SSM - Population Health 17 (2022) 100960

Contents lists available at ScienceDirect

SSM - Population Health

journal homepage: www.elsevier.com/locate/ssmph

Using sibling models to unpack the relationship between education and cognitive functioning in later life

Pamela Herd [a,*], Kamil Sicinski [b]

[a] Georgetown University, McCourt School of Public Policy, 37th and O Streets, NW. Old North, Suite 100, Washington, DC, 20057, USA
[b] Wisconsin Longitudinal Study, University of Wisconsin-Madison, USA

ARTICLE INFO

Keywords:
Education polygenic score
Cognitive functioning
Education

ABSTRACT

As the population ages and the prevalence of dementia increases, unpacking robust and persistent associations between educational attainment and later life cognitive functioning is increasingly important. We do know, from studies with robust causal designs, that policies that increase years of schooling improve later life cognitive functioning. Yet these studies don't illuminate *why* older adults with greater educational attainment have relatively preserved cognitive functioning. Studies focused on *why*, however, have been hampered by methodological limitations and inattention to some key explanations for this relationship. Consequently, we test explanations encompassing antecedent factors, specifically family environments, adolescent IQ, and genetic factors, as well as adult mediating mechanisms, specifically health behaviors and health. We employ the Wisconsin Longitudinal Study, which includes 80 years of prospectively collected data on a sample of 1 in every 3 high school graduates, and a selected sibling, from the class of 1957. Sibling models, and the inclusion of prospectively collected early and midlife covariates, allows us to address the explanatory and methodological limitations of the prior literature to better unpack the relationship between education and later life cognitive functioning. We find little evidence that early life genetic endowments and environments, or midlife health and health behaviors, explain the relationship. Adolescent cognition, however, does matter; higher educational attainment, linked to antecedent adolescent cognitive functioning, helps protect against lower levels of cognitive functioning in later life. Both adolescent cognition and education, however, independently associate with later life cognitive functioning at relatively similar magnitudes. Educational attainment's relationship to later life cognitive functioning is not simply a function of adolescent cognitive functioning.

1. Introduction

According to the 2020 Lancet commission report, one of the most robust correlates for dementia and cognitive impairment in later life is educational attainment (Livingston et al., 2020) At age 65, those with a college degree can anticipate spending 83 percent of their remaining life expectancy with good cognitive functioning, compared to those without a high school degree, who can anticipate spending less than half of remaining life expectancy with good cognitive functioning (Crimmins et al., 2018).

Yet, while there have been significant scientific breakthroughs in research, we have yet to develop effective clinical treatments that preserve cognitive functioning in later life (Elmaleh et al., 2019). And while increasing levels of education across cohorts has reduced population level dementia risk in younger cohorts, because of population aging, the numbers of individuals with cognitive impairment will increase. In this context, a better understanding of why those with higher educational attainment have relatively higher levels of cognitive functioning in later life is critical.

In tandem, two sets of studies have emerged to help unpack this relationship. One set, employing robust causal designs, has found that policies that increase years of schooling improve later life cognitive functioning. Yet these designs don't illuminate *why* older adults with greater educational attainment have relatively preserved cognitive functioning. Another set of studies focused on *why, however,* have been hampered by methodological limitations and inattention to some key explanations for this relationship. Consequently, we employ sibling models, which allow us to address the explanatory and methodological limitations of the former and latter approaches, to unpack the relationship between education and later life cognitive functioning.

* Corresponding author.
E-mail address: ph627@georgetown.edu (P. Herd).

https://doi.org/10.1016/j.ssmph.2021.100960
Received 18 May 2021; Received in revised form 5 November 2021; Accepted 6 November 2021
Available online 20 November 2021

Consequently, we test explanations encompassing antecedent factors, specifically family environments, adolescent IQ, and genetic factors, as well as adult mediating mechanisms, specifically health behaviors and health.

1.1. Causal evidence documenting the influence of additional years of schooling on later life cognitive functioning

There is robust evidence of a causal impact of additional years of schooling on later life cognitive functioning (Banks & Mazzonna, 2012; Glymour et al., 2008; Hamad et al., 2018; Schneeweis et al., 2014). Specifically, a series of recent studies have employed exogenous policy changes, such as the institution of mandatory schooling, which typically involve students completing one or two additional years of schooling. Even with these relatively small differences, these studies have demonstrated improved cognitive outcomes in later life with this additional schooling achieved earlier in life.

Importantly, however, these studies don't unpack why there such a strong relationship between higher educational attainment and better cognitive outcomes in later life. Instead, this approach *rules out*, for example, the influence of early life endowments, like genetics, and family environments (Banks & Mazzonna, 2012; Glymour et al., 2008; Hamad et al., 2018; Schneeweis et al., 2014). The policy changes are not caused by individual differences, like early life family resources, which might otherwise be correlated with both educational attainment and later life cognitive functioning. It also means, however, that this approach can't specify how early life endowments, like genetics or adolescent IQ, and family environments shape relationships between education and later life cognitive functioning, nor does this literature explore midlife health and health and health behaviors as possible adult mechanisms (Herd et al. Forthcoming).

In short, studies focused on exogenous policy changes do an excellent job demonstrating, in a causal fashion, that one to two years of additional schooling can be protective for cognitive functioning in later life. But these studies cannot help unpack the more dynamic early and mid life course relationships that link educational attainment to later life cognitive functioning, as we lay out in Fig. 1, and discuss below. They have not helped us unpack *why* there is such a robust relationship between educational attainment and later life cognitive functioning.

1.2. Explanations for the relationship between educational attainment and later life cognitive functioning

Another set of studies has attempted to tease out the life course mechanisms to help understand why there is such a robust relationship between educational attainment and cognitive functioning in later life

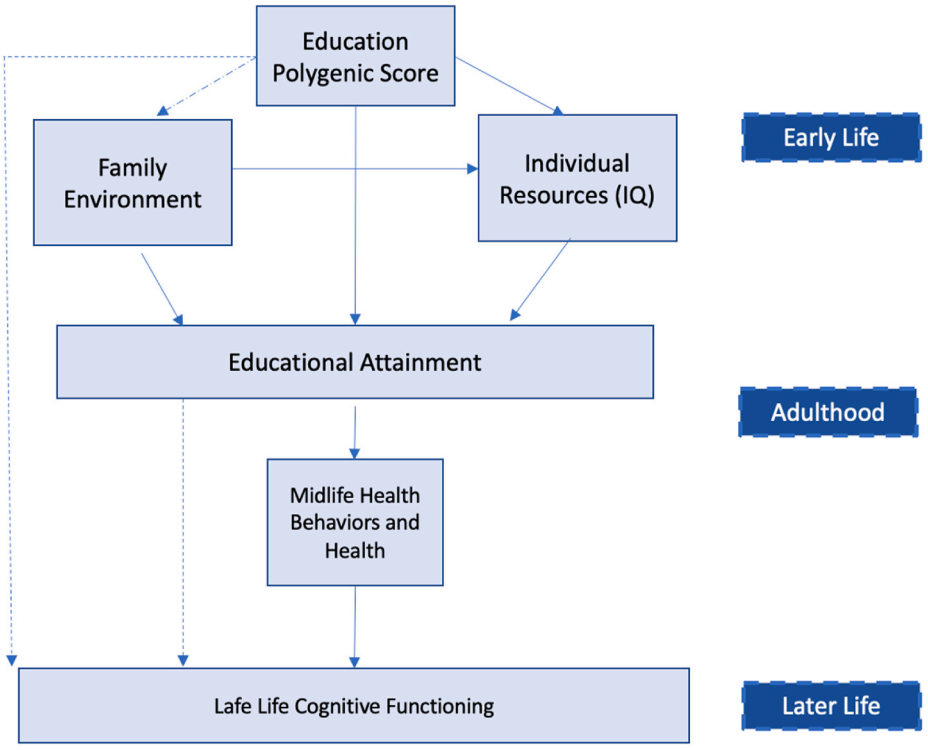


Fig. 1. Parsimonious model of proposed relationship*.

(Beck et al., 2018; Clouston et al., 2012; Greenfield et al. 2020a, 2020b; Moorman et al., 2018; Zhang et al., 2020; Walsemann & Ailshire, 2020). This work has found some evidence for different parts of the relationships documented in Fig. 1. But as we detail below, existing work has not been able to comprehensively examine this set of relationships, and it is also has methodological limitations in identifying and accounting for genetic influences, adolescent cognitive functioning, and early life environments, that may shape the life course relationship between educational attainment and later life cognitive functioning.

1.2.1. *Genetic endowments, education, and later life cognitive functioning*

The first limitation of existing work, with one significant exception (Ding et al., 2019), is not adequately accounting for, or testing for, genetic factors that may influence both educational attainment and later life cognitive outcomes. There is robust evidence that genetics has a strong influence on educational attainment; a polygenic score for educational attainment, what is effectively an individual cumulative genetic "risk" score, explains upward to 12 percent of the variance in educational attainment across a range of population samples (Lee et al., 2018). This magnitude is similar to the influence of one parent's educational attainment (Cesarini & Visscher, 2017; Lee et al., 2018). It is plausible that the same genetic factors that influence educational attainment, also influence later life cognitive functioning, either indirectly or directly. For example, studies exploring the mediating mechanisms between the education polygenic score and subsequent educational attainment, have found robust evidence that brain functioning and cognitive functioning in early life are a plausible mediating mechanism (Belsky et al., 2018; Judd et al., 2020; Lee et al., 2018).

