Casting doubt on the causal link between intelligence and age at first intercourse:

A cross-generational sibling comparison design using the NLSY

Abstract

Halpern *et al.* (2000) published a study based on early Add Health data with the provocative title "Smart Teens Don't Have Sex (or Kiss Much Either)." Several following papers reported the same result, a positive correlation between the intelligence of adolescent girls and age at first intercourse (AFI). However, the causal mechanism has not been carefully investigated. Harden and Mendle (2011) used Add Health data within a biometrical design and found that the relationship between intelligence and AFI was fully accounted for by shared environmental differences, suggesting at least the location of the causal mechanism - the part of the household environment shared by siblings that influences both child intelligence and AFI. In this study, we use an intergenerational sibling comparison design to investigate the causal link between intelligence and AFI, using the National Longitudinal Survey of Youth 1979 and the NLSY-Children/Young Adult data. We measured maternal IQ using the AFQT, child IQ using PPVT, PIAT, and Digit Span, and AFI, using respondent self-report. Our analytic method used Kenny's (2001) reciprocal standard dyad model. This model supported analyses treating the data as only between-family data (as in most past studies), and also allowed us to include both between- and within-family comparisons. These analyses included two forms, first a comparison of offspring of mothers in relation to maternal IQ, then a comparison of offspring themselves in relation to offspring IQ.

When we evaluated the relationship between maternal/child intelligence and AFI, using a between-family design, we replicated earlier results; smart teens do appear to delay sex. In the within-family analyses, the relationship between intelligence and AFI vanishes for both maternal intelligence and child intelligence. The finding is robust across gender and age. These results suggest that the cause of the intelligence-AFI link is not intelligence *per se*, but rather differences between families (parental education, SES, *etc.*) that correlate with family-level (but not individual-level) intelligence.

Keywords: Age at first intercourse; Cross-sectional Data; Discordant Sibling Design; Intelligence; Quasi-Experimental; Siblings

Casting doubt on the causal link between intelligence and age at first intercourse:

A cross-generational sibling comparison design using the NLSY

Teenage sexual activity has interested academics in demography, public health, and psychology for many years (Brooks-Gunn & Furstenberg, 1989; Kinsey, Pomeroy, Martin, & Gebhard, 1953; Santelli, Lindberg, Abma, McNeely, & Resnick, 2000). Anecdotal evidence from the popular media, (*e.g.*, MTV's reality television franchise, *16 and Pregnant*), and academic research converge to document a relatively consistent secular decline in age at first intercourse (AFI; see Bozon, 2003; Finer, 2007; Kann et al., 2014). Early AFI is associated with downstream consequences, including lower educational 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 (STIs; Kaestle, Halpern, Miller, & Ford, 2005). Moreover, beyond the obvious benefit of avoiding those negative outcomes, delaying AFI is associated with greater relationship satisfaction, perception of increased attractiveness, and higher household income (Harden, 2012). Because many of the negative consequences above are severe and long-reaching, it is important to identify the causal mechanisms associated with early AFI. One potential factor exerting causal influence on AFI is intelligence.

Higher levels of intelligence are associated with delaying first intercourse (Halpern et al., 2000; Mott, 1983; Paul, Fitzjohn, Herbison, & Dickson, 2000; Woodward, Fergusson, & Horwood, 2001), and also with delay of less-intimate sexual involvement (Halpern et al., 2000). Specifically, intelligent individuals may 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 life- and career-shattering outcomes (Halpern et al., 2000; Harden & Mendle, 2011). Although the link between intelligence and AFI has face validity, and has been confidently asserted (or often implied) as a causal link, a fundamental confound exists in most past research that limits our ability to infer causality.

Virtually all of the AFI-intelligence literature has used between-family designs. In analysis of data from such designs, a number of genetic and environmental influences, such as education and maternal intelligence, are confounded (D'Onofrio, Lahey, Turkheimer, & Lichtenstein, 2013; Harden, 2014; Lahey & D'Onofrio, 2010; Rodgers, Cleveland, van den Oord, & Rowe, 2000). By ignoring such confounds, the source of variance is ambiguous, and researchers that attribute the source to specific between- or within-family sources risk misattributions of causality (Rowe & Rodgers, 1997; Rutter, 2007). There is the potential for this type of confound in virtually all past research on the link between intelligence and AFI (Harden & Mendle, 2011; Harden, 2014; Plomin & Spinath, 2004; Rodgers, Rowe, & Buster, 1999; Rodgers, Rowe, & May, 1994). Thus, we need to critically evaluate whether intelligence has a causal influence on AFI or is rather a theoretically attractive confound. Sibling models have been used in the past to at least partially control for confounds that create causal ambiguity (*e.g.* East, 1996; East, Felice, & Morgan, 1993; Geronimus & Korenman, 1992; Rodgers & Rowe, 1990; Rodgers, Rowe, & Harris, 1992); several of these studies used the sibling structure in earlier versions of the NLSY. As in those early studies, we resolve some of these methodological challenges by using design innovations that emerge from the excellent cross-generational and longitudinal structure of the National Longitudinal Survey of Youth (NLSY; we use both the original NLSY79 survey and the NLSY-Children survey, described later).

Cause or Confound?

Numerous theories address the motivations for adolescents' initiation of first intercourse (see Rodgers, 1996 or Buhi & Goodson, 2007 for reviews), and even more specific precursors to first intercourse (Buhi & Goodson, 2007; D'Onofrio & Lahey, 2010; Kirby, 2002a; B. C. Miller et al., 1997; Santelli & Beilenson, 1992). Many of these theories emphasize biology/genetics, as adolescent pubertal development (and associated hormone changes) drives the onset of sexual behavior (W. B. Miller et al., 1999; Udry, 1979, 1994). Other theoretical frameworks use social/environmental

processes to explain developing sexual involvement in adolescence, such as Social Learning (DiBlasio & Benda, 1990; Hogben & Byrne, 1998), where social norms affect the likelihood of early sexual behavior; or Social Control theory (Hirschi, 2002), where societal and cultural influences reduce the likelihood that individuals will act on their natural tendency toward sexual involvement. Under these environmental theories the underlying biology is typically ignored, whereas under many of the biological/genetic theories, the environmental components are often ignored.

However, numerous articles have advocated integrative models (See Harden, Mendle, Hill, Turkheimer, & Emery, 2008; Harden, 2014; Udry, 1995). The integrative Biopsychosocial Model acknowledges both genetic and environmental contributions to human behavior (Engel, 1977; Petersen, 1987; Rodgers et al., 1999). Indeed, biology, psychology, and society/culture jointly influence adolescents' decisions to engage in sexual intercourse (Meschke, Zweig, Barber, & Eccles, 2000; Zimmer-Gembeck & Helfand, 2008).

Intelligence as a Cause

The short-term risks of early AFI are primarily negative, whereas the rewards for delay are primarily positive. These consequences extend into adulthood – early AFI has been related to adult delinquency (Harden et al., 2008), anti-social behavior, and substance abuse (Boislard & Poulin, 2011), whereas those with delayed AFI have higher household incomes in adulthood (Harden, 2012). It is intuitively appealing to believe that intelligent individuals are more likely to observe this potential risk-reward tradeoff, and through volition, act upon such observations by delaying first intercourse. Accordingly, intelligent individuals perceive the consequences of early AFI to negatively influence their lives and careers (Halpern et al., 2000; Harden & Mendle, 2011).

Indeed, most of the literature has contributed to the expectation that intelligence is causally connected to AFI. Those with higher educational goals delay their first intercourse (Boislard & Poulin, 2011; Schvaneveldt, Miller, Berry, & Lee, 2001), whereas those who engaged in early sexual intercourse reduced their educational goals compared

to earlier higher 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). Besides delaying first intercourse, smarter individuals appear to postpone all sexual/romantic activity (Halpern et al., 2000). Such blanket delays may be a proactive attempt to avoid "gateway" activities that might lead to intercourse. 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, p. 213).

However, Halpern et al. (2000) and many of the other studies we have referenced (e.g., C. Mathews et al., 2009; B. C. Miller et al., 1997; Paul et al., 2000) used between-family – typically cross-sectional – designs.[1] Such designs cannot logically distinguish between processes that act to create differences between families and processes that create differences among family members (Lahey & D'Onofrio, 2010; Rodgers et al., 2000). Thus the previous studies do not provide conclusive evidence that intelligence is the causal influence behind the AFI-intelligence relationship. Logically, other alternatives are that AFI has a causal link to intelligence (which is unlikely, for the obvious theoretical reasons, including that a child's intelligence precedes AFI in time) or that other confounds cause these two outcomes to correlate, but not causally. There are dozens, perhaps hundreds, of such confounds that can logically contend to explain the link between child intelligence and AFI; we review some of those confounds in the next section.

---

[1]Halpern et al. (2000) used the first wave of Add Health for most of their AFI-intelligence analyses.

Intelligence as a Confound

An equally valid set of explanations exist in which intelligence is not the causal factor behind the AFI-intelligence relationship, but rather one of dozens of correlated potentially explanatory processes. Instead, various confounds including family-level selection effects, or third variables at the individual or family level could be causing the relationship. Between-family influences such as SES and maternal intelligence could drive the relationship. Socioeconomic status is associated with the onset of first intercourse (Lammers, Ireland, Resnick, & Blum, 2000), which potentially explains many of the negative consequences linked with teenage pregnancy (Geronimus & Korenman, 1992), and is correlated with intelligence (Murray, 1998; Neisser et al., 1996; Strenze, 2007). Parental intelligence and parental education are also linked with child intelligence (Bouchard, 2004; Devlin, Daniels, & Roeder, 1997; Mercy & Steelman, 1982), and pose viable alternative explanations in which parents could be influencing, or even actively 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 better understand the causal relation between intelligence and AFI, we need to be able to untangle between- and within-family processes, using both data and designs that have the ability to separate these sources of variance.

Indeed, many such findings that link intelligence with various outcomes are quite possibly the result of misattributing between-family confounds to individual-level and within-family causes. The relationship between birth order and intelligence is a classic example of this misattribution (See Damian & Roberts, 2015; Rodgers et al., 2000; Rodgers, 2014). We briefly review that research arena here, because it illustrates the same challenge that occurs in studying the link between intelligence and AFI.

