetic Influence on Human Psychological Traits. A Survey. *Current Directions in Psychological Science*, *13*(4), 148–151. Retrieved from `http://cdp.sagepub.com/content/13/4/148`

Bozon, M. (2003). At what age do women and men have their first sexual intercourse? World comparisons and recent trends. *Population and Societies*(391), 1–4.

Buhi, E. R., & Goodson, P. (2007). Predictors of Adolescent Sexual Behavior and Intention: A Theory-Guided Systematic Review. *Journal of Adolescent Health*, *40*(1), 4–21.

Center for Human Resources Research. (2009). *NLSY79 CHILD YOUNG ADULT DATA USERS GUIDE: A Guide to the 1986 - 2006 Child Data 1994 - 2006 Young Adult Data.* Bureau of Labor Statistics, U.S. Department of Labor. Retrieved from

https://www.nlsinfo.org/pub/usersvc/Child-Young-Adult/2006ChildYA-DataUsersGuide.pdf

Clark, R. D., & Rice, G. A. (1982). Family constellations and eminence: The birth orders of Nobel Prize winners. *The Journal of Psychology*, *110*(2), 281–287.

Damian, R. I., & Roberts, B. W. (2015). The associations of birth order with personality and intelligence in a representative sample of US high school students. *Journal of Research in Personality*, *58*, 96–105.

Defense Manpower Data Center. (2012). *ASVAB Technical Bulletin No. 4 PP-ASVAB Forms 23-27* (Tech. Rep. No. 4).

Der, G., Batty, G. D., & Deary, I. J. (2009). Intelligence The association between IQ in adolescence and a range of health outcomes at 40 in the 1979 US National Longitudinal Study of Youth. *Intelligence*, *37*(6), 573–580.

Devlin, B., Daniels, M., & Roeder, K. (1997). The heritability of IQ. *Nature*, *388*(6641), 468–471.

DiBlasio, F. A., & Benda, B. B. (1990). Adolescent Sexual Behavior Multivariate Analysis of a Social Learning Model. *Journal of Adolescent Research*, *5*(4), 449–466.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised: PPVT-R*. American Guidance Service.

Dunn, L. M., & Markwardt, F. C. (1970). *Peabody individual achievement test*. American Guidance Service.

Engel, G. L. (1977). The need for a new medical model: a challenge for biomedicine. *Science*, *196*(4286), 129–136. Retrieved from http://www.sciencemag.org/content/196/4286/129://www.ncbi.nlm.nih.gov/pubmed/847460

Finer, L. B. (2007). Trends in premarital sex in the United States, 1954–2003. *Public Health Reports*, *122*(1), 73.

Galton, F. (1875). *English men of science: Their nature and nurture*. New York: D. Appleton and Co.

Halpern, C. T., Joyner, K., Udry, J. R., & Suchindran, C. (2000). Smart teens don't have sex (or kiss much either). *Journal of Adolescent Health*, *26*(3), 213–225.

Harden, K. P. (2012). True Love Waits? A Sibling-Comparison Study of Age at First Sexual Intercourse and Romantic Relationships in Young Adulthood. *Psychological Science*, *23*(11), 1324–36.

Harden, K. P. (2014). Genetic influences on adolescent sexual behavior: Why genes matter for environmentally oriented researchers. *Psychological bulletin*, *140*(2), 434–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23855958

Harden, K. P., & Mendle, J. (2011). Why Don't Smart Teens Have Sex? A Behavioral Genetic Approach. *Child Development*, *82*(4), 1327–1344.

Harden, K. P., Mendle, J., Hill, J. E., Turkheimer, E., & Emery, R. E. (2008). Rethinking timing of first sex and delinquency. *Journal of Youth and Adolescence*, *37*(4), 373–385.

Herrnstein, R. J., & Murray, C. (1994). *Bell curve: Intelligence and class structure in American life.* New York, NY: Free Press.

Hirschi, T. (2002). *Causes of delinquency*. Transaction publishers.

Hogben, M., & Byrne, D. (1998). Using social learning theory to explain individual differences in human sexuality. *Journal of Sex Research*, *35*(1), 58–71.