The one exception is recent work by Ding et al. (2019), which found a robust relationship between the education polygenic score and cognitive functioning in the Health and Retirement Study (HRS). Moreover, they found that the relationship remained robust even after controlling for phenotypic educational attainment. But the study, while the first to estimate these relationships, has a few potential limitations. First, it did not account for early life cognitive functioning. The influence of education polygenic score, and education more generally, may be via cognitive development earlier in life, which in turn influences educational attainment, rather than a more direct influence of genetics on later life cognitive trajectories.

The second issue is that the estimate of the influence of the education polygenic score may have been biased. A number of recent studies have argued that polygenic scores, particularly those generated on samples of unrelated individuals, may not reflect pure direct genetic effects due to 'population phenomena' (Morris et al., 2020). Of particular relevance, the education polygenic score may pick up the influence of childhood environments (Armstrong-Carter et al., 2021; Kong et al., 2018; Mills & Tropf, 2020). For example, recent work has found that mothers with higher education polygenic scores were more likely to cognitively stimulate their children via activities like reading, all of which contributed to their child's better academic outcomes, over and above genetics their children directly inherited; this is termed genetic nurture or specifically passive gene-environment correlation (Scarr & McCartney, 1983, pp. 424–435; Wertz et al., 2019). This is why Fig. 1 includes the dotted arrow from the education polygenic score to family environment. It's also the case that even outside of family environments, those with higher education polygenic scores may seek out broader environments that further enhance their cognitive and educational outcomes, what is termed evocative gene-environment correlation (Scarr & McCartney, 1983, pp. 424–435).

While other polygenic scores may have similar issues, the education polygenic score may be more exposed due to factors like assortative mating; people with similar levels of educational attainment are more likely to partner, implying more genetic similarity among partners than the population as a whole. The correlations between genotype and phenotype (education) that result can cause biased estimates of the resulting influence of genotype on phenotype in future generations (Brumpton et al., 2020; Kong et al., 2018; Morris et al., 2020). That said, existing work has demonstrated only a small influence of assortative mating (Kong et al., 2018; Selzam et al., 2019). Currently, the primary proposed mechanism to address many of these possible biases are through the use of sibling or family-based models (Morris et al., 2020), but these models were not possible in the data employed by Ding et al. (2019).

1.2.2. *Early life cognition, educational attainment, and later life cognitive functioning*

Aside from genetics, one of the most obvious questions is the role of early life cognition as an antecedent influence on the relationship between educational attainment and later life cognitive functioning. There is some evidence that adolescent cognition has a large influence on later life cognitive functioning, with education, not surprisingly, being a key proposed mechanism between the two (as illustrated in Fig. 1) (Cox et al., 2016; Anderson et al., 2020). But because most data lack early life measures of cognition, there are just a few studies—and the findings are mixed. A few studies have found that education's role is confined to being on a pathway from early life cognition to later life cognition (Cox et al., 2016; Anderson et al., 2020). Other work, however, has found that education is not only important as a mechanism from adolescent cognition to later life cognitive functioning, but also as an independent influence (as proposed in Fig. 1) (Greenfield, Moorman, & Rieger, 2020; Greenfield, Akincigil, & Moorman, 2020:; Moorman et al., 2018).

One limitation of the few prior studies parsing out the relationship between early life cognition, educational attainment and later life cognitive functioning, however, is their limited accounting for early life environments. While a robust literature has found an important role for parental socioeconomic status in influencing relationships between early life cognition, education, and later life cognitive outcomes (Greenfield & Moorman, 2019: Moorman et al., 2018; Zhang et al., 2020), there is a broad array of early life conditions for which these studies lack observable measures, Developmental psychologists and education researchers have found early life conditions, ranging from supportive parental relationships to school and neighborhood factors, can influence both cognitive development in children and educational attainment (Brooks-Gunn et al., 1993; Englund et al., 2004; Flouri, 2006; Luster & McAdoo, 1996; Nielson 2006; Sautz and Matzel 2018). Children who grow up in households without sufficiently nurturing family relationships, with lower educational and economic resources, and in neighborhoods and communities with more limited resources, have lower levels of cognitive functioning in early life, and achieve lower levels of educational attainment, though these outcomes are not interchangeable (Brooks-Gunn et al., 1993; Englund et al., 2004; Flouri, 2006; Herd et al., 2021; Luster & McAdoo, 1996; Nielson 2006; Sautz and Matzel 2018; Sewell et al., 1971).

There is also a risk that genetic endowments, for which there are no observable measures, may bias findings. The education polygenic score does predict both adolescent IQ and educational attainment, but a robust body of evidence shows that genetic influence on early life cognition is different, and likely more robust, than genetic influence on educational attainment (Branigan et al., 2013; Lee et al., 2018). And the education polygenic score may not fully capture all of the genetic factors that influence educational attainment (Lee et al., 2018).

Consequently, to best parse out these relationships, a broader

accounting of unobserved early life environments, and to some extent genetic endowments (aside from the education polygenic score), may be important to understand these relationships. Sibling models are one way to do this; focusing on within family sibling variation accounts for a range of unobserved early life environment and, at least partially, genetic endowments that might influence these relationships. One study that employed sibling models, did find independent relationships between both adolescent cognition and education and later life cognitive functioning in the Wisconsin Longitudinal Study (Cook & Fletcher, 2015). The study, however, was not focused on unpacking life course relationships, or specifically relationships between adolescent IQ, educational attainment, and later life cognitive functioning, so doesn't offer insights into how a broader array of genetic factors, like the education polygenic score, or midlife health and health behaviors might affect the relationship between education and later life cognitive functioning. It does provide a model, however, for the potential strength of employing sibling models.

1.2.3. Midlife health and health behaviors as mediators between educational attainment and later life cognitive functioning

In addition to antecedent explanations, including genetics, adolescent cognition, and family environment, adult mediating health behaviors and health may be another explanation for the relationship between education and later life cognitive functioning. There has been surprisingly little empirical work testing health and health behaviors. Yet, they are among the most plausible adult mediating mechanisms because of the robust relationships between education, health behaviors, chronic conditions and later life cognitive functioning.

Education strongly patterns health behaviors including smoking, obesity, and physical activity, as well as chronic conditions like heart disease, stroke, and diabetes (Herd et al., 2007; Lantz et al. 2001; Marshall et al., 2015). In many cases, like with diabetes, the risk is double for those without high school degrees compared to those with college degrees (Beckles et al., 2016). Education's influence on these chronic conditions is both indirect, via health behaviors, and also more direct (Herd et al., 2007; Lantz et al. 2001).

These health behaviors and chronic conditions, in turn, predict lower levels of cognitive functioning in later life (Sabia et al., 2009; Meuller et al., 2020; Cadar et al. 2012). In short, there is robust evidence that modifiable health behaviors like exercise and smoking, with some evidence regarding obesity, influence risk for outcomes like dementia (Beydoun et al. 2014). Moreover, strokes, cardiovascular disease, hypertension, and diabetes have all been linked to cognitive impairment, especially vascular forms (Lee et al., 2021; Stefanidis et al., 2018).

But the findings from the surprisingly few studies testing the role of health and health behaviors as mediators in the relationship between education and later life cognitive functioning have been mixed. A recent English study found that health behaviors explain up to half of educational disparities in dementia (Deckers et al., 2019). Ding and colleagues (1992) also found that behaviors did help explain the relationship between the education polygenic score and cognitive decline. Liu and colleagues (2020) found that exercise just slightly reduced these relationships. In contrast, however, Hayward and colleagues (2021) found that declines in dementia prevalence across recent cohorts was driven by increasing educational attainment, but not by trends in health behaviors and chronic conditions. Existing studies have also not accounted for the potential role for unobserved early life conditions and genetic endowments, which may pattern both health behaviors and chronic conditions and educational attainment (Freedman et al., 2008; Pudrovska et al., 2014; Zhang et al., 2020).

1.3. Research questions

Given the limitations of the prior literature, we use the Wisconsin Longitudinal Study, which includes 80 years of prospectively collected data on a sample of 1 in every 3 high school graduates from the class of 1957, as well as a randomly selected sibling, to test the role of early life endowments and environments, as well as midlife health behaviors and health, on later life cognitive functioning. The sibling model design, combined with a robust observable polygenic score for educational attainment and an adolescent cognitive measure, as well as midlife health behaviors and health, allows us to better unpack the life course dynamics of these relationships.

Specifically, we will answer the following three questions. First, does the education polygenic score for educational attainment predict later life cognitive functioning? Second, what is the relationship between adolescent cognition, educational attainment and later life cognitive functioning, and how do unobserved early life environments and endowments influence these relationships? Finally, do adult health behaviors and health mediate any of the relationship between educational attainment and later life cognitive functioning?

2. Materials and methods

2.1. Data

The Wisconsin Longitudinal Study (WLS) provides a unique opportunity to explore the relationship between educational attainment and cognitive functioning in later life. It has three key advantages. First, it is the only sample of older adults that includes a sibling sample design, which better accounts for early life environments and reduces the risk for genetic confounders. Second, it is the only nearly full life course longitudinal US based sample of older adults that includes prospectively collected measures of adolescent cognitive functioning, as well as other key prospectively collected measures of health and cognition. Third, it's cohort design focused on those born around 1939, which helps rule out, for example, period effects that might confound analyses focused on educational disparities. The WLS is uniquely situated to explore the questions posed in this study (NRC 2013).