Between-family studies that rely on cross-sectional data have consistently found

that first born children have higher IQs than later born children (Belmont & Marolla, 1973; Zajonc, 1976). Yet within-family studies have typically found a non-significant relationship (Berbaum & Moreland, 1980; Galbraith, 1982; Retherford & Sewell, 1991; Rodgers, 1984; Rodgers et al., 2000; also see Barclay, 2015 and Bjerkedal, Kristensen, Skjeret, & Brevik, 2007 for recent exceptions that have found small, but significant within-family effects in large national studies). Moreover, when designs that can distinguish within- and between-family variance have been conducted, the methodological source of the IQ-birth order effects have emerged from the between-family variance (Black, Devereux, & Salvanes, 2011; Rodgers, 1984; Rodgers et al., 2000; Wichman, Rodgers, & Maccallum, 2006; Wichman, Rodgers, & MacCallum, 2007). Potential causes of this confound include maternal age at first birth, parental IQ, parental education, and SES (Page & Grandon, 1979; Rodgers, 2001; Rodgers et al., 2008, also see Anastasi (1956) for an insightful overview, written prior to the IQ-birth order debate).

In summary, if there is a valid within-family link between intelligence and birth order (which is, definitionally, a within-family variable), it is at one extreme of small magnitude and only detectable in large national datasets, and at the other extreme non-existent (even in large U.S. datasets). Similarly, the link between intelligence and AFI may be mostly or completely spurious. For example, socioeconomic status is associated with the onset of first intercourse (Lammers et al., 2000), which implicates a between-family process as explanatory of early AFI and teenage pregnancy association with negative consequences.

Prior Within-Family Analyses

Two past studies have explicitly separated between- and within-family influences on the AFI-intelligence relationship (Harden & Mendle, 2011; Nedelec, Schwartz, Connolly, & Beaver, 2012). Harden and Mendle used 536 same-sex identical and fraternal twin pairs from the National Longitudinal Study of Adolescent to Adult Health Study (Add Health) to "test[ ] whether relations between intelligence, academic

achievement and age at first sex were due to unmeasured genetic and environmental differences between families." Twins who differed in their intelligence or their academic achievement did not differ in their age at first intercourse. They concluded that "the association between intelligence and age at first sex could be attributed entirely to *unmeasured environmental differences between families*"(italics added, our own emphasis). Nedelec, Schwartz, Connolly, and Beaver (2012) conducted an exploratory analysis of MZ twin pairs from the same sample employed by Harden and Mendle, and Halpern et al. (2000), using intelligence difference scores to predict various social outcomes. They found consistent null results, though their samples were small and their statistical analyses were substantially underpowered.

Current Study, Summary

To summarize, the current study examines the relationship between intelligence and age at first intercourse, using maternal siblings and their children from a multi-generational nationally representative sample, the NLSY. This examination extends the intelligence literature in several key ways. First, we test whether the relationship between intelligence and AFI existed in either or both between- and within-family analyses, using data in which we can explicitly separate those sources of variance. Second, we evaluated the alternative explanation that maternal intelligence influences child AFI, using the cross-generational structure of the NLSY. Third, we replicated our findings using two different age periods. Fourth, we address overlapping questions to those studied in Halpern *et al.* (2000, Harden and Mendle (2011), and Nedelec *et al.* (2012), using a different data source.[2]

We made the following predictions, based primarily upon Harden & Mendle (2011):

Between Families,

1. Does Generation 2 (*i.e.*, children's) intelligence predict Generation 2 AFI?: We expect intelligence to be associated with age of first intercourse because there is a

---

[2]All of those previous studies used the Add Health dataset.

sizable body of literature reporting that result (Kirby, 2002b; Manlove, 1998; Raffaelli
& Crockett, 2003; Rodgers et al., 1994).

  2. Does Generation 1 (*i.e.*, mother's) intelligence predict Generation 2 AFI?: We
also expect maternal intelligence to be associated with age of first intercourse because
the heritability of intelligence is moderate-to-high (Bouchard, 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 weaker, but existent. However, if the
intelligence-AFI relationship is the product of between-family confounds, then we would
expect that the cross-generational association 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. In other words,
maternal intelligence can serve as a proxy for many of the between-family confounds
that are of concern in the current study. Comparably sized effects would also be
consistent with a between-family confound. Given that 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,

  3. Does Generation 2 intelligence predict Generation 2 AFI?: We do not expect
to find within-family differences in intelligence correlating with AFI, given that Harden
& Mendle (2011) did not report an effect.

  4. Does Generation 1 intelligence predict Generation 2 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.

  5. Is the relationship consistent across cross-sectional and within-family designs?:
Doubtful, we do not expect the results to be consistent across methods because both
Harden & Mendle (2011) and Meredith (2013) found no within-family effect, whereas
the traditional findings from between-family studies find an effect (Kirby, 2002b;
Manlove, 1998; Raffaelli & Crockett, 2003). The birth order literature is one example in
which apparent causal links mitigate or disappear when variance is properly attributed

to within- versus between-family sources.

Method

Research Design and Analytic Approach

We adapted Kenny and colleagues (2001; 2006) reciprocal standard dyad model to facilitate sibling comparisons. Sibling-based quasi-experimental models are particularly effective for incorporating genetic and environmental design elements (Lahey & D'Onofrio, 2010; Rutter, 2007). Our model uses differences between both pairs of mothers and pairs of adolescents, which provide the measures that explicitly account for within-family variance. Further, within-family differences create a powerful control for virtually all background heterogeneity (variance) associated with both genetic and environmental differences (Lahey & D'Onofrio). We compare individuals from within the family in the context of the following models. First, we predict the difference in AFI, $Y_{i\Delta}$, for a given pair of NLSY-Children, indexed as i, in the following model:

$$Y_{i\Delta} = \beta_0 + \beta_1 \bar{Y}_i + \beta_2 \bar{X}_i + \beta_3 X_{i\Delta} \tag{1}$$

where,

$$Y_{i\Delta} = Y_{i1} - Y_{i2}; \ X_{i\Delta} = X_{i1} - X_{i2}, \text{ and} \tag{2}$$

$$Y_{i1} = \max(Y_{ij}); \ Y_{i2} = \min(Y_{ij}) \tag{3}$$

In this model, the relative difference in kin outcomes (Y_Δ; *e.g.*, AFI) is predicted from the mean level of Y ($\bar{Y}$; *e.g.* mean AFI), the mean level of X ($\bar{X}$; *e.g.*, intelligence), and the between-kin intelligence difference ($X_{i\Delta}$). The mean levels, reflecting between-family variance, support causal inference through at least partial control for genes and shared environment in previous generations. Within this model, there is explicit separation of within-family variance (with Y_Δ and X_Δ), and between-family variance (with $\bar{Y}$ and $\bar{X}$).

For our analytic models, we use either the complete version or subsets of this overall analytic model to achieve different goals (elaborated below). We can use only the between-family independent variables, only the within-family independent variable, or both (the first and the third models achieve the goals of the current study).

Thus, this model explicitly untangles between- and within-family influences. If there is a true causal link between intelligence and AFI, then we expect kin differences in intelligence to be significantly associated with kin differences in AFI. If the effect is spurious – only a function of between-family confounds – 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 dataset(NLSY79) is based on a nationally representative household probability sample, jointly sponsored by the U.S. Bureau of Labor Statistics and the U.S. Department of Defense. On December 31, 1978, 12,686 adolescents were sampled within a household probability sample from 8,770 households. The initial sample consisted of three subsamples:

- a cross–sectional household probability sample of 6,111 non-institutionalized adolescents residing in the United States on December 31st of 1978;

- a separate over-sampled civilian subsample of 5,295 racial minorities and disadvantaged whites;

- a representative sample of 1,280 youth serving in the U.S. Military on September 30th, 1978.

In the two civilian samples, subjects had birthdates that ranged from January 1, 1957 to December 31, 1964, and were between the ages of 14 and 21 on December 31, 1978; military subject's birthdates ranged from January 1, 1957 to December 31, 1961, and were between 17 and 21 years old. Participants were surveyed annually until 1994, and then surveyed biennially to the present. 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. Note that because there are no siblings within the military sample, it is irrelevant for the current research, as all military respondents are screened out by the

requirement of having siblings within the sample. More information about the sampling

process and the data can be found on the Bureau of Labor Statistics (BLS) website:

`http://www.bls.gov/nls/nlsy79.htm`

In 1986, all biological children of the female NLSY79 participants were surveyed

for the first round of the NLSY79 Children and Young Adults (NLSY-Children) survey.

Now that childbearing is complete among the NLSY79 females (who were 49 to 57 in

the 2014 survey), a total of 11,512 respondents are in the NLSY-Children surveys,

which is continuing on a biennial basis. Participants in the NLSY79 will typically be

referred to as the Generation 1 (Gen1) sample, whereas the NLSY-Children will be

referred to as the Generation 2 (Gen2) sample.

Tetrads

To conduct our study using the requisite within-family information, we require

sister pairs in Generation 1 who both had children. The children of these sisters are

cousin pairs. In the original NLSY79 and NLSY-Children surveys, identification of level

of sibling relatedness in the NLSY was primarily inferential. NLSY79 twins, full

siblings, half siblings, and adoptive siblings were distinguishable indirectly from

respondent and maternal information about birthdates and the biological father(s).

NLSY-Children respondents within a given family were all full- or half-siblings, because

they were (by design) the biological children of the NLSY79 females. In 2006, both

NLSY surveys included explicit indicators of the level of sibling relatedness. Our

research team has recently completed a multi-year project to reliably and validly

identify the kinship pairs within these two datasets (Rodgers et al., 2016). Sibling and

cousin pairs from these kinship links are used in the current study.

Specifically, Mother-Child-Aunt-Nibling (MCAN) tetrads were created using the

NLSY Kinship Links (Rodgers et al., 2016) and supporting R package (Beasley et al.,

2015). The oldest two female kin (Mother, Aunt) were selected from each NLSY79

household (note that additional female Generation 1 sister pairs from families with

three or more sisters – a relatively small number – were excluded). Three tetrad designs

were employed, in which the genders of Generation 2 were the defining feature:

- Mother-Daughter-Aunt-Niece (MDAN) tetrads included the oldest Generation 2 female child from each of the two Generation 1 sisters,

- Mother-Son-Aunt-Nephew (MSAN) tetrads included the oldest Generation 2 male child from each of the Generation 1 sisters, and

- The first two types of tetrads were combined together into Mother-Child-Aunt-Nibling (MCAN) included the firstborn child from each of the Generation 1 sisters. (Note: "Nibling" refers to a niece or nephew with unspecified gender; compare to "Sibling.")

Measures

Generation 1 AFI.   NLSY-79 subjects indicated their AFI a maximum of three times, in response to questions in 1983, 1984, and 1985. The 1984 and 1985 questions were included to assess those with non-response in 1983, but in fact many female respondents were surveyed multiple times. Further, females were asked for additional related information (Year of First Intercourse, Month of First Intercourse) in 1984 and 1985. The average AFI for women was 18.52 (sd = 2.12; n = 5562); men was 17.16 (sd = 2.36; n = 5640). Subjects with missing AFI data were excluded from analyses.