Hutchinson, M. K., Jemmott, J. B., Jemmott, L. S., Braverman, P., & Fong, G. T. (2003). The role of mother–daughter sexual risk communication in reducing sexual risk behaviors among urban adolescent females: a prospective study. *Journal of adolescent health*, *33*(2), 98–107.

Kaestle, C. E., Halpern, C. T., Miller, W. C., & Ford, C. A. (2005). Young age at first sexual intercourse and sexually transmitted infections in adolescents and young adults. *American Journal of Epidemiology*, *161*(8), 774–780.

Kass, R. A., Mitchell, K. J., Grafton, F. C., & Wing, H. (1982). *Factor Structer of the Armed Services Vocational Aptitude Battery (ASVAB), Forms 8, 9, and 10: 1981 Army Applicant Sample* (Tech. Rep.). US Army Research Institute for the Behavioral and Social Sciences.

Kearney, M. S., & Levine, P. B. (2014, jan). Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing. Retrieved from http://www.nber.org/papers/w19795

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. Guilford Press.

Kenny, D. A., Mohr, C. D., & Levesque, M. J. (2001). A social relations variance partitioning of dyadic behavior. *Psychological Bulletin*, *127*(1), 128.

Kirby, D. (2002a). Antecedents of adolescent initiation of sex, contraceptive use, and pregnancy. *American journal of health behavior*, *26*(6), 473–485.

Kirby, D. (2002b). Effective approaches to reducing adolescent unprotected sex, pregnancy, and childbearing. *Journal of sex research*, *39*(1), 51–57.

Lahey, B. B., & D'Onofrio, B. M. (2010). All in the Family: Comparing Siblings to Test Causal Hypotheses Regarding Environmental Influences on Behavior. *Current Directions in Psychological Science*, *19*(5), 319–323.

Lammers, C., Ireland, M., Resnick, M., & Blum, R. (2000). Influences on adolescents' decision to postpone onset of sexual intercourse: A survival analysis of virginity among youths aged 13 to 18 years. *Journal of Adolescent Health*, *26*(1), 42–48.

Leitenberg, H., & Saltzman, H. (2000). A Statewide Survey of Age at First Intercourse for Adolescent Females and Age of Their Male Partners: Relation to Other Risk Behaviors and Statutory Rape Implications. *Archives of Sexual Behavior*, *29*(3), 203–215.

Lord, F. M. (1963). Elementary Models for Measuring Change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 22–38). Madison: University of Wisconsin Press.

Maier, M. H., & Sims, W. H. (1986). *The ASVAB score scales: 1980 and World War II* (Tech. Rep.). Alexandria, VA: Center for Naval Analyses.

Manlove, J. (1998). The influence of high school dropout and school disengagement on the risk of school-age pregnancy. *Journal of Research on Adolescence*, *8*(2), 187–220.

Mathews, C., Aarø, L. E., Flisher, A. J., Mukoma, W., Wubs, A. G., & Schaalma, H. (2009). Predictors of early first sexual intercourse among adolescents in Cape Town, South Africa. *Health Education Research*, *24*(1), 1–10.

Mathews, J. J. (1977). *Analysis Aptitude Test for Selection of Airmen for the Radio Communications Analysis Specialist Course: Development and Validation* (Tech. Rep.). Brooks Air Force Base, Texas: Air Force Human Resources Lab.

Mayberry, P. W., & Hiatt, C. M. (1992). *Computing AFQT Scores from Historical Data* (Tech. Rep.). DTIC Document.

McGrevy, D. F., Knouse, S. B., & Thompson, R. A. (1974). *Relationships among an individual intelligence test and two air force screening and selection tests* (Tech. Rep.). DTIC Document.

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

Meredith, K. M. (2013). *Is AFI All in the Family? A Multi-Level Family Study of Age of First Intercourse* (Dissertation). University of Oklahoma.

Meschke, L. L., Zweig, J. M., Barber, B. L., & Eccles, J. S. (2000). Demographic, Biological, Psychological, and Social Predictors of the Timing of First Intercourse. *Journal of Research on Adolescence*, *10*(3), 315–338.

Miller, B. C., Norton, M. C., Curtis, T., Hill, E. J., Schvaneveldt, P. L., & Young, M. H. (1997). The timing of sexual intercourse among adolescents family, peer, and other antecedents. *Youth and Society*, *29*(1), 54–83.