The WLS is based on a randomly selected 1/3 sample of all 1957 Wisconsin high school graduates (Herd et al., 2014), and a randomly selected sibling for each eligible graduate. Participants were originally empaneled at age 18 (1957), which was followed with data collection at ages 25 (1964), 36 (1975), 54 (1993), 65 (2003-4), and 72 (2011–12). Data on cognition, which are detailed below, are drawn from administrative records (standardized IQ tests administered to all Wisconsin high school students starting in the 1920s) and from cognitive assessments administered in 2011.

In regards to genetics, WLS collected saliva samples in 2007 and 2011 from respondents using Oragene kits and a mailback protocol patterned closely on a previous study (Rylander-Rudqvist et al., 2006). Compliance to the DNA request was about 5 percentage points higher among males, but broadly matched response rates for other data in the WLS (Herd et al., 2014). After quality control, respondents were genotyped at ~710,000 markers (before imputation) utilizing the Omni-Express beadchip. Documentation is accessible at http://www.ssc.wisc.edu/wlsresearch/.

Genetic data and measures of adolescent cognition, as well as some additional key health measures, were available for approximately 7000 graduates and siblings whom completed the 2011 data collection, which included 4837 graduates and 2221 sibling participants. Because we needed sibling pairs, however, the final sample included 2920 siblings and graduates or 1460 pairs, reflecting mismatched graduate and sibling cases (e.g. a sibling that did not have a graduate who completed a 2011 interview due to either mortality or attrition). We did a range of sensitivity analyses to address how attrition may have affected the final sample. First, we included a weight to address selection in terms of who provided genetic data. Second, we ran simplified models that excluded health and health behavior variables, which constituted about 500 missing cases, to see how it affected the education, education polygenic score, and adolescent cognition findings. Neither produced evidence that the findings substantially differed. Finally, we analyzed differences

Table 1
Descriptive Statistics for the Wisconsin Longitudinal Study Participants.

	Full sample		Original Graduate Participant		Siblings	
Variables	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Cognitive Outcomes (standardized): 2011						
Letter Fluency	0.078	1.01	0.035	0.994	0.122	1.024
Digit Ordering*	0.067	1.018	0.058	1.007	0.076	1.029
Immediate Recall*	0.032	1.007	0.056	0.968	0.008	1.045
Delayed Recall*	-0.025	1.049	-0.009	1.014	-0.042	1.082
Demographics						
Female	0.532	0.499	0.519	0.5	0.546	0.498
Age in 2011 (difference from mean)	-0.118	4.62	0.618	0.903	-0.858	6.399
Education Polygenic Score	0.084	0.992	0.1	0.991	0.068	0.993
Years of Education	13.955	2.409	13.857	2.342	14.055	2.472
1957 IQ Score (standardized)	0.148	1.034	0.078	0.992	0.22	1.07
Chronic Health Conditions						
Ever had a stroke	0.057	0.232	0.055	0.227	0.06	0.237
Has diabetes	0.164	0.37	0.166	0.372	0.162	0.369
Has hypertension	0.59	0.492	0.612	0.487	0.567	0.496
Has or had cardiovascular diseases	0.243	0.429	0.25	0.433	0.236	0.425
Health Behaviors						
BMI						
Underweight '11 (BMI<18.5)	0.004	0.067	0.002	0.045	0.007	0.083
Normal weight '11 (18.5<=BMI<=24.9)	0.368	0.482	0.376	0.485	0.359	0.48
Overweight '11 (25.0<=BMI<=29.9)	0.415	0.493	0.435	0.496	0.395	0.489
Obese '11 (BMI>=30.0)	0.213	0.409	0.188	0.391	0.238	0.426
Exercise						
Exercised >3 times per week ('92)	0.195	0.396	0.209	0.407	0.18	0.384
Exercised 1 or 2 times per week ('92)	0.153	0.36	0.141	0.348	0.166	0.372
Exercised one to three times per month	0.175	0.38	0.168	0.374	0.183	0.387
Exercised less than once per month ('92	0.403	0.491	0.43	0.495	0.375	0.484
Smoking						
Never smoked	0.438	0.496	0.434	0.496	0.441	0.497
Former smoker	0.483	0.5	0.491	0.5	0.474	0.5
Current smoker	0.079	0.269	0.074	0.262	0.083	0.276
	N=2920		N=1466		N=1466	

* These measures were collected on a random 80 percent subsample. Consequently, the sample size is 2174 for these outcomes.

among the model covariates and outcomes, when comparing the final paired sibling analytic sample with the full ~7000 cases available in the 2011 survey. The only statistically significant difference across all measures was the education polygenic score, which was slightly higher in the final analytic sample. This is driven by the siblings, which are on average, somewhat more select than the graduate sample.[1] That said, sensitivity analyses that compared the full graduate sample to the analytic graduate sample for this study (meaning they had had a paired sibling), did not meaningfully differ.

[1] The graduate sample constitutes ~69 percent of surviving participants. It is important to keep in mind that the WLS is unusual in having a 100 percent response rate from the initial sample frame. The highest initial response rate for a Health and Retirement Study cohort, for example, is 80 percent for the 1931–1939 birth cohort. Consequently, the total response rate is higher in the WLS, even as the length of the study (60 years) is substantially longer.

2.2. Outcomes measures

We employ measures of late life cognitive functioning that were collected in 2005 and 2011 when participates were ~age 65 and ~age 72. These include: letter fluency, which captures verbal fluency and requires individuals to recall as many words that start with a certain letter in 60 s; an immediate and delayed recall task, which captures memory and requires individuals to recall a set of ten words immediately and then ~10 min later; and a number series task, which requires participants to remember and repeat a series of numbers in the correct order. Table 1 provides descriptive statistics on the outcome and covariate measures on the full sample, and separately for the original empaneled graduate and their paired sibling.

2.3. Covariates

We control for age and sex. Note that there is very limited variance in age for the graduate sample that all graduated from high school in 1957. Approximately 90 percent of siblings fall on either side of the graduates age by 7 years. In addition, we include the following measures:

Adolescent Measures of Cognitive Functioning: 1) WLS is one of just a few longitudinal aging cohort studies with an early life cognitive measure. This measure is derived from the Henmon-Nelson IQ test administered to WLS participants during their junior year in high school (1956). It was a 30-minute test consisting of 90 items in order of increasing difficulty. It included vocabulary, sentence completion, disarranged sentences, classification, logical selection, series completion, directions, analogies, anagrams, proverb interpretation, and arithmetic problems. It highly correlates (0.83) with IQ tests more commonly administered today, especially the WAIS (Watson & Klett, 1975).

Polygenic Score for Educational Attainment: Polygenic scores summarize predictive information in the genome with respect to a particular trait. The scores use weights based on genome-wide association studies (GWAS) conducted in other samples. GWAS studies generally find associations between the most common kind of genetic variation, specifically single nucleotide polymorphisms (SNPs), and outcomes ranging from height and cancers to personality and educational attainment. Each individual, in these studies, has somewhere between 1 and 2.5 million variants (SNPs) that have been identified (Lee et al., 2018). GWAS then identifies the strength of the associations between each variant and the outcome.

The polygenic score for educational attainment that we use is based on a 1.1-million-person GWAS (Lee et al., 2018).[2] Weights of individual variants are multiplied by the count of trait-associated alleles for each SNP and summed across all variants. Polygenic scores that provide robust, out-of-sample, predictions have been developed for outcomes such as height, body mass index (BMI), psychiatric disorders, and smoking (for a review see Ware & Faul, 2021). Detailed information on the construction of the education polygenic scores can be in supporting documentation (https://www.ssc.wisc.edu/wlsresearch/documentation/GWAS), as well as in additional papers (Belsky et al., 2018; Herd et al., 2019). The score is standardized to have a mean of 0 and SD of 1. (Fig. 1)

Population Stratification Controls: We also residualize the education polygenic score for the first 10 principal components, which were estimated from the genome-wide SNP data to account for allele frequency differences across ancestral groups (population stratification) in our analytic sample. The sibling model design also helps address this. Population stratification is a key issue in studies of this kind. In short, there is the risk that results are confounded by ancestry differences. Consequently, study findings can be sensitive to the inclusion of controls for population stratification.

Educational Attainment*:* We employ "years" of schooling, which is a summary measure derived from highest level of schooling students obtained based on measures collected between 1975 and 2011. The measure ranges from 8 years of schooling to 20, which includes those with PhDs. The measure is standardized with 2.5 'year's equal to 1 standard deviation.

Health Behavior and Health Measures: We include covariates for some key health related mediating mechanisms that could explain the relationship between educational attainment and cognitive function. We classify respondents into never, former or current smokers based on their survey responses in 2011. Body mass index (BMI) is grouped into standard Center for Disease Control categories for obese, overweight, and normal weight, based on reports of height and weight in 1993 when respondents were ~age 54. We also include measures of how physically active participants were in 1993, specifically whether they exercised more than 2 times a week, 1–2 times a week, 1–3 times a month or never. The measures of chronic conditions include a series of self-reported medical conditions (diabetes, coronary heart disease/heart attack/angina/congestive heart failure/other heart problems, hypertension, stroke) collected in both 2004 and 2011.