We used the repeated questions to estimate the test-retest reliability of self-reported AFI and AFI difference scores. In Table 1, 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 the upper triangle indicates the number of respondents that reported AFI for both respective years. The test-retest correlations are moderate to high (r > .75) across all viable pairings, suggesting that our subjects are reliably reporting AFI.

Table 1

Gen1 Self-Reported AFI Correlations and Sample Sizes (1983-1985)

	1983	1984	1985
1983	8432	3765	88
1984	0.86	4516	0
1985	0.76	NA	424

Generation 2 AFI.   Over the lifetime of the NLSY-Children survey, participants were asked approximately the same questions to assess AFI that their mothers were asked. However, Generation 2 respondents were only asked for AFI information once they had reached age 15 or later. 7.1% (812) of the sample had not reached age 15. The exact nature of the question varied slightly 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.

We calculated NLSY-Children AFI, using a multi-step process for three reasons: (1) to account for the diversity of AFI questions across survey administrations, (2) to incorporate multiple reports by the same participant, which occasionally differed, and (3) to account for the imprecision of AFI reporting (*e.g.*, a subject who reports AFI at 16 could be any age between 16 years and 0 days old through 16 years and 364 days old).

For each survey, we transformed year of first intercourse into an age variable, AFI. If subjects reported both age and year within the same survey and ages were different, we averaged the AFI scores. Across surveys, we identified the earliest possible AFI and the latest possible AFI for each subject. We designated these two AFIs as the Minimum AFI and the Maximum AFI respectively, thus identifying the full range of possible AFIs for each participant.[3] We used this AFI range to calculate the expected value of AFI by averaging the Maximum and Minimum AFI. Using this method, the average Generation

---

[3]We added 1 year to the Maximum AFI to address the imprecision of self-reported age. The expected value of AFI of any subject does not equal the reported AFI. For example, a subject who reports AFI at 16 could be anywhere from 16 years and 0 days old to 16 years and 364 days old.

2 AFI was 16.01 (sd = 2.30; n = 6288).[4]

   After transforming all AFI scores, we recoded impossible AFIs as missing. A score was defined as impossible if the reported AFI exceeded participant's age at time of survey (new $\overline{\text{AFI}}$ = 15.99, sd = 2.30, n = 6235). Next, we excluded all AFIs below age 12 (new $\overline{\text{AFI}}$ = 16.14, sd = 2.10, n = 6087). Finally we excluded subjects who reported AFI prior to their self-reported menarche (new $\overline{\text{AFI}}$ = 16.16, sd = 2.09, n = 6047). We excluded those below age 12 because those responses likely are the result of misunderstanding, non-consensual sexual activity, or other forms of unreliability. We excluded those with pre-menarchal AFI because we focus on AFI that could potentially link to reproduction and fertility. Subjects with missing data were excluded from analyses. 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) of any race-gender categories. For a complete portrayal of summary statistics, see Table 2 and Figure 1.

---

[4]Taking the average of all AFIs (without addressing expected value), results in 15.49 (sd = 2.30; n = 6288). Adding in expected value of .5 changes this value to 15.99, approximately our computed age.

Table 2

Gen2 Mean AFI by Gender, Race, and Gender by Race

	AFI	
	Mean	Sd
Overall	16.16	2.097
Male	15.88	2.152
Female	16.47	1.991
Hispanic	16.22	2.140
Black	15.66	2.012
Non-Black, Non-Hispanic	16.57	2.054
Hispanic Male	15.92	2.163
Black Male	15.04	1.958
Non-Black, Non-Hispanic Male	16.54	2.061
Hispanic Female	16.60	2.050
Black Female	16.26	1.877
Non-Black, Non-Hispanic Female	16.61	2.048

Figure 1. Smoothed Density Plots of Gen1 AFI by Race and Sex
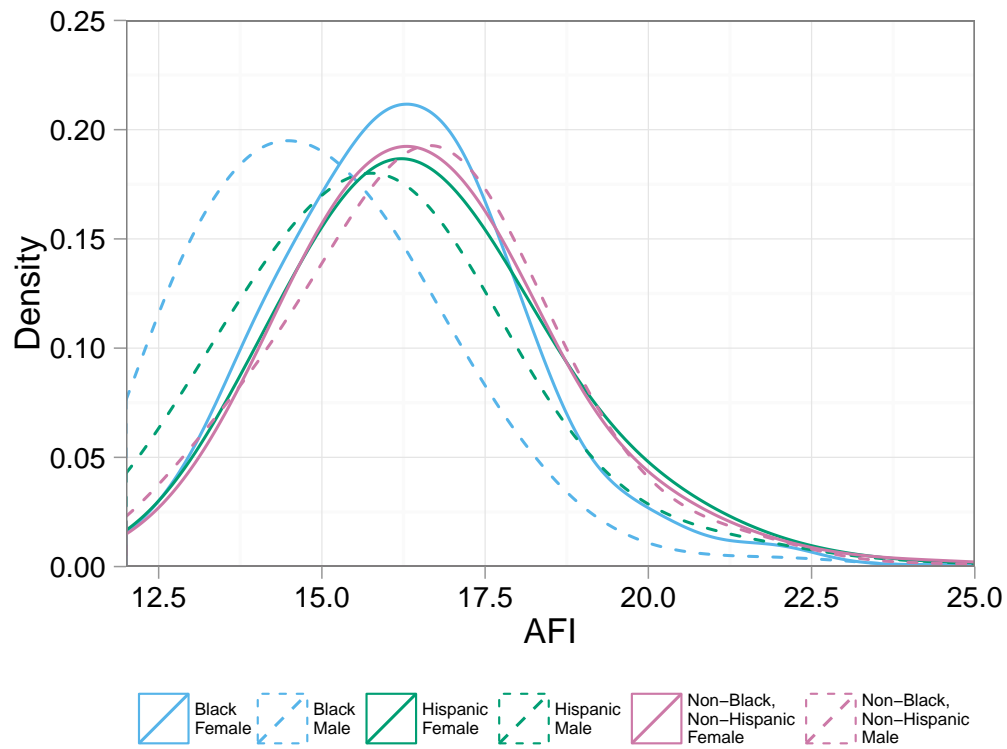
Generation 1 Intelligence. The Armed Services Vocational Aptitude Battery (ASVAB; Form 8A; Palmer et al., 1988) was administered to Gen1 participants in 1980. The Armed Forces Qualification Test (AFQT) is contained within the ASVAB, and has been used in the US military as a measure of general trainability (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). Other administrations of the pencil and paper ASVAB reveal that all the AFQT subscales have high coefficient α internal consistency ($\alpha_{AR} = .91$; $\alpha_{WK} = .92$; $\alpha_{PC} = .81$; $\alpha_{MK} = .87$; Kass, Mitchell, Grafton, & Wing, 1982). Reported reliability of the overall AFQT (version 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:

$$AFQT = AR + MK + 2VE, \qquad (4)$$

$$\text{where, } VE = PC + WK. \qquad (5)$$

Many researchers have used the AFQT standard score as a proxy for general intelligence (g) (Herrnstein & Murray, 1994; Der, Batty, & Deary, 2009). Indeed, the U.S. 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, & Ree, 1984; J. J. Mathews, 1977).

Generation 2 Intelligence. NLSY-Children respondents, beginning at age five and as a consistent part of the survey, complete 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

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

The standard scores of the PPVT-R, PIATs, and Digit Span are considered valid and reliable assessments of cognitive ability (Mott & Baker, 1995). However, subjects were surveyed on a biennial basis. Thus we could not use cognitive tests at a fixed age. Instead, we aggregated scores across a 4 year window, and targeted the midpoint between 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 by measuring intelligence prior to the occurrence of AFI. In the case of missing subtests, we allowed age 11 standard scores to replace age 9 standard scores, and age 8 standard scores to replace age 10 standard scores. Our replacement strategy ensured that the average age of testing matched the average of our targeted ages. To obtain intellectual ability measures for each NLSY-children respondent, we fit a confirmatory factor analysis model (using Mplus; Muthén & Muthén, 2014) and their robust maximum likelihood estimator option. A single-factor model fit moderately well (RMSEA = .101; CFI = .973; TLI = .946), and we used this model to construct a unidimensional scale score for each respondent. We used factor scores obtained from this model as our measure of NLSY-Children intelligence.

Replicability and Reliability. 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. The test-retest reliabilities of Gen2 intelligence across our three aggregations is reported in the lower triangle of Table 3. 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. The test-retest correlations are very high ($r >$ .90) across all pairings, suggesting that our method captures consistent measures of intelligence across ages.

Table 3

Gen2 Aggregated Intelligence Correlations and Sample Sizes (Ages 9.5, 10.5, 11.5)

	Age 9.5	Age 10.5	Age 11.5
Age 9.5	8254	7974	7669
Age 10.5	0.95	8143	7838
Age 11.5	0.90	0.96	7970

Reliability of Difference Scores

Our design assumes that the difference scores of our measures are reliable. We have reported the test-retest reliability of Gen2 intelligence and Gen1 AFI in earlier sections. Here, we report the test-retest reliability of the pairwise differences of those measures.

AFI. Comparing sibling differences in AFI as reported in 1983 and 1984 (n = 783 pairs) we found a moderate correlation (r = 0.76). The sample of sibling pairs with complete information in 1985 was too small (n = 12 pairs) to compare to the other two years. Regardless, sibling differences in self-reported AFI appear satisfactory for research purposes. We note that the common concern with unreliability of difference scores is at least partially mitigated when the separate scores defining the differences are reliable themselves. Because we could not estimate test-retest reliabilities for Generation 2, we calculated the reliability using Lord's (1963) equation. Generation 2 Mean AFI difference scores were also moderately reliable (r = 0.73) and comparable to Generation 1 sibling differences.

Intelligence. Cousin differences in intelligence as assessed at ages 9.5, 10.5, and 11.5 were correlated using three different linking methods (Mixed, Daughters, Sons). Reliabilities across linking methods was consistent and high (min r = .86; max r = .95). However, we could not estimate test-retest reliabilities for Generation 1; we calculated the reliability using Lord's (1963) equation. The calculated reliability of Generation 1's differences in AFQT was 0.70, which was marginally acceptable, and obviously lower than the empirical correlation we derived for cousin differences.