Miller, W. B., Pasta, D. J., MacMurray, J., Chiu, C., Wu, H., & Comings, D. (1999). Dopamine receptor genes are associated with age at first sexual intercourse. *Journal of Biosocial Science*, *31*(01), 43–54.

Mott, F. L. (1983). Early fertility behavior among American youth: Evidence from the 1982 national longitudinal surveys of labor force behavior of youth. In *America public health association meetings, dallas.*

Mott, F. L., & Baker, P. C. (1995). *The NLSY children 1992: Description and evaluation.* Center for Human Resource Research, Ohio State University.

Murray, C. (1998). *Income inequality and IQ.* American Enterprise Inst. for Public Policy Research.

Muthén, L. K., & Muthén, B. O. (2014). *Mplus.*

Nedelec, J. L., Schwartz, J. A., Connolly, E. J., & Beaver, K. M. (2012). Exploring the association between IQ and differential life outcomes: Results from a longitudinal sample of monozygotic twins. *Temas em Psicologia*, *20*(1), 31–43.

Neisser, U., Boodoo, G., Bouchard, Jr., T. J., Boykin, A. W., Brody, N., Ceci, S. J., ... Urbina, S. (1996). *Intelligence: Knowns and unknowns.* (Vol. 51) (No. 2).

Open Science Collaboration. (2015, aug). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716–aac4716. Retrieved from http://www.sciencemag.org/content/349/6251/aac4716.abstract

Page, E. B., & Grandon, G. M. (1979). Family configuration and mental ability: Two theories contrasted with US data. *American Educational Research Journal*, *16*(3), 257–272.

Palmer, P., Hartke, D. D., Ree, M. J., Welsh, J. R., & Valentine, Jr., L. D. (1988). *Armed Services Vocational Aptitude Battery (ASVAB): Alternate Forms Reliability (Forms 8, 9, 10, and 11)* (Tech. Rep.). Brooks Air Force Base, Texas: Air Force Human Resources Laboratory.

Paul, C., Fitzjohn, J., Herbison, P., & Dickson, N. (2000, aug). The determinants of sexual intercourse before age 16. *Journal of Adolescent Health*, *27*(2), 136–147.

Petersen, A. C. (1987). The nature of biological-psychosocial interactions: The sample case of early adolescence. *Biological-psychosocial interactions in early adolescence. Child psychology.*, 35–61.

Plomin, R., & Spinath, F. M. (2004). Intelligence: genetics, genes, and genomics. *Journal of personality and social psychology*, *86*(1), 112.

Raffaelli, M., & Crockett, L. J. (2003). Sexual risk taking in adolescence: the role of self-regulation and attraction to risk. *Developmental psychology*, *39*(6), 1036.

Retherford, R. D., & Sewell, W. H. (1991). Birth order and intelligence: Further tests of the confluence model. *American Sociological Review*, 141–158.

Rodgers, J. L. (1996). Sexual transitions in adolescence. In J. A. Graber, J. Brooks-Gunn, & A. C. Petersen (Eds.), *Transitions through adolescence: Interpersonal domains and context* (pp. 85–110). Mahwah, New Jersey.

Rodgers, J. L. (2014). Are birth order effects on intelligence really Flynn Effects? Reinterpreting Belmont and Marolla 40years later. *Intelligence*, *42*, 128–133.

Rodgers, J. L., Beasley, W. H., Bard, D. E., Meredith, K. M., Hunter, M., Johnson, A. B., ... Rowe, D. C. (2015). *The NLSYKinship Links: Using the NLSY79 and NLSY-Children Data to Conduct Genetically-Informed and Family-Oriented Research.*

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*(6), 599–612.

Rodgers, J. L., Kohler, H.-P., McGue, M., Behrman, J. R., Petersen, I., Bingley, P., & Christensen, K. (2008). Education and Cognitive Ability as Direct, Mediating, or Spurious Influences on Female Age at First Birth: Behavior Genetic Models Fit to Danish Twin Data. *American Journal of Sociology*, *114*(S1), S202–32.

Rodgers, J. L., Rowe, D. C., & Buster, M. (1999). Nature, nurture and first sexual intercourse in the USA: fitting behavioural genetic models to NLSY kinship data. *Journal of Biosocial Science*, *31*(01), 29–41.