2.4. Statistical analysis

We employ a series of sibling fixed effect OLS regression. The sequential models reflect the life course ordering of these covariates as presented in Fig. 1. We further note that models only including the education polygenic score, and the polygenic score and adolescent cognition, are critical given the potential for collider bias and unobserved variable bias that may specifically affect the relationships between the education polygenic score, educational attainment, and later life cognitive functioning (Akimova et al., 2021). That said, the use of sibling models, and the inclusion of the adolescent cognitive functioning measure, substantially reduces this risk and is an improvement as compared to existing research.

We include analyses that both include and exclude sibling fixed effects to test how this impacts the polygenic score, given population phenomena issues (e.g. Morris et al., 2020), and also to see how unobserved childhood environments and genetic endowments influence the relationship between adolescent cognition, education, midlife health behaviors and health, and later life cognitive functioning. When we exclude fixed effects, we adjust standard errors to account for sibling clusters. Models that only included the graduates, as opposed to simply eliminating the fixed effect, produced the same pattern of results. These analyses are available on request. Appendix includes full model results that include and exclude fixed effects.

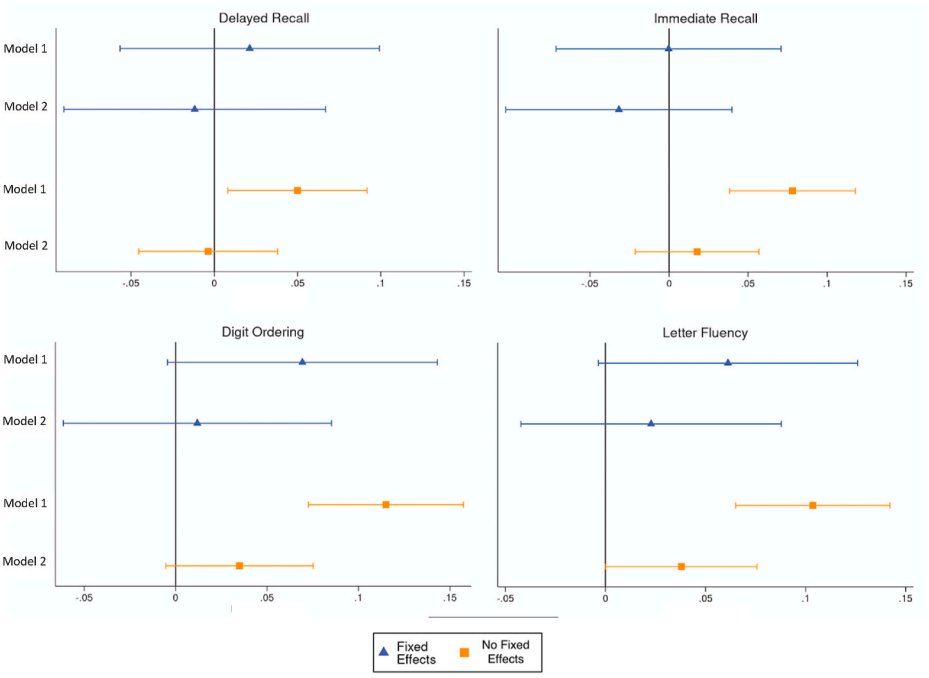
3. Results

The results focus on the cognition measures from 2011 when participants were ~age 72. As already noted, the education polygenic score is residualized for the first 10 principal components estimated from the genome-wide SNP data to account for allele differences across ancestral groups (population stratification) in the analytic sample. The model coefficients underlying the figures are located in Appendix Tables A.1-A.4.

3.1. The influence of the education polygenic score on later life cognitive functioning

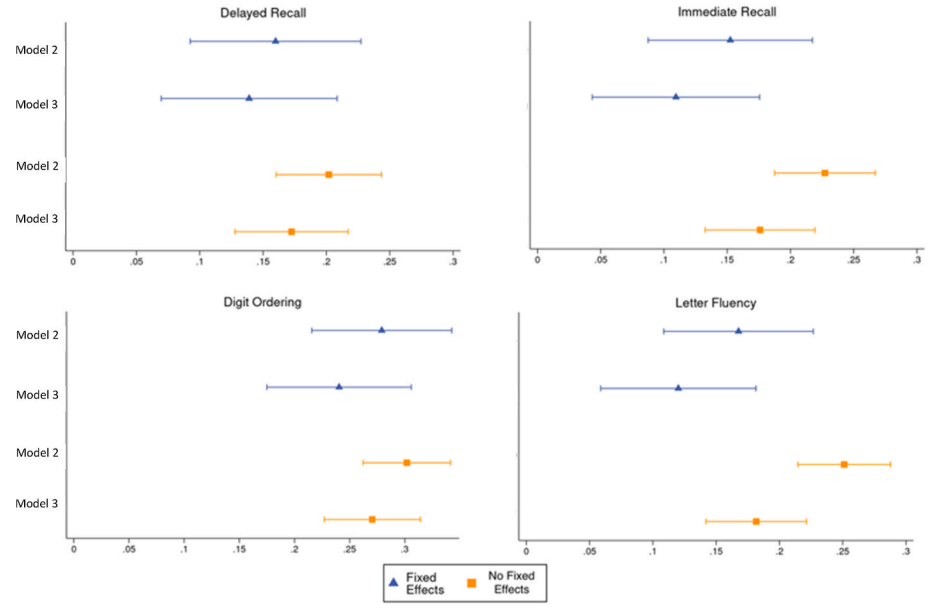
Fig. 2 displays the relationship between the education polygenic score and later life cognitive functioning, both including and excluding sibling fixed effects, across all four cognitive outcomes. The first thing to note is that in non-fixed effect models, controlling for adolescent cognitive functioning explains nearly all of the relationship between the education polygenic score and later life cognitive functioning. This was the case for immediate recall, delayed recall, and digit ordering. The exception to this was letter fluency, but the coefficient size was cut by 60

[2] One SD is equivalent to 2.4 'years' for the education variable.



*Coefficients for the education polygenic score are presented.
Model 1 adjusts for age and sex.
Model 2 adjusts for age, sex, and adolescent cognitive functioning.

Fig. 2. Later life cognition outcomes regressed on the education polygenic score*.

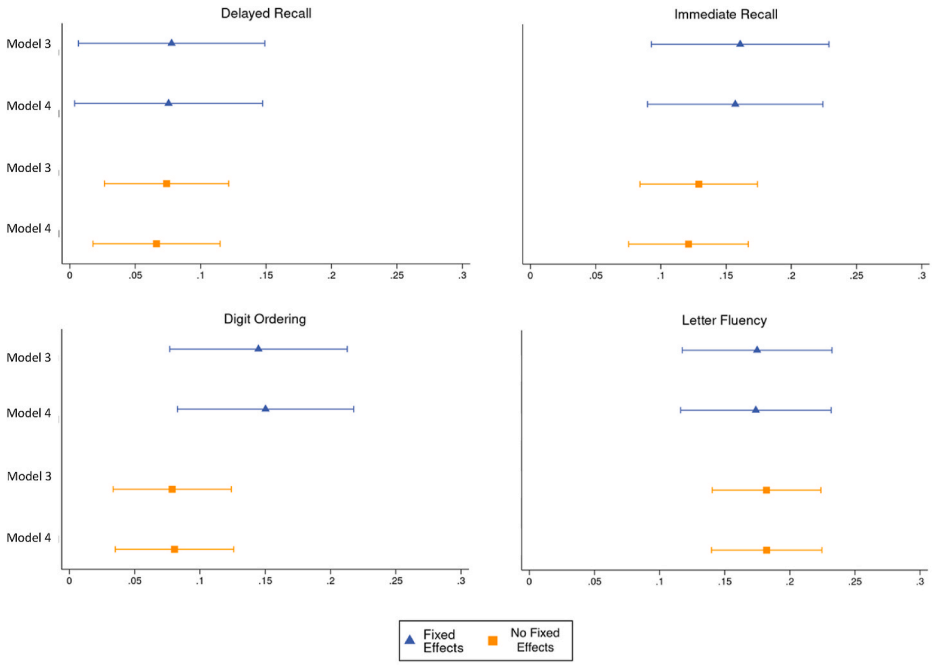


*Coefficients for adolescent IQ are presented.
Model 2 adjusts for age, sex, and the education polygenic score.
Model 3 adjusts for age, sex, the education polygenic score, and educational attainment.

Fig. 3. Later life cognition outcomes regressed on adolescent cognition*.

percent. The inclusion of educational attainment then reduced the coefficient to near zero as shown in the Appendix tables.