Results

We examined the relationship between AFI and intelligence using two designs: a between-family design, and a combination between- and within-family design (which includes between-family variance in the differences between the family means, and within-family variance in the sibling/cousin differences). The results are organized by those two designs. The between-family analyses report the relationships between the

within-family average AFI and various measures of ability. The combination between-family/within-family analyses add to the between-family analyses the difference scores, testing whether differences in AFI can be explained by differences in various measures of ability, controlling for the between-family variance.

Between-Family Analyses

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 intelligence and of Gen2 intelligence. We evaluated the influences both independently and simultaneously. All intelligence 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. In Tables 4 - 6, we have reported results for three different linking methods:

- The mixed model, which contains the firstborn child of each sister;

- The daughters model, which contains the firstborn daughters; and

- The sons model, which contains the firstborn sons).

In the spirit of transparency, we have reported all three methods. However, because all three linking methods reported similar findings, we will focus the results section on the mixed model and only discuss the other two methods when they deviate. We have also provided zero-order and pairwise semi-partial correlations for all between-family variables (Tables C1 and C2), using the mixed-model data and the ppcor R library (Kim, 2015).

Gen1 Mean Intelligence → Gen2 Mean AFI. Gen1 sister averages (NLSY79 mothers) of standardized AFQT scores were used to predict Gen2 averages of gender-standardized AFI. Table 4 displays the results by Gen2 category. The mixed model reports the averages of the firstborn child (both males and females) of each maternal sister (n = 342). A one-unit increase in the average standardized intelligence of the children's mothers predicted a statistically significant increase of .013 standard

deviations in average Gen2 AFI. When we transform the coefficient into a standardized beta weight ($\beta_{Gen1Intell} = .299$), a standard-deviation increase in the averaged mother's intelligence predicts a .299 standard-deviation increase in the average of the cousins' AFI. The adjusted R^2 was .087.

Gen2 Mean Intelligence → Gen2 Mean AFI. Gen2 averages of standardized intelligence scores were used to predict Gen2 averages of gender-standardized AFI. Table 5 displays the results by Gen2 category. The mixed model reports the averages of the firstborns of each of the NLSY79 mothers (sisters) (n = 344). A one-unit increase in the average standardized intelligence of the children predicted a statistically significant $\approx .075$ standard-deviation increase in average Gen2 cousins' AFI. When we transform the coefficient into a standardized beta weight ($\beta_{Gen2Intell} = .128$), a standard-deviation increase in the averaged cousins' intelligence predicts a .128 standard-deviation increase in the average of the cousins' AFI. The adjusted R^2 was .014.

Joint Mean Intelligence → Gen2 Mean AFI. Results from the Gen1 maternal sister averages of standardized AFQT scores and Gen2 averages of standardized intelligence scores predicting Gen2 averages of gender-standardized AFI are displayed in Table 6. In the mixed model, Gen1 (maternal) intelligence was significantly associated with Gen2 AFI (p < .01), while Gen2 (child) intelligence was not significantly associated with Gen2 AFI. A one-unit increase in the average standardized intelligence of the children's mothers predicted .013 standard-deviation increase in average Gen2 AFI, after controlling for Gen2 cousin averages of standardized intelligence scores. When we transform the coefficients into standardized beta weights ($\beta_{Gen1Intell} = .303$; $\beta_{Gen2Intell} = -.003$), a standard-deviation increase in the averaged mother's intelligence predicts a .303 standard-deviation increase in the average of the cousins' AFI. The adjusted R^2 was .086. The total variance explained by the Joint model ($R^2 = 9.1\%$) is nearly identical to the Gen1 model($R^2=9\%$). Gen2 intelligence explains an additional .1% of the variance.

When we broaden our sample to all Mother-Child pairs, we see that the

relationship between Gen 2 intelligence and Gen2 AFI is small ($r = .139$), and smaller than the relationship between Gen1 intelligence and Gen2 AFI ($r = .215$; see Figure 2).

Table 4

Between-Family: Gen1 Intelligence Predicts Gen2 AFI

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen1 Mean Intel	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)
Constant	−0.75*** (−1.01, −0.50)	−0.80*** (−1.09, −0.51)	−0.86*** (−1.14, −0.57)
Sample Size	342	264	282
R^2	0.09	0.10	0.11
Adjusted R^2	0.09	0.10	0.10
Residual Std. Error	0.70 (df = 340)	0.70 (df = 262)	0.69 (df = 280)
F Statistic	33.50*** (df = 1; 340)	29.40*** (df = 1; 262)	33.20*** (df = 1; 280)

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

Gen1-sister averages (NLSY79 mothers) of standardized AFQT scores predict Gen2 averages of gender-standardized AFI.

Table 5

Between-Family: Gen2 Intelligence Predicts Gen2 AFI

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean Intel	0.08** (0.01, 0.14)	0.09** (0.01, 0.16)	0.07* (−0.004, 0.14)
Constant	−0.02 (−0.10, 0.06)	−0.02 (−0.11, 0.07)	−0.04 (−0.12, 0.05)
Sample Size	344	267	283
R^2	0.02	0.02	0.01
Adjusted R^2	0.01	0.02	0.01
Residual Std. Error	0.74 (df = 342)	0.74 (df = 265)	0.73 (df = 281)
F Statistic	5.70** (df = 1; 342)	5.32** (df = 1; 265)	3.43* (df = 1; 281)

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

Gen2-cousin averages of standardized intelligence scores predict Gen2 averages of gender-standardized AFI.

Table 6

*Between-Family: Gen1 & Gen2 Intelligence Predict Gen2 AFI*

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen1 Mean Intel	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)
Gen2 Mean Intel	−0.002 (−0.07, 0.07)	−0.000 (−0.08, 0.08)	−0.01 (−0.09, 0.06)
Constant	−0.77*** (−1.05, −0.48)	−0.81*** (−1.13, −0.50)	−0.89*** (−1.22, −0.57)
Sample Size	337	260	278
R^2	0.09	0.10	0.11
Adjusted R^2	0.09	0.10	0.10
Residual Std. Error	0.70 (df = 334)	0.71 (df = 257)	0.69 (df = 275)
F Statistic	16.70*** (df = 2; 334)	14.90*** (df = 2; 257)	16.50*** (df = 2; 275)

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

Gen1-sister averages (NLSY79 mothers) of standardized AFQT scores and Gen2-cousin averages
of standardized intelligence scores predict Gen2 averages of gender-standardized AFI.

Figure 2. Gen2 AFI vs. Gen1 & Gen2 Intelligence
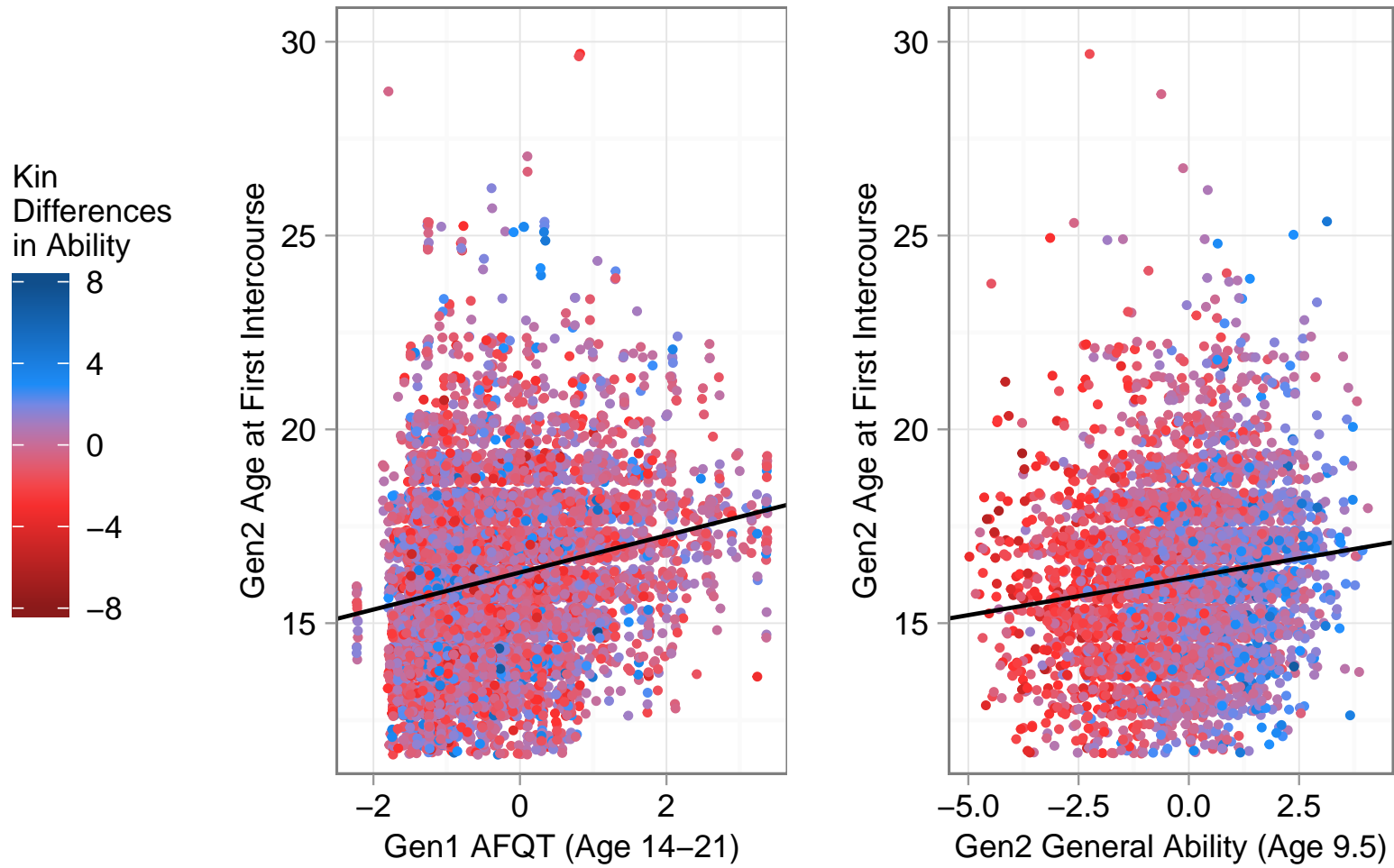
Combination Between- and Within-Family Analyses

We replicated the between-family analyses reported in the previous subsection, adding within-family difference scores as independent variables to the between-family means. Using the discordant sibling model, we predicted the differences in Generation 2 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 Generation 2 linking method variants, just as we did in the between family analyses: the mixed model reports the differences of the firstborns of each sister, the daughters model reports the differences of the firstborn girls, and the sons model reports the differences of the firstborn sons. In the spirit of transparency, we have reported all three methods. However, because all three linking methods resulted in similar findings, we will focus the results section on the mixed model and only discuss the other two methods when they deviate. We have also provided zero-order and pairwise semi-partial correlations for all between-family variables (Tables C1 and C3), using the mixed-model data and the ppcor R library (Kim, 2015).