Rowe, D. C., & Rodgers, J. L. (1997). Poverty and Behavior : Are Environmental Measures Nature and Nurture ? *Developmental Review*, *17*(3), 358–375.

Rutter, M. (2007). Proceeding From Observed Correlation to Causal Inference: The Use of Natural Experiments. *Perspectives on Psychological Science*, *2*(4), 377–395.

Santelli, J. S., & Beilenson, P. (1992). Risk factors for adolescent sexual behavior, fertility, and sexually transmitted diseases. *Journal of School Health*, *62*(7), 271–279.

Savage, D. S., & Freeman, M. J. (2009). *16 and Pregnant.* MTV.

Schvaneveldt, P. L., Miller, B. C., Berry, E. H., & Lee, T. R. (2001). Academic goals, achievement, and age at first sexual intercourse: Longitudinal, bidirectional influences. *Adolescence*, *36*(144), 767.

Spriggs, A. L., & Halpern, C. T. (2008). Timing of sexual debut and initiation of postsecondary education by early adulthood. *Perspectives on sexual and reproductive health*, *40*(3), 152–161.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, *35*(5), 401–426.

Udry, J. R. (1979). Age at menarche, at first intercourse, and at first pregnancy. *Journal of Biosocial Science*, *11*(04), 433–441.

U.S. Coast Guard. (2004). *Armed Services Vocational Aptitude Battery (ASVAB) renorming.* (Tech. Rep.). U.S. Department of Homeland Security.

U.S. Department of Defense. (1982). *Profile of American Youth: 1980 nationwide administration of the Armed Services Vocational Aptitude Battery.* Washington, DC: Office of the Assistant Secretary of Defense (Manpower, Installations and Logistics).

U.S. Department of the Army. (2013). *Pocket Recruiter Guide.* Fort Knox, KY: Government Printing Agency.

Waters, B. K., Laurence, J. H., Camara, W. J., & Green, B. F. (1987). *Personnel enlistment and classification procedures in the US military.* National Academy Press.

Wechsler, D. (1974). *Manual for the Wechsler intelligence scale for children, revised.* Psychological Corporation.

Wellings, K., Nanchahal, K., Macdowall, W., McManus, S., Erens, B., Mercer, C. H., ... Field, J. (2001, dec). Sexual behaviour in Britain: Early heterosexual experience. *Lancet*, *358*(9296), 1843–1850.

Welsh, J. R., Kucinkas, S. K., & Curran, L. T. (1990). *Armed Services Vocational Aptitude Battery (ASVAB: Integrative Review of Validity Studies* (Tech. Rep.). Brooks Air Force Base, Texas: Manpower and Personnel Division Air Force Human Resources Laboratory.

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*(1), 117–127.

Wichman, A. L., Rodgers, J. L., & Maccallum, R. C. (2007). Birth order has no effect on intelligence: a reply and extension of previous findings. *Personality and social psychology bulletin*, *33*(9), 1195–1200.

Wilbourn, J., Valentine, Jr., L. D., & Ree, M. J. (1984). *Relationship of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 8, 9, and 10 to Air Force Technical School Final Grades* (Tech. Rep.). Brooks Air Force Base, Texas: Manpower and Personnel Division.

Woodward, L. J., Fergusson, D. M., & Horwood, L. J. (2001). Risk Factors and Life Processes Associated with Teenage Pregnancy: Results of a Prospective Study From Birth to 20 Years. *Journal of Marriage and Family*, *63*(4), 1170–1184.

Zajonc, R. B. (1976). Family configuration and intelligence: Variations in scholastic aptitude scores parallel trends in family size and the spacing of children. *Science*.

Zajonc, R. B., & Bargh, J. (1980). The Confluence Model : Parameter Estimation for Six Divergent Data Sets on Family Factors and Intelligence. *Intelligence*, *4*, 349–361.

Zimmer-Gembeck, M. J., & Helfand, M. (2008). Ten years of longitudinal research on US adolescent sexual behavior: Developmental correlates of sexual intercourse, and the importance of age, gender and ethnic background. *Developmental Review*, *28*(2), 153–224.