Second, the figures demonstrate how the education polygenic score coefficients are affected with the inclusion and exclusion of the fixed effect. For all cognitive outcomes, in models only including controls for age and sex, the inclusion of the fixed effect either reduces the coefficient to zero or substantially reduces the coefficient size, as well as substantially increasing the standard errors, though the differences



*Coefficients for educational attainment are presented.
Model 3 adjusts for age, sex, the education polygenic score, and IQ
Model 4 adjusts for age, sex, the education polygenic score, IQ, Health Behaviors and Health

Fig. 4. Later life cognition outcomes regressed on education*.

across fixed and non-fixed effect models do not reach statistical significance. For both delayed and immediate recall, the coefficient for the education polygenic score was never substantively correlated, nor did it achieve statistical significance in any of the models that included sibling fixed effects. The education polygenic score was more correlated with digit ordering and letter fluency, as compared to memory items, but both still had large standard errors. Further, the inclusion of adolescent cognition reduces the education polygenic score coefficients to near zero. In sum, in models that exclude a sibling fixed effect, the relationship between the education polygenic score and later life cognitive functioning was mediated by adolescent cognitive functioning. In sibling fixed effect models only adjusting for age and sex, there was almost no relationship between the education polygenic score and later life cognitive functioning. It is worth noting that the education polygenic score is robustly correlated with adolescent IQ (see Appendix Table A.5), in models with and without fixed effects, though the EPGS is not robustly correlated with later life cognitive functioning, likely pointing to differing factors driving early versus later life cognitive functioning.

3.2. Adolescent cognition, educational attainment, and later life cognitive functioning

Fig. 3 provides evidence as to the relationship between adolescent cognition, educational attainment and later life cognitive functioning, including the role of family environments in explaining these relationships. First, Fig. 3 does demonstrate that the main effects of adolescent cognition are reasonably large. For example, in the fixed effect models, a 1(SD) increase in adolescent cognition is correlated with a 0.16 (SD) increase in later life delayed recall performance. This is equivalent to about 5 years of age (see Appendix Table A.1). Fig. 3, however, also demonstrates that the inclusion of educational attainment does explain some of the relationship between adolescent cognition and later life cognitive functioning, ranging from roughly 10 and 15 percent for digit ordering and delayed recall performance to 25 and 30 percent for immediate recall and letter fluency performance respectively. Higher educational attainment resulting from higher adolescent cognitive functioning helps protect against lower levels of cognitive functioning in later life. Education's role in influencing later life cognitive functioning, is, in part, a function of it acting as a mechanism via adolescent cognitive functioning. Fig. 3 also demonstrates, however, that unobserved family environments or genetic endowments, or some mixture of the two, play a significantly large role in explaining the influence of adolescent cognition on later life cognitive functioning. The adolescent cognition coefficient is 12–50 percent smaller with the inclusion of family fixed effects, with the smaller impact of the fixed effect for digit ordering.

Fig. 4 demonstrates that educational attainment remains significantly correlated with later life cognitive functioning, even after accounting for adolescent cognition. In terms of magnitude, the range is from 0.14 to 0.18 SD increases in cognitive performance in later life, for every one standard deviation increase in educational attainment.[2] This is equivalent to about 5 years of age (see Appendix tables). When comparing the standardized adolescent cognition coefficient presented in Fig. 3 with the standardized education coefficients in Fig. 4, the relative influence of adolescent cognition and education vary across cognitive outcome. In the sibling fixed effects models, for immediate recall and letter fluency, the standardized education coefficients are

roughly 50–60 percent larger than the adolescent cognition coefficient, whereas for digit ordering and delayed recall the standardized adolescent cognition coefficients are 60–80 percent larger. If capturing the total influence, however, adolescent cognition has a somewhat larger influence.

Fig. 4 also demonstrates there is no influence of unobserved early family environments and/or genetic endowments on the relationship between education and later life cognitive functioning. Unlike with adolescent cognition, the inclusion and exclusion of fixed effects has almost no influence on the education coefficient, except in the case of digit ordering where the coefficient size increased, and just slightly increases the standard errors. We should note that the inclusion and exclusion of sibling fixed effects also had no impact on the education coefficient even when adolescent IQ is not included in these models (estimates not presented here). This implies that the relationship between educational attainment and later life cognitive functioning is quite robust to unobserved early life environments and genetic endowments, unlike the relationship between adolescent cognition and later life cognitive functioning.

3.3. Health behaviors and chronic conditions as midlife mediators

Finally, Fig. 4 also presents findings on the role that midlife health behaviors and chronic conditions might play in mediating relationships between educational attainment and later life cognitive functioning. In short, we find little evidence they serve as important midlife mediators. The education coefficient remains nearly the same with the inclusion of these midlife mediators. There are also almost no meaningful differences between the fixed effect and non-fixed models. The appendix tables also detail that just a few of these behaviors and chronic conditions, specifically being overweight, being a smoker, and having had a stroke, are significantly correlated with lower levels of cognitive functioning.

4. Discussion

4.1. Summary of findings

In the context of family fixed effects models, we examined the role of key early life endowments (genetics and adolescent cognition) and family environments as antecedents to educational attainment, as well as health and health behaviors as midlife mediators, to better understand the relationship between educational attainment and later life cognitive functioning. We find a robust independent influence of education on later life cognition, with neither early life genetic endowments and environments, nor midlife health and health behaviors influencing these relationships. Adolescent cognition, however, does partly explain why educational attainment shapes later life cognitive functioning; higher educational attainment resulting from higher adolescent cognitive functioning helps protect against lower levels of cognitive functioning in later life. Adolescent cognition's relationship with later life cognitive functioning however, is influenced by unobserved genetic and family environment factors. Both adolescent cognition and education, however, independently influence later life cognitive functioning at relatively similar magnitudes. Educational attainment's relationship to later life cognitive functioning is not simply a function of adolescent cognitive functioning.

4.2. Interpretation

While many studies have shown a robust correlation between educational attainment and later life cognitive functioning, we were able to address prior limitations in the literature to better unpack these relationships. The education polygenic score, which does robustly predict educational attainment (Herd et al., 2019), including in the context of sibling models in WLS, did not have a direct influence on later life cognitive functioning above and beyond adolescent cognition in models without fixed effects, and had almost no association in family fixed effect models. These findings were in contrast to a recent paper (Ding et al., 2019), which found the polygenic score was predictive of cognitive decline, even after accounting for phenotypic educational attainment. The differences in findings likely reflect our inclusion of adolescent cognition, as well sibling fixed effect models, which have been proposed to address possible population phenomena issues with the education polygenic score (Morris et al., 2020). That said, the Ding et al. (2019) sample included more variation on educational attainment because it encompassed those without high school degrees. This may have influenced the difference in findings across these two studies.

Adolescent cognition, however, plays an important role in these processes. While education explains 10 to 30 percent of the relationship between adolescent cognition and later life cognitive functioning, adolescent cognition remains robustly associate with later life cognitive functioning. We did, however, find that the size of adolescent cognition's influence was sensitive to properly accounting for unobserved early life family environments and/or genetic factors, with the coefficient 12 to 50 percent smaller with the inclusion of family fixed effects. Existing studies in designs without family fixed effects may be underestimating the influence of unobserved early life environments and other genetic endowments on these relationships (Anderson et al., 2020; Cox et al., 2016). At the same time, we show that the failure to account for early life cognitive functioning in studies examining the relationship between educational attainment and later life cognitive functioning is missing an important part of the puzzle.

We also found that educational attainment robustly correlates with later life cognitive functioning, at magnitudes not dissimilar from adolescent cognitive functioning, and that its influence did not differ in sibling fixed effects models as compared to those without fixed effects. In short, things like family environments and socioeconomic resources, which are adjusted for in sibling models, did not appear to explain education's influence on later life cognitive functioning. The difference in this finding for educational attainment, as compared to adolescent cognition, could reflect a more pernicious role in how early life conditions shape cognitive functioning that may have long term implications for cognitive health in later life. In contrast, the benefits from education may compensate for those early life conditions. Alternately, it may be unobserved genetic confounders; indeed, there is evidence that genetic influences on early life cognition are stronger than for educational attainment (Branigan et al., 2013).

Finally, we found no evidence that health behaviors and health explained the relationship between educational attainment and later life cognitive functioning. This does map onto prior general work on health disparities that has challenged the role, especially, of individualized health behaviors in explaining educational disparities (Lantz et al. 2001). In short, the focus on exercise or nutrition, while possibly having some protective effect for cognitive functioning in later life, will not help reduce educational disparities. The evidence from this study is that focusing interventions on Individual behaviors and health will not

reduce educational disparities in later life cognitive functioning. That said, we do want to note some caution given the possibility this finding may reflect the more limited variation in educational attainment in this sample.

4.3. Limitations and future directions

There are a few study limitations to note, with implications for future research. First, sibling models are powerful, but they cannot account for all unobserved variables, be they genetic or environmental. For example, they would be unlikely to fully account for evocative genetic nurture, where people with higher polygenic education scores might seek out environments that further enhance cognitive functioning—-though early in life children's ability to seek out those environments in ways that would influence adolescent cognitive functioning is constrained by their family environment. We also can't rule out the potential for further confounding and unobserved variable and collider bias that may specifically affect the relationships between the education polygenic score, educational attainment, and later life cognitive functioning (Akimova et al., 2021). That said, the use of sibling models, and the inclusion of the adolescent cognitive functioning measure, substantially reduces this risk and is an improvement as compared to existing research. Moreover, the basic finding that the relationship between the education polygenic score and later life cognitive functioning was not robust held in very parsimonious models. We also note that a new preprint (Fletcher et al., 2021) has questioned whether recent recommendations to utilize sibling fixed effects models (e.g. Morris et al., 2020) may actually induce error with polygenic scores, including whether it introduces bias to account for genetic nurture effects on early environments. Again, however, we were able to test relationships both with and without sibling fixed effects. Even when not employing fixed effects, the relationship between the education polygenic score and later life cognitive functioning, is explained by accounting for adolescent cognition; in addition to the sibling fixed effects results, this is a novel empirical finding. The bigger story holds regardless of model choice; the education polygenic score is not a good explanation for the relationship between educational attainment and later life cognitive functioning.