Gen1 Intelligence Differences → Gen2 AFI Differences. Generation 1 maternal sister differences in standardized AFQT scores were used to predict Generation 2 (cousin) differences of gender-standardized AFI, controlling for Generation 1 sister averages of standardized AFQT scores and Generation 2 averages of gender-standardized AFI. Table 7 displays the results by Generation 2 linking method. The mixed model reports the averages and differences of the firstborns of each sister (n = 336). Generation 2 averages of gender-standardized AFI (between-family measures) were significant predictors of Generation 2 differences in gender-standardized AFI (p < .01). A one-unit increase in the average gender-standardized AFI predicted 0.303 standard-deviation increase in average Gen2 AFI difference, controlling for all other variables in the model. When we transform the coefficients into standardized beta weights ($\beta_{Gen2MeanAFI} = .283$; $\beta_{Gen1MeanIntell} = -.057$; $\beta_{Gen1DiffIntell} = .03$), a standard-deviation increase in the averaged cousin's AFI predicts a .283

standard-deviation increase in the difference between the cousins' AFI. The adjusted R^2 was .066. The Gen1 intelligence difference variable was not statistically significant.

In the sons model, the Generation 1 maternal 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 .0083 decrease in the AFI difference between cousins. When we transform the coefficients into standardized beta weights ($\beta_{Gen2MeanAFI} = .353$; $\beta_{Gen1MeanIntell} = -.174$; $\beta_{Gen1DiffIntell} = .006$), a standard-deviation increase in the averaged cousin's AFI predicts a .353 standard-deviation increase in the difference between the cousins' AFI, while a standard-deviation increase in the averaged mother's AFQT predicts a .174 standard-deviation decrease in the difference between the cousins' AFI. All other variables were not significant, including all kin-difference variables (the within-family measures). Note that this between-family result is somewhat anomalous, because it is in the opposite direction to the other results. The adjusted R^2 was = .106.

In the sons model for the age 10.5 and 11.5 replications (see Appendices A and B), the Generation 1 cousin average of standardized intelligence scores was a significant predictor of differences in Generation 2 AFI (p < .05; see Tables A4 & B4). The adjusted R^2s were similar.

Gen2 Intelligence Differences → Gen2 AFI Differences.   Gen2 cousin differences in standardized intelligence scores were used to predict Gen2 differences of gender-standardized AFI, controlling for Gen2 cousin averages of standardized intelligence scores and gender-standardized AFI (to account for between-family variance). Table 8 displays the results by Generation 2 categories. The mixed model reports the averages and differences of the firstborns of each sister (n = 291). Generation 2 averages of gender-standardized AFI were significant predictors of Generation 2 differences in gender-standardized AFI (p < .01). A one-unit increase in the average gender-standardized AFI predicted 0.357 standard-deviation increase in average Gen2 AFI difference, controlling for all other variables in the model. The adjusted R^2 was = .103. When we transform the coefficients into standardized beta

weights ($\beta_{Gen2MeanAFI} = .337$; $\beta_{Gen2MeanIntell} = -.091$; $\beta_{Gen2DiffIntell} = .068$), a standard-deviation increase in the averaged cousin's AFI predicts a .337 standard-deviation increase in the difference between the cousins' AFI. The Gen2 intelligence difference variable was not statistically significant.

In the sons model, the Generation 2 cousin average of standardized intelligence scores was a significant predictor of differences in Generation 2 AFI (p < .05). A one-unit increase in the average standardized intelligence of the children predicted a .107 decrease in the AFI difference between cousins(again, the sons model result is in the opposite direction to other results). When we transform the coefficients into standardized beta weights ($\beta_{Gen2MeanAFI} = .372$; $\beta_{Gen2MeanIntell} = -.15$; $\beta_{Gen2DiffIntell} = .021$), a standard-deviation increase in the averaged cousin's AFI predicts a .372 standard-deviation increase in the difference between the cousins' AFI, whereas a standard-deviation increase in the averaged cousin's intelligence predicts a .15 standard-deviation decrease in the difference between the cousins' AFI. All other variables were not significant, including all kin-difference variables. The adjusted R^2 was .132).

In the mixed, daughters, and sons models for the age 10.5 and 11.5 replications (see Appendices A and B), the Generation 2 cousin average of standardized intelligence scores were significant predictors of differences in Generation 2 AFI (p < .05; see Tables A5 & B5). Regardless, all kin-difference variables were not significant. The adjusted R^2s were similar.

Moreover, when we replicate this analysis using the larger Gen1 sample of siblings, we find that the between-family relationship of Gen1 AFI and Gen1 intelligence is moderate ($r = .291$), but within the family, the AFI of the smarter sibling was no later than the less smart sibling's. In Figure 3, we have illustrated this finding using a scatter plot for the individual subjects, and two marginal distributions to show the mean levels of AFQT and AFI, grouped by sibling classification. The distribution above the x-axis shows the distributions of AFQT based on sibling classification. The blue distribution shows the standardized AFQT score for the siblings who were .33 standard-deviations

smarter than their sibling, the red distribution shows the standardized AFQT score of siblings who were at least .33 standard deviations less smart than their sibling, and the purple distribution shows those siblings who were within at least .33 standard deviations of one another. As expected, the blue distribution of smarter siblings had a higher mean than the other groups. The distribution across from the y-axis shows the distribution of AFI based on those same sibling classifications. All three of these distributions are indistinguishable, which suggests that siblings do not differ in their AFI across these intelligence categories. This plot illustrate that there is no within-family effect for intelligence on AFI. Had there been a within-family effect of intelligence on AFI, the marginal distributions of AFI would be visually separable, and would follow the pattern for the separate distributions of AFQT.

Joint Intelligence Differences → Gen2 AFI Differences. Gen1 maternal sister differences in standardized AFQT scores and Gen2 cousin differences in standardized intelligence scores were used simultaneously to predict Gen2 differences of gender-standardized AFI, controlling for Gen1-sister averages of standardized AFQT scores, Gen2-cousin averages of standardized intelligence scores, and Gen2-cousin averages of gender-standardized AFI. Table 9 displays the results by Generation 2 categories. The mixed model reports the averages and differences of the firstborns of each sister (n = 285). Gen2 averages of gender-standardized AFI were significant predictors of Generation 2 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 standard-deviation increase in Generation 2 AFI difference, controlling for all other variables in the model. All other variables were not significant, including all kin difference variables. When we transform the coefficients into standardized beta weights ($\beta_{Gen2MeanAFI} = .324$; $\beta_{Gen2MeanIntell} = -.091$; $\beta_{Gen1MeanIntell} = .017$; $\beta_{Gen2DiffIntell} = .059$; $\beta_{Gen1DiffIntell} = .02$), a standard-deviation increase in the averaged cousin's AFI predicts a .345 standard-deviation increase in the difference between the cousins' AFI. The adjusted R^2 was .090.

Table 7

Within-Family: Gen1 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Differences in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.30*** (0.19, 0.42)	0.33*** (0.19, 0.46)	0.38*** (0.25, 0.50)
Gen1 Mean Intel	−0.003 (−0.01, 0.002)	−0.002 (−0.01, 0.004)	−0.01*** (−0.01, −0.003)
Gen1 Dif Intel	0.001 (−0.003, 0.01)	0.001 (−0.01, 0.01)	0.000 (−0.005, 0.01)
Constant	1.18*** (0.89, 1.47)	1.18*** (0.84, 1.52)	1.50*** (1.18, 1.83)
Sample Size	336	258	278
R^2	0.07	0.08	0.12
Adjusted R^2	0.07	0.07	0.11
Residual Std. Error	0.76 (df = 332)	0.78 (df = 254)	0.73 (df = 274)
F Statistic	8.83*** (df = 3; 332)	7.64*** (df = 3; 254)	12.00*** (df = 3; 274)
Notes:			*p<0.1; **p<0.05; ***p<0.01

Gen1-sister differences in standardized AFQT scores predict Gen2 differences in gender-standardized AFI.

Table 8

Within-Family: Gen2 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Differences in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.36*** (0.24, 0.47)	0.40*** (0.26, 0.53)	0.40*** (0.27, 0.53)
Gen2 Mean Intel	−0.06 (−0.14, 0.01)	−0.07 (−0.16, 0.02)	−0.11** (−0.19, −0.02)
Gen2 Dif Intel	0.03 (−0.02, 0.08)	0.04 (−0.02, 0.09)	0.01 (−0.04, 0.06)
Constant	1.05*** (0.96, 1.14)	1.07*** (0.97, 1.18)	1.07*** (0.97, 1.16)
Sample Size	291	223	238
R^2	0.11	0.13	0.14
Adjusted R^2	0.10	0.12	0.13
Residual Std. Error	0.77 (df = 287)	0.79 (df = 219)	0.75 (df = 234)
F Statistic	12.10*** (df = 3; 287)	11.20*** (df = 3; 219)	13.00*** (df = 3; 234)

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

Gen2-cousin differences in standardized intelligence scores predict Gen2 differences in gender-standardized AFI.

Table 9

*Within-Family: Gen1 & Gen2 Differences in Intelligence Predict Gen2 Differences in AFI*

	Dependent variable: Differences in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.34*** (0.22, 0.47)	0.38*** (0.23, 0.52)	0.43*** (0.29, 0.56)
Gen2 Mean Intel	−0.06 (−0.15, 0.02)	−0.07 (−0.17, 0.03)	−0.07 (−0.16, 0.03)
Gen1 Mean Intel	0.001 (−0.01, 0.01)	0.001 (−0.01, 0.01)	−0.01 (−0.01, 0.001)
Gen2 Dif Intel	0.03 (−0.02, 0.07)	0.03 (−0.02, 0.09)	0.01 (−0.04, 0.06)
Gen1 Dif Intel	0.001 (−0.004, 0.01)	−0.001 (−0.01, 0.01)	0.001 (−0.005, 0.01)
Constant	1.00*** (0.64, 1.36)	1.01*** (0.59, 1.43)	1.37*** (0.97, 1.77)
Sample Size	285	217	235
R^2	0.11	0.13	0.15
Adjusted R^2	0.09	0.10	0.13
Residual Std. Error	0.77 (df = 279)	0.79 (df = 211)	0.75 (df = 229)
F Statistic	6.61*** (df = 5; 279)	6.04*** (df = 5; 211)	8.08*** (df = 5; 229)

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

Gen1-sister differences in standardized AFQT scores and Gen2-cousin differences in standardized intelligence scores predict Gen2 differences in gender-standardized AFI.

Figure 3. Gen1 AFI vs. Gen1 Intelligence with Marginal Distributions by Sibling Difference in AFQT
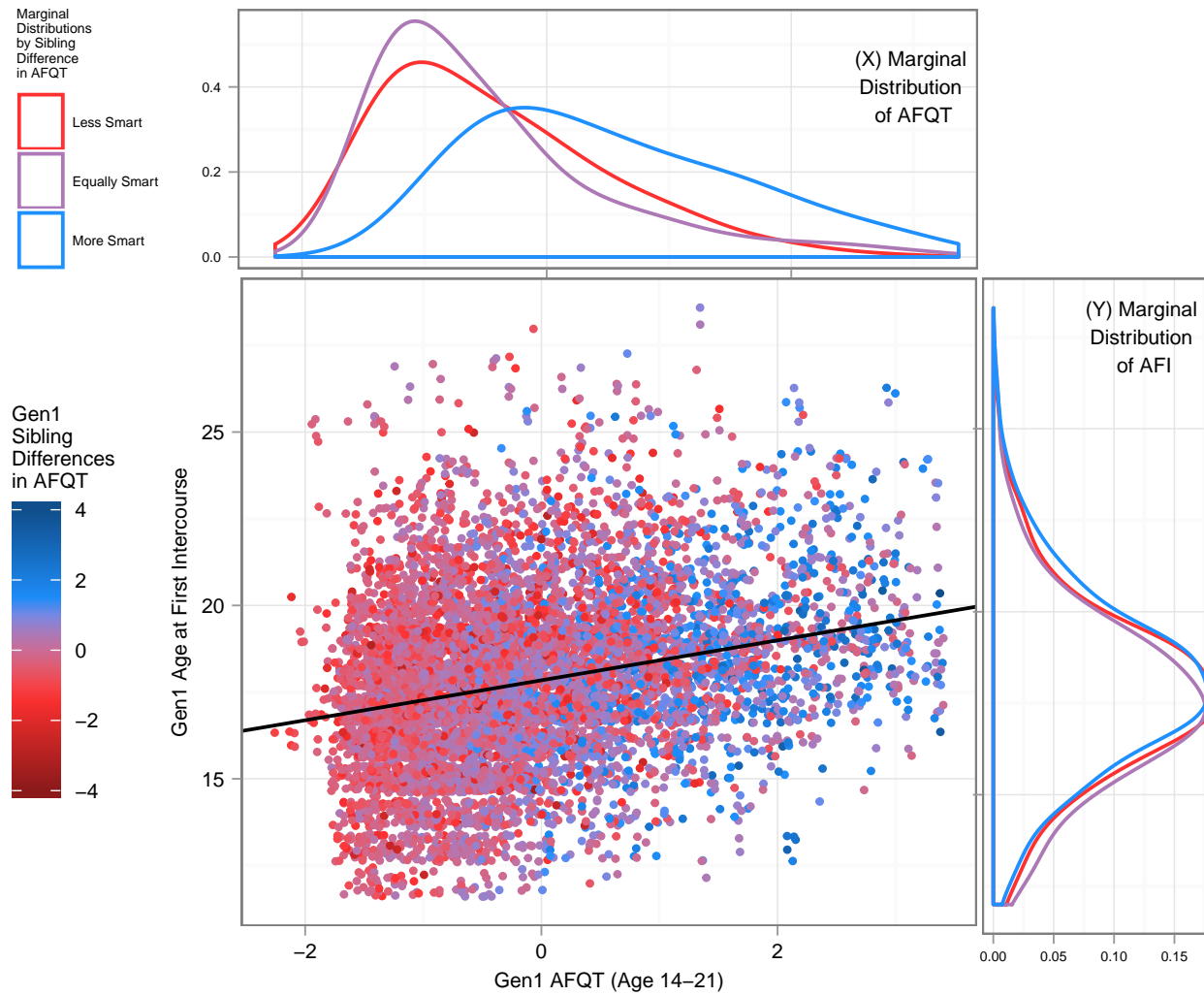
Discussion

This article presents analysis of the relationship between AFI and intelligence using two different designs: a between-family design, and a combination between- and within-family design (*i.e.*, a within-family design includes between-family variance within it). The between-family design allowed us to replicate results obtained by previous researchers who used cross-sectional samples. The combination between- and within-family design allowed us to separately account for within- and between-family variance, to determine the source of the explanatory processes. The logic of this separation allows us to get much closer to evaluating intelligence differences within the family to address issues of causality. The results revealed a stark contrast between the two methods, and cast doubt on the validity of past causal assertions.

Between- vs. Within-Family Variance

Between-Family Results.   Notably, the between-family analyses showed a relationship between intelligence and AFI. Thus, we were able to replicate the findings of various researchers (Halpern et al., 2000; Mott, 1983; Paul et al., 2000; Woodward et al., 2001), and confirm hypotheses 1 and 2. Our within-generation findings[5] were comparable in effect size ($d_{Gen1} = .554$; $d_{Gen2} = .281$) to Halpern *et al.*'s 2000 finding ($d_{Halpern} = .542$). The effects were small to medium in size. The effect for Gen1 was more similar to Halpern et al. (2000)'s finding, likely because both findings used intelligence assessed at a later age (≈ 16 years), compared to the Gen2 assessment at age 9.5.

The relationship between AFI and intelligence was substantially stronger between maternal intelligence and child AFI than between the child's own intelligence and child AFI, suggesting that family-level variables rather than individual-level intelligence is the likely source of the relationship. If the child's own intelligence had been the primary causal influence on AFI we would have expected a considerably weaker

---

[5]We report within-generation between-family effects for intelligence correlated with AFI (*e.g.* Gen1 intelligence correlated with Gen1 AFI)

cross-generational association between AFI and intelligence. Instead we find that the within-generation association is the weaker effect, suggesting that the child's own AFI is likely derivative of the child's mother's intelligence (or other between-family correlates), which are the more likely causal influence. Thus, the "new" and alternative interpretation of this finding would be that maternal intelligence or correlates are driving the effect, and that past between-family analyses finding a link between child's intelligence and AFI are likely because child's intelligence is indirectly measuring maternal intelligence and other between-family correlates.

There are interpretations that would support the plausibility of this result, including maternal intelligence as a direct causal influence on their children's AFI. Smarter mothers might be more effective at encouraging their children to delay intercourse – perhaps by effectively conveying the riskiness of sexual intercourse (Hutchinson et al., 2003; C. Mathews et al., 2009). Considering that intelligence is moderately-to-highly heritable (Bouchard, 2004) and thus highly correlated across generations, this alternative explanation would still be consistent with the traditional between family findings, which typically do not control for maternal intelligence (Halpern et al., 2000; Mott, 1983; Paul et al., 2000; Woodward et al., 2001). We note that the Harden & Mendle (2011) findings, using biometrically-informed data, implicated the shared environment – a between-family source of variance – in this causal process. Our results are entirely consistent with theirs, using a different dataset and a different methodological approach to identify important sources of variance.

Within-Family Results.   In the within-family analyses, the relationship between intelligence and AFI vanishes for both maternal intelligence and child intelligence. The child of the smarter Generation 1 mother was not more likely to delay intercourse compared to the child of the less-smart Generation 1 mother. In spite of our finding that Generation 1 intelligence was a relatively stronger predictor of Generation 2 AFI than Generation 2 intelligence, we did not find that within family differences in Generation 1 intelligence were associated with differences in Generation 2 AFI.

These results cast doubt on the alternative explanation for the between-family

results we posed in the previous section. If Gen1-maternal intelligence was driving the effect as a proximal cause, we would have expected to find a significant within-family link from maternal intelligence to child AFI, which we did not. Thus, within our sample, it appears that intelligence is not the proximal cause of delayed AFI. Rather, maternal and child intelligence appear to be are indirect measures of many other between-family 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 the home environment, which could be measured as a composite of parental income, intelligence, education, family interaction, *etc.*

Concluding Remarks

We interpret these results in relation to two previous findings. First, Rodgers et al. (2008) used Danish twin data, and found that the link from education/cognitive ability to maternal age at first birth (AFB) was entirely accounted for by between-family variance: "variance in AFB emerges from [IQ and education] differences between families, not differences between sisters within the same family" (Rodgers et al., 2008, p. 202). We have exactly the same type of result in the current study. Second, Harden and Mendle's (2011) results, obtained from the Add Health data, use intelligence as a predictor and AFI as an outcome, just as we did with the NLSY. Their biometrical finding of meaningful shared environmental variance is consistent with our finding of only between-family variance. Their design identifies that the source of the covariation between AFI and intelligence is in the shared environment.

Our findings cast further doubt on the direct and causal influence of intelligence. By explicitly parsing between- and within-family variance, we tested the causal link between AFI and intelligence. We employed numerous replications. We varied when we measured intelligence (Age 9.5, 10.5, 11.5; Age 16); how we measured intelligence (composite of five measures; AFQT from the ASVAB); which cohort we used (NLSYC; NLSY79); how we created kinship pairs (firstborn, first female, first male); and how

related our kin-pairs were (cousins; siblings). We examined and eliminated both maternal intelligence and child intelligence as having any within-family causal etiology.

Given our findings and those of Harden & Mendle (2011), we find no evidence for intelligence being a direct causal influence on AFI. Instead, we direct future researchers to look at the general family environment,"little e," and other between-family factors correlated with maternal intelligence as likely causes of AFI. Although "smart teens don't have sex (or kiss much either)" (Halpern et al., 2000) at a descriptive level, their reasons for delaying these activities do not appear to be caused by their smartness.

References

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

Barclay, K. J. (2015). A within-family analysis of birth order and intelligence using population conscription data on Swedish men. *Intelligence*, *49*, 134–143.

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.

Bjerkedal, T., Kristensen, P., Skjeret, G. A., & Brevik, J. I. (2007). Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families. *Intelligence*, *35*(5), 503–514.

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.

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, T. J. (2004). Genetic influence on human psychological traits: A survey. *Current Directions in Psychological Science*, *13*(4), 148–151.

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

Brooks-Gunn, J., & Furstenberg, F. F. (1989). Adolescent sexual behavior. *American Psychologist*, *44*(2), 249–257.

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.

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.

Der, G., Batty, G. D., & Deary, I. J. (2009). 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.

D'Onofrio, B. M., & Lahey, B. B. (2010). Biosocial influences on the family: A decade review. *Journal of Marriage and Family*, *72*(3), 762–782.