Another limitation is that the WLS is a dominantly white sample of high school graduates, though it is the case that just over 20 percent of the WLS sample grew up in poor households (Hauser & Sweeney, 1997). The findings should be understood in this context. That said, sample homogeneity is advantageous in this case, given the use of the education polygenic score. Because of potential issues with population stratification, the more homogenous sample reduces the risk for confounding. A reduced risk for unobserved variable bias in a more homogenous sample generally holds true for the analysis as a whole. That said, it's critical to test these questions in more diverse samples, as well as make broader investments in diverse samples that actually allow for this kind of analysis.

It's also worth noting that this is a single cohort, with a modal birth year of 1939. While a single cohort can bring advantages, in terms of ruling out potential confounders, these relationships may not hold for other cohorts. One key cohort change is increasing educational attainment. Indeed, there is some evidence that dementia rates may be falling as a result (Langa et al., 2017). Because this study only includes those with a high school degree, in some ways the WLS looks more like younger cohorts. But it's also the case that those with college degrees in this cohort are somewhat more select given broader access to post-secondary schooling for more recent cohorts. More generally, there is evidence that the influence of the education polygenic score on educational attainment changes across cohorts, with some evidence of an increasing influence due to increasing access to higher education (Branigan and Freese 2013; Herd et al., 2019). In short, removing structural barriers to higher education allows for genetic influence to translate into higher educational attainment (Herd et al., 2019). That said, Ding et al. (2019) did not find any differences across cohorts in the relationship between the education polygenic score and later life cognitive functioning. Nonetheless, exploring cohort differences may provide novel insights into the mechanisms underlying these relationships.

Finally, while sibling models allow us to broadly test the influence of early life environments, without additional observable data that varies within sibling pairs, we can't unpack those early life environments, such as the influence of socioeconomic resources versus emotionally nurturing environments. Indeed, the more limited variation present within sibling pairs is a more generic challenge with this modeling approach.

5. Conclusion

A growing body of evidence has documented that educational attainment is a robust predictor of cognitive functioning in later life. But the antecedents of educational attainment, which may influence both the propensity to attain higher levels of schooling, as well as independently influence later life cognitive functioning, are less well understood, as are midlife health behaviors and chronic conditions as mediators. The findings here document that genetic factors that influence educational attainment do not appear to play a significant role in shaping later life cognitive functioning. We do find, however, that adolescent cognitive functioning does play a role in shaping the influence of educational attainment on later life cognitive functioning, as well as acting as an important independent influence. Nonetheless, education remains robustly associated with later life cognitive functioning even after accounting for key antecedent influences and midlife health behavior and chronic condition mediators.

Author statement

Pamela Herd: Conceptualization, Methodology, Writing. **Kamil Sicinski**: Visualization, Software, Data Analysis.

Ethical statement

We have no financial and personal relationships with other people or organizations that could inappropriately influence or bias work.

Declaration of competing interest

We have no conflict of interest or financial issues to disclose.

Acknowledgements

Funding: This work was supported by R01AG041868-01A1, National Institute on Aging (NIA), A Longitudinal Resource for Genetic Research in Behavioral and Social Sciences and R01AG060737, National Institute on Aging (NIA), Wisconsin Longitudinal Study-Initial Lifetime's Impact on Alzheimer's Disease and Related Dementias (WLS-ILIAD Study).

APPENDIX

Table A1
Delayed Recall Performance at age ~72 Regressed on the Early and Mid Life Course Covariatees.

	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect
Early Life Course								
Education Polygenic Score	0.021	0.050**	-0.012	-0.004	-0.025	-0.015	-0.035	-0.020
	(0.040)	(0.021)	(0.040)	(0.021)	(0.040)	(0.021)	(0.040)	(0.022)
Adolescent IQ Score			0.160***	0.202***	0.139***	0.172***	0.130***	0.168***
			(0.034)	(0.021)	(0.035)	(0.023)	(0.036)	(0.023)
Years of education					0.032**	0.031***	0.031**	0.027***
					(0.015)	(0.010)	(0.015)	(0.010)
Mid Life Course Health Behaviors and Health								
Smoking (Reference=Never)								
Former smoker							0.066	0.088*
							(0.066)	(0.045)
Current smoker							-0.082	-0.062
							(0.111)	(0.083)
BMI (reference=Normal Weight)								
Underweight '92 (BMI<18.5)							-0.252	-0.402
							(0.378)	(0.288)
Overweight '92 (25.0<=BMI<=29.9)							-0.015	-0.027
							(0.070)	(0.050)
Obese '92 (BMI>=30.0)							-0.058	-0.080
							(0.092)	(0.065)
Exercise (Reference=None)								
Exercised >3 times per week ('92)							0.024	0.083
							(0.086)	(0.064)
Exercised 1 or 2 times per week ('92)							0.004	0.060
							(0.089)	(0.063)
Exercised one to three times per month ('92)							0.101	0.119**
							(0.082)	(0.059)
Chronic Conditions								
Had stroke?							-0.293**	-0.236**
							(0.141)	(0.097)
Has diabetes?							-0.065	-0.039
							(0.089)	(0.060)
Has hypertension?							0.057	0.048
							(0.065)	(0.044)
Had heart problems?							-0.066	-0.048
							(0.069)	(0.050)
Female	0.431***	0.428***	0.415***	0.418***	0.436***	0.439***	0.442***	0.448***
	(0.054)	(0.042)	(0.054)	(0.041)	(0.055)	(0.041)	(0.061)	(0.045)
Age in 2011 (difference from mean)	-0.036***	-0.040***	-0.032***	-0.035***	-0.032***	-0.035***	-0.031***	-0.034***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.005)
_cons	-0.265***	-0.263***	-0.278***	-0.284***	-0.736***	-0.719***	-0.738***	-0.733***
	(0.030)	(0.030)	(0.029)	(0.029)	(0.217)	(0.145)	(0.239)	(0.161)
r2	0.076	0.076	0.093	0.112	0.096	0.116	0.110	0.127
N	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000

Standard errors in parentheses.
Dependent variable is the delayed recall score in 2011.
Dependent variable is standardized using 2011 data.
* p<0.10, ** p<0.05, *** p<0.01
Adolescent IQ is standardized. Educational attainment is not. But the education coefficients included in the figures are standardized.

Table A2
Immediate Recall Performance at age ~72 Regressed on the Early and Mid Life Course Covariatees.

	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect
Early Life Course								
Education Polygenic Score	-0.000	0.078***	-0.032	0.018	-0.060	-0.003	-0.071*	-0.011
	(0.036)	(0.020)	(0.036)	(0.020)	(0.037)	(0.020)	(0.037)	(0.021)
Adolescent IQ Score			0.152***	0.227***	0.109***	0.176***	0.095***	0.170***
			(0.033)	(0.020)	(0.034)	(0.022)	(0.034)	(0.022)
Years of education					0.067***	0.054***	0.065***	0.050***
					(0.014)	(0.010)	(0.014)	(0.010)
Mid Life Course Health Behaviors and Health								
Smoking (Reference=Never)								
Former smoker							0.075	0.073*
							(0.056)	(0.042)
Current smoker							-0.132	-0.074
							(0.096)	(0.075)
BMI (reference=Normal Weight)								
Underweight '92 (BMI<18.5)							-0.327	-0.270
							(0.431)	(0.291)
Overweight '92 (25.0<=BMI<=29.9)							-0.142**	-0.072
							(0.064)	(0.048)
Obese '92 (BMI>=30.0)							-0.061	-0.102*
							(0.084)	(0.061)
Exercise (Reference=None)								
Exercised >3 times per week ('92)							0.089	0.086
							(0.075)	(0.060)
Exercised 1 or 2 times per week ('92)							0.016	0.004
							(0.082)	(0.059)
Exercised one to three times per month ('92)							0.008	0.078
							(0.077)	(0.057)
Chronic Conditions								
Had stroke?							-0.492***	-0.432***
							(0.129)	(0.103)
Has diabetes?							0.017	0.007
							(0.082)	(0.057)
Has hypertension?							0.002	0.010
							(0.058)	(0.043)
Had heart problems?							-0.087	-0.033
							(0.065)	(0.048)
Female	0.412***	0.401***	0.396***	0.390***	0.439***	0.426***	0.425***	0.418***
	(0.053)	(0.040)	(0.053)	(0.039)	(0.054)	(0.039)	(0.057)	(0.041)
Age in 2011 (difference from mean)	-0.034***	-0.040***	-0.030***	-0.033***	-0.029***	-0.033***	-0.027***	-0.032***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.004)	(0.006)	(0.005)
_cons	-0.194***	-0.192***	-0.207***	-0.217***	-1.154***	-0.975***	-1.056***	-0.909***
	(0.029)	(0.030)	(0.029)	(0.029)	(0.208)	(0.139)	(0.218)	(0.151)
r2	0.077	0.080	0.094	0.129	0.111	0.142	0.139	0.159
N	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000
rho	0.410		0.397		0.397		0.402	

Standard errors in parentheses.
Dependent variable is the immediate recall score in 2011.
Dependent variable is standardized using 2011 data.
* p<0.10, ** p<0.05, *** p<0.01
Adolescent IQ is standardized. Educational attainment is not. But the education coefficients included in the figures are standardized.