D'Onofrio, B. M., Lahey, B. B., Turkheimer, E., & Lichtenstein, P. (2013). Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. *American Journal of Public Health*, *103 Suppl*(SUPPL.1), S46–55.

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.

East, P. L. (1996). The younger sisters of childbearing adolescents: Their attitudes, expectations, and behaviors. *Child Development*, *67*(2), 267–282.

East, P. L., Felice, M. E., & Morgan, M. C. (1993). Sisters' and girlfriends' sexual and childbearing behavior: Effects on early adolescent girls' sexual outcomes. *Journal of Marriage and Family*, *55*(4), 953.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, *196*(4286), 129–136.

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

Galbraith, R. C. (1982). Sibling spacing and intellectual development: A closer look at the confluence models. *Developmental Psychology*, *18*(2), 151–173.

Geronimus, A. T., & Korenman, S. (1992). The socioeconomic consequences of teen childbearing reconsidered. *Quarterly Journal of Economics*, *107*(4), 1187–1214.

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.

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

Hirschi, T. (2002). *Causes of delinquency.* Piscataway, NJ: 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.

Kann, L., Kinchen, S., Shanklin, S. L., Flint, K. H., Hawkins, J., Harris, W. A., . . . Zaza, S. (2014). *Youth risk behavior surveillance – United States, 2013* (Vol. 63; Tech. Rep. No. 4).

Kass, R. A., Mitchell, K. J., Grafton, F. C., & Wing, H. (1982). *Factor structure 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.

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.

Kim, S. (2015). ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for statistical applications and methods*, *22*(6), 665.

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., & Gebhard, P. H. (1953). *Sexual behavior in the human male.* Philadelphia: W. B. Saunders Company.

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.). Alexandria, VA: Center for Naval Analyses.

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.). Brooks Air Force Base, Tx: Air Force Human Resources Lab.

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. Paper presented at the *America Public Health Association Meeting*, Dallas, TX.

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: Statistical analyses with latent variables. User's guide version 7.3.*

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. *American Psychologist*, *51*(2), 77–101.

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, L. D., & 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: DTIC Document.

Paul, C., Fitzjohn, J., Herbison, P., & Dickson, N. (2000). 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. In R. M. Lerner & T. T. Foch (Eds.), *Biological-psychosocial interactions in early adolescence: Child psychology.* (pp. 35–61). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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. (1984). Confluence effects: Not here, not now! *Developmental Psychology*, *20*(2), 321.

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. (2001). What causes birth order-intelligence patterns? The admixture hypothesis, revived. *American Psychologist*, *56*(6-7), 505–510.

Rodgers, J. L. (2014). Are birth order effects on intelligence really Flynn Effects? Reinterpreting Belmont and Marolla 40 years later. *Intelligence*, *42*(1), 128–133.

Rodgers, J. L., Beasley, W. H., Bard, D. E., Meredith, K. M., Hunter, M., Johnson, A. B., . . . Rowe, D. C. (2016). The NLSYKinship Links: Using the NLSY79 and NLSY-Children data to conduct genetically-informed and family-oriented research. *Behavior Genetics*, *46*(4), 538–551.

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.

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. (1990). Adolescent sexual activity and mildly deviant behavior: Sibling and friendship effects. *Journal of Family Issues*, *11*(3), 274–293.

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.

Rodgers, J. L., Rowe, D. C., & Harris, D. F. (1992). Sibling differences in adolescent sexual behavior: Inferring process models from family composition patterns. *Journal of Marriage and Family*, *54*(1), 142–152.

Rodgers, J. L., Rowe, D. C., & May, K. O. (1994). DF analysis of NLSY IQ/achievement data: Nonshared environmental influences. *Intelligence*, *19*(2), 157–177.

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

Santelli, J. S., Lindberg, L. D., Abma, J., McNeely, C. S., & Resnick, M. (2000). Adolescent sexual behavior: Estimates and trends from four nationally representative surveys. *Family Planning Perspectives*, 156–194.

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.

Udry, J. R. (1994). The nature of gender. *Demography*, *31*(4), 561–573.

Udry, J. R. (1995). Sociology and biology: What biology do sociologists need to know? *Social Forces*, *73*(4), 1267–1278.

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). 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, 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*,
   192, 227–236.

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.

Appendix A

Age 10.5 Replication

Between-Family Analyses

Table A1

Between-Family: Gen1 Intelligence Predicts Gen2 AFI

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen1 Mean Intel	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)
Constant	−0.75*** (−1.01, −0.50)	−0.80*** (−1.09, −0.51)	−0.86*** (−1.14, −0.57)
Sample Size	342	264	282
R^2	0.09	0.10	0.11
Adjusted R^2	0.09	0.10	0.10
Residual Std. Error	0.70 (df = 340)	0.70 (df = 262)	0.69 (df = 280)
F Statistic	33.50*** (df = 1; 340)	29.40*** (df = 1; 262)	33.20*** (df = 1; 280)

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

Table A2

Between-Family: Gen2 Intelligence Predicts Gen2 AFI

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean Intel	0.10*** (0.04, 0.16)	0.10*** (0.04, 0.17)	0.09*** (0.02, 0.15)
Constant	−0.02 (−0.09, 0.06)	−0.01 (−0.10, 0.07)	−0.03 (−0.12, 0.05)
Sample Size	345	267	283
R²	0.03	0.03	0.02
Adjusted R²	0.03	0.03	0.02
Residual Std. Error	0.73 (df = 343)	0.74 (df = 265)	0.73 (df = 281)
F Statistic	10.50*** (df = 1; 343)	8.99*** (df = 1; 265)	6.91*** (df = 1; 281)

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

Table A3

Between-Family: Gen1 & Gen2 Intelligence Predict Gen2 AFI

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen1 Mean Intel	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)
Gen2 Mean Intel	0.02 (−0.05, 0.09)	0.02 (−0.06, 0.09)	0.01 (−0.07, 0.08)
Constant	−0.71*** (−1.01, −0.42)	−0.77*** (−1.10, −0.44)	−0.84*** (−1.18, −0.51)
Sample Size	338	260	278
R^2	0.09	0.10	0.11
Adjusted R^2	0.09	0.10	0.10
Residual Std. Error	0.70 (df = 335)	0.71 (df = 257)	0.70 (df = 275)
F Statistic	16.90*** (df = 2; 335)	14.70*** (df = 2; 257)	16.50*** (df = 2; 275)

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

Within-Family Analyses

Table A4

Within-Family: Gen1 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Difference in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.30*** (0.19, 0.42)	0.33*** (0.19, 0.46)	0.38*** (0.25, 0.50)
Gen1 Mean Intel	−0.003 (−0.01, 0.002)	−0.002 (−0.01, 0.004)	−0.01*** (−0.01, −0.003)
Gen1 Dif Intel	0.001 (−0.003, 0.01)	0.001 (−0.01, 0.01)	0.000 (−0.005, 0.01)
Constant	1.18*** (0.89, 1.47)	1.18*** (0.84, 1.52)	1.50*** (1.18, 1.83)
Sample Size	336	258	278
R^2	0.07	0.08	0.12
Adjusted R^2	0.07	0.07	0.11
Residual Std. Error	0.76 (df = 332)	0.78 (df = 254)	0.73 (df = 274)
F Statistic	8.83*** (df = 3; 332)	7.64*** (df = 3; 254)	12.00*** (df = 3; 274)

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

Table A5

Within-Family: Gen2 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Difference in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.33*** (0.21, 0.44)	0.37*** (0.23, 0.50)	0.36*** (0.23, 0.49)
Gen2 Mean Intel	−0.08** (−0.15, −0.01)	−0.09** (−0.17, −0.003)	−0.11*** (−0.19, −0.03)
Gen2 Dif Intel	0.02 (−0.02, 0.07)	0.02 (−0.03, 0.07)	0.002 (−0.04, 0.05)
Constant	1.04*** (0.95, 1.13)	1.06*** (0.96, 1.17)	1.05*** (0.96, 1.15)
Sample Size	287	219	234
R^2	0.10	0.12	0.13
Adjusted R^2	0.09	0.11	0.11
Residual Std. Error	0.76 (df = 283)	0.78 (df = 215)	0.74 (df = 230)
F Statistic	10.50*** (df = 3; 283)	9.57*** (df = 3; 215)	11.00*** (df = 3; 230)

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

Table A6

Within-Family: Gen1 & Gen2 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Difference in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.31*** (0.19, 0.44)	0.35*** (0.20, 0.49)	0.39*** (0.25, 0.52)
Gen2 Mean Intel	−0.08* (−0.16, 0.01)	−0.09* (−0.19, 0.01)	−0.06 (−0.15, 0.03)
Gen1 Mean Intel	0.000 (−0.01, 0.01)	0.001 (−0.01, 0.01)	−0.01** (−0.01, −0.000)
Gen2 Dif Intel	0.02 (−0.03, 0.06)	0.02 (−0.03, 0.07)	0.000 (−0.05, 0.05)
Gen1 Dif Intel	0.002 (−0.004, 0.01)	−0.001 (−0.01, 0.01)	0.002 (−0.004, 0.01)
Constant	1.02*** (0.66, 1.38)	1.02*** (0.60, 1.44)	1.44*** (1.04, 1.84)
Sample Size	282	214	232
R^2	0.09	0.11	0.14
Adjusted R^2	0.08	0.09	0.12
Residual Std. Error	0.76 (df = 276)	0.78 (df = 208)	0.73 (df = 226)
F Statistic	5.71*** (df = 5; 276)	5.08*** (df = 5; 208)	7.26*** (df = 5; 226)

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

Appendix B

Age 11.5 Replication

Between-Family Analyses

Table B1

Between-Family: Gen1 Intelligence Predicts Gen2 AFI

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen1 Mean Intel	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)
Constant	−0.75*** (−1.01, −0.50)	−0.80*** (−1.09, −0.51)	−0.86*** (−1.14, −0.57)
Sample Size	342	264	282
R^2	0.09	0.10	0.11
Adjusted R^2	0.09	0.10	0.10
Residual Std. Error	0.70 (df = 340)	0.70 (df = 262)	0.69 (df = 280)
F Statistic	33.50*** (df = 1; 340)	29.40*** (df = 1; 262)	33.20*** (df = 1; 280)

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

Table B2

Between-Family: Gen2 Intelligence Predicts Gen2 AFI

	Dependent variable: Average of Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean Intel	0.10*** (0.04, 0.16)	0.12*** (0.05, 0.19)	0.09*** (0.03, 0.16)
Constant	−0.02 (−0.09, 0.06)	−0.01 (−0.10, 0.08)	−0.04 (−0.12, 0.05)
Sample Size	339	262	278
R^2	0.03	0.04	0.03
Adjusted R^2	0.03	0.04	0.02
Residual Std. Error	0.73 (df = 337)	0.74 (df = 260)	0.72 (df = 276)
F Statistic	11.70*** (df = 1; 337)	10.80*** (df = 1; 260)	7.65*** (df = 1; 276)

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

Table B3

Between-Family: Gen1 & Gen2 Intelligence Predict Gen2 AFI

| | *Dependent variable:* Average of Gen2 AFI | | |
	Mixed	Daughters	Sons
Gen1 Mean Intel	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)
Gen2 Mean Intel	0.01 (−0.06, 0.08)	0.02 (−0.06, 0.10)	−0.01 (−0.08, 0.07)
Constant	−0.72*** (−1.03, −0.42)	−0.75*** (−1.10, −0.40)	−0.86*** (−1.21, −0.51)
Sample Size	333	255	274
R^2	0.09	0.10	0.10
Adjusted R^2	0.08	0.09	0.10
Residual Std. Error	0.71 (df = 330)	0.71 (df = 252)	0.70 (df = 271)
F Statistic	15.80*** (df = 2; 330)	13.80*** (df = 2; 252)	15.60*** (df = 2; 271)

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

Within-Family Analyses

Table B4

Within-Family: Gen1 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Differences in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.30*** (0.19, 0.42)	0.33*** (0.19, 0.46)	0.38*** (0.25, 0.50)
Gen1 Mean Intel	−0.003 (−0.01, 0.002)	−0.002 (−0.01, 0.004)	−0.01*** (−0.01, −0.003)
Gen1 Dif Intel	0.001 (−0.003, 0.01)	0.001 (−0.01, 0.01)	0.000 (−0.005, 0.01)
Constant	1.18*** (0.89, 1.47)	1.18*** (0.84, 1.52)	1.50*** (1.18, 1.83)
Sample Size	336	258	278
R^2	0.07	0.08	0.12
Adjusted R^2	0.07	0.07	0.11
Residual Std. Error	0.76 (df = 332)	0.78 (df = 254)	0.73 (df = 274)
F Statistic	8.83*** (df = 3; 332)	7.64*** (df = 3; 254)	12.00*** (df = 3; 274)

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

Table B5

Within-Family: Gen2 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Differences in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.34*** (0.22, 0.46)	0.37*** (0.23, 0.51)	0.38*** (0.25, 0.51)
Gen2 Mean Intel	−0.10*** (−0.17, −0.03)	−0.12*** (−0.21, −0.03)	−0.13*** (−0.21, −0.05)
Gen2 Dif Intel	0.04 (−0.01, 0.08)	0.03 (−0.02, 0.08)	0.01 (−0.03, 0.06)
Constant	1.05*** (0.96, 1.14)	1.07*** (0.96, 1.17)	1.08*** (0.98, 1.17)
Sample Size	286	223	230
R^2	0.12	0.13	0.15
Adjusted R^2	0.11	0.12	0.14
Residual Std. Error	0.75 (df = 282)	0.77 (df = 219)	0.73 (df = 226)
F Statistic	12.50*** (df = 3; 282)	11.00*** (df = 3; 219)	13.20*** (df = 3; 226)

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

Table B6

Within-Family: Gen1 & Gen2 Differences in Intelligence Predict Gen2 Differences in AFI

	Dependent variable: Differences in Gen2 AFI		
	Mixed	Daughters	Sons
Gen2 Mean AFI	0.33*** (0.20, 0.45)	0.34*** (0.19, 0.48)	0.42*** (0.28, 0.55)
Gen2 Mean Intel	−0.10** (−0.18, −0.01)	−0.13** (−0.23, −0.03)	−0.07 (−0.16, 0.02)
Gen1 Mean Intel	0.001 (−0.01, 0.01)	0.002 (−0.01, 0.01)	−0.01** (−0.01, −0.000)
Gen2 Dif Intel	0.03 (−0.01, 0.08)	0.04 (−0.02, 0.09)	0.01 (−0.04, 0.06)
Gen1 Dif Intel	0.002 (−0.004, 0.01)	−0.002 (−0.01, 0.005)	0.001 (−0.004, 0.01)
Constant	0.98*** (0.61, 1.35)	0.92*** (0.49, 1.36)	1.47*** (1.06, 1.89)
Sample Size	278	215	226
R^2	0.11	0.12	0.16
Adjusted R^2	0.10	0.10	0.14
Residual Std. Error	0.75 (df = 272)	0.78 (df = 209)	0.73 (df = 220)
F Statistic	6.86*** (df = 5; 272)	5.97*** (df = 5; 209)	8.50*** (df = 5; 220)

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

Appendix C

Correlations

Table C1

Zero-Order Correlations

	G1 Dif Intel	G1 Mean Intel	G2 Dif AFI	G2 Dif Intel	G2 Mean AFI	G2 Mean Intel
G1 Dif Intel	1					
G1 Mean Intel	0.01	1				
G2 Dif AFI	0.02	0.04	1			
G2 Dif Intel	0.16	−0.01	0.05	1		
G2 Mean AFI	−0.05	0.3	0.29	−0.05	1	
G2 Mean Intel	−0.01	0.43	−0.07	0.04	0.13	1

−1   −0.6   −0.2   0.2   0.6   1
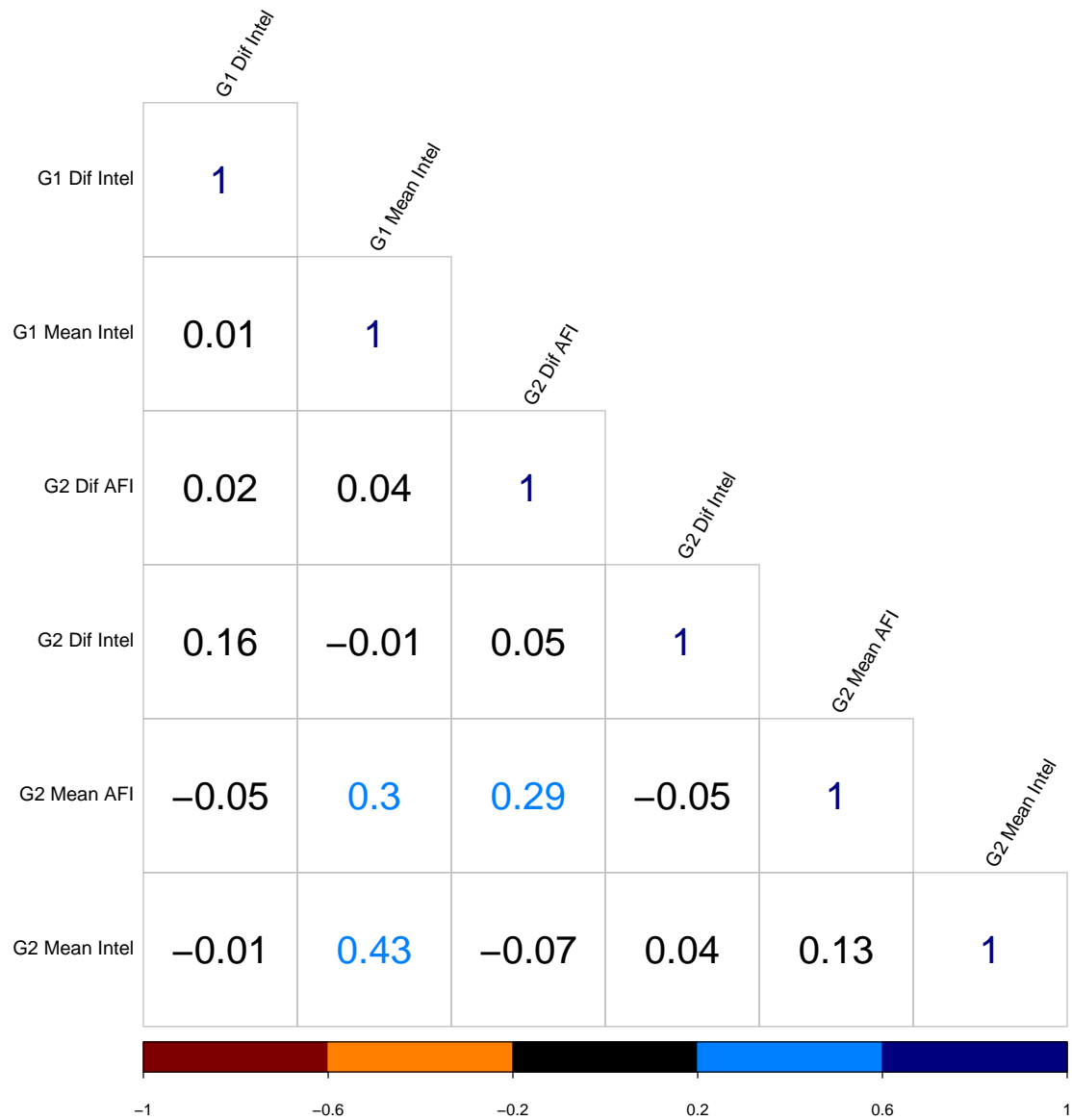
Table C2

Semi-Partial Correlations for Between-Family Joint Model, Predicting Gen2 Averages of AFI

	Semi-Partial Correlation (p-value)		
Predictor	Mixed	Daughters	Sons
Gen1 Mean Intel	0.25 (<.0001)	0.27 (<.0001)	0.272 (<.0001)
Gen2 Mean Intel	−0.003 (0.96)	0 (0.997)	−0.02 (0.73)

Notes: Joint Model is in Table 6. The semi-partial correlation of Gen2 Mean AFI with each variable controls for all other variables in the table.

Table C3

Pairwise Semi-Partial Correlations for Within-Family Joint Model, Predicting Gen2 Differences in AFI

	Semi-Partial Correlation (p-value)		
Predictor	Mixed	Daughters	Sons
Gen2 Mean AFI	0.29 (<.0001)	0.29 (<.0001)	0.36 (<.0001)
Gen2 Mean Intel	-0.075 (0.21)	-0.075 (0.21)	-0.081 (0.22)
Gen1 Mean Intel	0.013 (0.83)	0.013 (0.83)	-0.089 (0.18)
Gen2 Dif Intel	0.06 (0.31)	0.06 (0.31)	0.018 (0.79)
Gen1 Dif Intel	0.021 (0.73)	0.021 (0.73)	0.016 (0.81)

Notes: Joint Model is in Table 9. The semi-partial correlation of Gen2 differences in AFI with each variable controls for all other variables in the table.