Table A3
Digit Ordering Task Performance at age ~72 Regressed on the Early and Mid Life Course Covariatees.

	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect
Early Life Course								
Education Polygenic Score	0.069*	0.115***	0.012	0.035*	-0.013	0.022	-0.015	0.026
	(0.038)	(0.022)	(0.037)	(0.021)	(0.038)	(0.021)	(0.038)	(0.021)
Adolescent IQ Score			0.279***	0.302***	0.240***	0.270***	0.228***	0.264***
			(0.032)	(0.020)	(0.033)	(0.022)	(0.034)	(0.022)
Years of education					0.060***	0.033***	0.062***	0.033***
					(0.014)	(0.010)	(0.014)	(0.010)
Mid Life Course Health Behaviors and Health								
Smoking (Reference=Never)								
Former smoker							0.114*	0.140***
							(0.062)	(0.043)
Current smoker							0.068	0.161**
							(0.100)	(0.080)
BMI (reference=Normal Weight)								
Underweight '92 (BMI<18.5)							-0.055	0.011
							(0.182)	(0.134)
Overweight '92 (25.0<=BMI<=29.9)							0.055	0.058
							(0.062)	(0.048)
Obese '92 (BMI>=30.0)							0.153*	0.124**
							(0.088)	(0.059)
Exercise (Reference=None)								
Exercised >3 times per week ('92)							0.036	0.046
							(0.087)	(0.065)
Exercised 1 or 2 times per week ('92)							0.086	0.088
							(0.081)	(0.062)
Exercised one to three times per month ('92)							0.100	0.066
							(0.078)	(0.054)
Chronic Conditions								
Had stroke?							-0.258**	-0.257***
							(0.122)	(0.083)
Has diabetes?							-0.012	-0.008
							(0.080)	(0.056)
Has hypertension?							0.032	0.049
							(0.061)	(0.044)
Had heart problems?							-0.041	-0.095**
							(0.066)	(0.048)
Female	0.025	0.051	-0.003	0.036	0.035	0.059	0.071	0.088**
	(0.055)	(0.041)	(0.053)	(0.039)	(0.053)	(0.039)	(0.057)	(0.042)
Age in 2011 (difference from mean)	-0.030***	-0.032***	-0.023***	-0.024***	-0.022***	-0.024***	-0.020***	-0.021***
	(0.006)	(0.004)	(0.005)	(0.004)	(0.006)	(0.004)	(0.006)	(0.004)
_cons	0.045	0.026	0.021	-0.006	-0.832***	-0.469***	-1.002***	-0.635***
	(0.029)	(0.033)	(0.029)	(0.031)	(0.206)	(0.140)	(0.220)	(0.151)
r2	0.026	0.035	0.083	0.121	0.096	0.126	0.112	0.140
N	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000	2264.000
N_clust	1131.000	1131.000	1131.000	1131.000	1131.000	1131.000	1131.000	1131.000
rho	0.409		0.395		0.400		0.399	

Standard errors in parentheses.
Dependent variable is the digit ordering score in 2011.
Dependent variable is standardized using 2011 data.
* p<0.10, ** p<0.05, *** p<0.01

Adolescent IQ is standardized. Educational attainment is not. But the education coefficients included in the figures are standardized.

Table A4
Letter Fluency Performance at age ~72 Regressed on the Early and Mid Life Course Covariatees.

	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect	Fixed Effect	No Fixed Effect
Early Life Course								
Education Polygenic Score	0.061*	0.104***	0.023	0.038**	-0.011	0.008	-0.017	0.006
	(0.033)	(0.020)	(0.033)	(0.019)	(0.033)	(0.019)	(0.033)	(0.019)
Adolescent IQ Score			0.168***	0.251***	0.120***	0.182***	0.113***	0.178***
			(0.030)	(0.019)	(0.031)	(0.020)	(0.031)	(0.020)
Years of education					0.073***	0.076***	0.072***	0.076***
					(0.012)	(0.009)	(0.012)	(0.009)
Mid Life Course Health Behaviors and Health								
Smoking (Reference=Never)								
Former smoker							0.058	0.141***
							(0.053)	(0.037)
Current smoker							0.013	0.094
							(0.102)	(0.071)
BMI (reference=Normal Weight)								
Underweight '92 (BMI<18.5)							-0.215	0.037
							(0.338)	(0.315)
Overweight '92 (25.0<=BMI<=29.9)							-0.092*	-0.045
							(0.056)	(0.041)
Obese '92 (BMI>=30.0)							-0.033	-0.035
							(0.075)	(0.050)
Exercise (Reference=None)								
Exercised >3 times per week ('92)							0.012	-0.030
							(0.075)	(0.055)
Exercised 1 or 2 times per week ('92)							0.008	-0.045
							(0.072)	(0.052)
Exercised one to three times per month ('92)							0.073	0.032
							(0.069)	(0.051)
Chronic Conditions								
Had stroke?							-0.250**	-0.199***
							(0.103)	(0.074)
Has diabetes?							-0.026	-0.097**
							(0.073)	(0.049)
Has hypertension?							-0.035	-0.045
							(0.052)	(0.038)
Had heart problems?							0.012	0.037
							(0.058)	(0.043)
Female	0.274***	0.270***	0.259***	0.264***	0.308***	0.317***	0.298***	0.315***
	(0.047)	(0.037)	(0.047)	(0.036)	(0.047)	(0.036)	(0.050)	(0.038)
Age in 2011 (difference from mean)	-0.031***	-0.031***	-0.026***	-0.024***	-0.026***	-0.023***	-0.024***	-0.021***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)
_cons	-0.060**	-0.060**	-0.074***	-0.090***	-1.105***	-1.161***	-1.055***	-1.154***
	(0.026)	(0.029)	(0.026)	(0.027)	(0.175)	(0.128)	(0.192)	(0.138)
r2	0.053	0.054	0.073	0.115	0.093	0.139	0.101	0.149
N	2921.000	2921.000	2921.000	2921.000	2921.000	2921.000	2921.000	2921.000
N_clust	1457.000	1457.000	1457.000	1457.000	1457.000	1457.000	1457.000	1457.000
rho	0.407		0.392		0.388		0.386	

Standard errors in parentheses.
Dependent variable is the letter fluency score in 2011.
Dependent variable is standardized using 2011 data.
* p<0.10, ** p<0.05, *** p<0.01

Adolescent IQ is standardized. Educational attainment is not. But the education coefficients included in the figures are standardized.

Table A5
Adolescent IQ Regressed on the Education Polygenic Score.

	Fixed Effect	No Fixed Effect
Education Polygenic Score	0.229***	0.263***
	(0.029)	(0.020)
Female	0.087**	0.025
	(0.042)	(0.037)
Age in 2011 (difference from mean)	-0.029***	-0.030***
	(0.004)	(0.004)
_cons	0.079***	0.110***
	(0.022)	(0.030)
r2	0.069	0.081
N	2921.000	2921.000

Standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Note: Older participants have lower adolescent IQ scores both because of the Flynn effect (Flynn 2007), as well as because some members of the 1957 graduating class were born prior to 1939 and likely graduated later because of academic struggles. Indeed these participants also had lower levels of academic performance in high school.

References

Akimova, E., Breen, R., Brazel, D., & Mills, M. (2021). Gene-environment dependencies lead to collider bias in models with polygenic scores. *Scientific Reports, 11*(1), 1–9.

Anderson, E., Howe, L., Wade, K., Ben-Shlomo, Y., Hill, W., Deary, I., Sanderson, E., Zheng, J., Korologou-Linden, R., Stergiakouli, E., & Davey Smith, G. (2020). Education, intelligence and Alzheimer's disease. *International Journal of Epidemiology, 49*(4), 1163–1172.

Armstrong-Carter, E., Wertz, J., & Domingue, B. (2021). Genetics and child development. *Child Development Perspectives, 15*(1), 57–64.

Banks, J., & Mazzonna, F. (2012). The effect of education on old age cognitive abilities. *The Economic Journal, 122*(560), 418–448.

Beck, A., Franz, C., Xian, H., Vuoksimaa, E., Tu, X., Reynolds, C., Panizzon, M., McKenzie, R., Lyons, M., Toomey, R., & Jacobson, K. (2018). Mediators of the effect of childhood socioeconomic status on late midlife cognitive abilities. *Innovation in aging, 2*(1), igy003.

Beckles, G., & Chou, C. (2016). Disparities in the prevalence of diagnosed diabetes—United States, 1999–2002 and 2011–2014. *Morbidity and Mortality Weekly Report, 65*(45), 1265–1269.

Belsky, D., Domingue, B., Wedow, R., Arseneault, L., Boardman, J., Caspi, A., … Harris, K. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences, 115*(31), E7275–E7284.

Branigan, A. R., McCallum, K. J., & Freese, J. (2013). Variation in the heritability of educational attainment: An international meta-analysis. *Social Forces, 92*(1), 109–140.

Brooks-Gunn, J., Duncan, G., Klebanov, P., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology, 99*(2), 353–395.

Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F., Harrison, S., Vie, G., … Davies, N. (2020). Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nature Communications, 11*(1), 1–13.

Cesarini, D., & Visscher, P. (2017). Genetics and educational attainment. *Npj Science of Learning, 2*(1), 1–7.

Clouston, S., Kuh, D., Herd, P., Elliott, J., Richards, M., & Hofer, S. (2012). Benefits of educational attainment on adult fluid cognition. *International Journal of Epidemiology, 41*(6), 1729–1736.

Cook, C. J., & Fletcher, J. (2015). Can education rescue genetic liability for cognitive decline? *Social Science & Medicine, 127*, 159–170.

Cox, S., Dickie, D., Ritchie, S., Karama, S., Pattie, A., Royle, N., … Deary, I. (2016). Associations between education and brain structure at age 73 years, adjusted for age 11 IQ. *Neurology, 87*(17), 1820–1826.

Crimmins, E., Saito, Y., Kim, J., Zhang, Y., Sasson, I., & Hayward, M. (2018). Educational differences in the prevalence of dementia and life expectancy with dementia: Changes from 2000 to 2010. *Journal of Gerontology: Series B, 73*(suppl_1), S20–S28.

Deckers, K., Cadar, D., van Boxtel, M., Verhey, F., Steptoe, A., & Köhler, S. (2019). Modifiable risk factors explain socioeconomic inequalities in dementia risk. *Journal of Alzheimer's Disease, 71*(2), 549–557.

Ding, X., Barban, N., Tropf, F., & Mills, M. (2019). The relationship between cognitive decline and a genetic predictor of educational attainment. *Social Science & Medicine, 239*, Article 112549.

Elmaleh, D. R., Farlow, M. R., Conti, P. S., Tompkins, R. G., Kundakovic, L., & Tanzi, R. E. (2019). Developing effective Alzheimer's disease therapies: Clinical experience and future directions. *Journal of Alzheimer's Disease, 71*(3), 715–732.

Englund, M., Luckner, A., Whaley, G., & Egeland, B. (2004). Children's achievement in early elementary school: Longitudinal effects of parental involvement, expectations, and quality of assistance. *Journal of Educational Psychology, 96*(4), 723.

Flouri, E. (2006). Parental interest in children's education, children's self-esteem and locus of control, and later educational attainment. *British Journal of Educational Psychology, 76*(1), 41–55.

Freedman, V., Martin, L., Schoeni, R., & Cornman, J. (2008). Declines in late-life disability: The role of early-and mid-life factors. *Social Science & Medicine, 66*(7), 1588–1602.

Glymour, M., Kawachi, I., Jencks, C., & Berkman, L. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology & Community Health, 62*(6), 532–537.

Greenfield, E., Akincigil, A., & Moorman, S. (2020). Is college completion associated with better cognition in later life for people who are the least, or most, likely to obtain a bachelor's degree? *Journal of Gerontology: Series B, 75*(6), 1286–1291.

Greenfield, E., & Moorman, S. (2019). Childhood socioeconomic status and later life cognition. *Journal of Aging and Health, 31*(9), 1589–1615.

Greenfield, E., Moorman, S., & Rieger, A. (2020). Life course pathways from childhood socioeconomic status to later-life cognition. *Journal of Gerontology: Series B*. https://doi.org/10.1093/geronb/gbaa062

Hamad, R., Elser, H., Tran, D., Rehkopf, D., & Goodman, S. (2018). How and why studies disagree about the effects of education on health: A systematic review and meta-analysis of studies of compulsory schooling laws. *Social Science & Medicine, 212*, 168–178.

Hauser, R., & Sweeney, M. (1997). Does adolescent poverty affect the life chances of high school graduates? In G. J. Duncan, & J. Brooks-Gunn (Eds.), *(1997). Consequences of growing up poor*. Russell Sage Foundation.

Hayward, M., Farina, M., Zhang, Y., Kim, J., & Crimmins, E. (2021). *The importance of improving educational attainment for dementia prevalence trends from 2000 to 2014, among older non-Hispanic Black and white Americans*. The Journals of Gerontology: Series B.

Herd, P., Carr, D., & Roan, C. (2014). Cohort profile: Wisconsin longitudinal study (WLS). *International Journal of Epidemiology, 43*(1), 34–41.

Herd, P., Freese, J., Sicinski, K., Domingue, B., Mullan Harris, K., Wei, C., & Hauser, R. (2019). Genes, gender inequality, and educational attainment. *American Sociological Review, 84*(6), 1069–1098.

Herd, P., Goesling, B., & House, J. (2007). Socioeconomic position and health: The differential effects of education versus income on the onset versus progression of health problems. *Journal of Health and Social Behavior, 48*(3), 223–238.
Herd, P., Mills, M., & Dowd, J. (2021). Reconstructing sociogenomics research: Dismantling biological race and genetic essentialism. *Journal of Health and Social Behavior, 62*(3), 419–435.
Judd, N., Sauce, B., Wiedenhoeft, J., Tromp, J., Chaarani, B., Schliep, A., … Klingberg, T. (2020). Cognitive and brain development is independently influenced by socioeconomic status and polygenic scores for educational attainment. *Proceedings of the National Academy of Sciences, 117*(22), 12411–12418.
Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., Benonisdottir, S., Oddsson, A., Halldorsson, B. V., Masson, G., & Gudbjartsson, D. F. (2018). The nature of nurture. *Science, 359*(6374), 424–428.
Langa, K. M., Larson, E. B., Crimmins, E. M., Faul, J. D., Levine, D. A., Kabeto, M. U., & Weir, D. R. (2017). A comparison of the prevalence of dementia in the United States in 2000 and 2012. *JAMA Internal Medicine, 177*(1), 51–58.
Lee, Y., Cho, Y., & Oh, H. (2021). Multiple chronic conditions and risk of cognitive impairment and dementia among older Americans. *Aging, Neuropsychology, and Cognition, 28*(4), 493–507.
Lee, J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics, 50*(8), 1112–1121.
Liu, Y., & Lachman, M. (2020). Education and cognition in middle age and later life: The mediating role of physical and cognitive activity. *Journal of Gerontology: Series B, 75*(7), e93–e104.
Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., … Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet commission. *The Lancet, 396*(10248), 413–446.
Luster, T., & McAdoo, H. (1996). Family and child influences on educational attainment: A secondary analysis of the high/scope Perry Preschool data. *Developmental Psychology, 32*(1), 26.
Marshall, I., Wang, Y., Crichton, S., McKevitt, C., Rudd, A., & Wolfe, C. (2015). The effects of socioeconomic status on stroke risk and outcomes. *The Lancet Neurology, 14*(12), 1206–1218.
Mills, M., & Tropf, F. (2020). Sociology, genetics, and the coming of age of sociogenomics. *Annual Review of Sociology, 46*, 553–581.
Moorman, S., Carr, K., & Greenfield, E. (2018). Childhood socioeconomic status and genetic risk for poorer cognition in later life. *Social Science & Medicine, 212*, 219–226.
Morris, T., Davies, N., Hemani, G., & Smith, G. (2020). Population phenomena inflate genetic associations of complex social traits. *Science Advances, 6*(16), Article eaay0328.
National Research Council. (2013). The future of the sociology of aging: An agenda for action. In *Committee on population: Division of behavioral and social Sciences and education*. Washington, DC: National Academy Press.
Nielsen, F. (2006). Achievement and ascription in educational attainment: Genetic and environmental influences on adolescent schooling. *Social Forces, 85*(1), 193–216.
Pudrovska, T., Logan, E., & Richman, A. (2014). Early-life social origins of later-life body weight: The role of socioeconomic status and health behaviors over the life course. *Social Science Research, 46*, 59–71.
Scarr, S., & McCartney, K. (1983). *How people make their own environments: A theory of genotype→ environment effects* (pp. 424–435). Child development.
Schneeweis, N., Skirbekk, V., & Winter-Ebmer, R. (2014). Does education improve cognitive performance four decades after school completion? *Demography, 51*(2), 619–643.
Sewell, W. (1971). Inequality of opportunity for higher education. *American Sociological Review, 36*, 793–809.
Stefanidis, K., Askew, C., Greaves, K., & Summers, M. (2018). The effect of non-stroke cardiovascular disease states on risk for cognitive decline and dementia: A systematic and meta-analytic review. *Neuropsychology Review, 28*(1), 1–15.
Walsemann, K. M., & Ailshire, J. A. (2020). Early educational experiences and trajectories of cognitive functioning among US adults in midlife and later. *American Journal of Epidemiology, 189*(5), 403–411.
Ware, E., & Faul, J. (2021). Genomic data measures and methods: A primer for social scientists. In *Handbook of aging and the social Sciences* (pp. 49–62). Academic Press.
Watson, C., & Klett, W. (1975). The Henmon-Nelson, Cardall-Miles, slosson and quick tests as predictors of WAIS IQ. *Journal of Clinical Psychology*.
Wertz, J., Belsky, J., Moffitt, T., Belsky, D., Harrington, H., Avinun, R., Poulton, R., Ramrakha, S., & Caspi, A. (2019). Genetics of nurture. *Developmental Psychology, 55*(7), 1461.
Zhang, Z., Liu, H., & Choi, S. (2020). Early-life socioeconomic status, adolescent cognitive ability, and cognition in late midlife. *Social Science & Medicine, 244*, Article 112575.