

Cohort Profile

Cohort Profile: High School and Beyond

Eric Grodsky,¹ Jennifer Manly,² Chandra Muller³ and
John Robert Warren ^{4,*}

¹Department of Sociology, University of Wisconsin—Madison, Madison, Wisconsin, USA, ²Division of Neurology, Columbia University, New York, New York, USA, ³Department of Sociology, University of Texas at Austin, Austin, Texas, USA and ⁴Minnesota Population Center, University of Minnesota—Twin Cities, Minneapolis, Minnesota, USA

* Corresponding author. Minnesota Population Center, University of Minnesota, Twin Cities, 909 Social Science, 267 19th Ave South, Minneapolis, MN 55455, USA. E-mail: warre046@umn.edu

Received 10 September 2021; Editorial decision 14 February 2022; Accepted 25 February 2022

Why was the cohort set up?

High School and Beyond (HSB) was launched with funding from the U.S. Department of Education's National Center for Education Statistics (NCES) as part of its Secondary Longitudinal Studies Program. Sociologist James S. Coleman directed the design of the study. From 1980 through 1992, the purpose of HSB was to document the educational and labour force development of young people and to study their development as they entered post-secondary institutions, the work force, the military and adult family life.

HSB sample members—all of whom were high school sophomores (10th grade of U.S. secondary school) or seniors (12th and final year of US secondary school) in the spring of 1980—occupy an important position at the end of the American Baby Boom cohort. They are the first cohort to enter adult lives after the American Civil Rights movements: after it became normative for women's educational attainments to exceed those of men's; after it became normative for women to work in paid jobs without interruption for child rearing; and after the decline of generous pensions and affordable health insurance for most workers. The HSB cohort is more racially and ethnically diverse than earlier cohorts, in part because it was the first to come of age after the U.S. Immigration and Nationality Act of 1965.

From 2014 onward, however, the purpose of HSB has been to document the long-run relationship between

education and sample members' mid-life health, cognitive status, mortality, labour force status, economic status, family and other outcomes. Of particular interest has been assessing how early life educational and other contexts influence: mid-life cognitive functioning and impairment; longevity; health; and employment and economic outcomes. No other large, diverse, nationally representative US cohort study includes such a wealth of information about early life—especially educational—contexts and detailed measures of later life outcomes.

Who is in the cohort?

In 1980, HSB students were selected through a two-stage stratified probability sample with schools as the first-stage units and students as the second-stage units.¹ Except for special strata—e.g. alternative schools, high-performing private schools, predominantly Cuban schools, Catholic schools—schools were selected with probability proportional to enrolment in a stratified way across regions and urban/rural areas. Within each school, 36 seniors and 36 sophomores were randomly selected; if there were fewer than 36 seniors or 36 sophomores, all eligible students were selected. As designed, the sample included 1120 schools (from a frame of 26 100 schools with grades 10 or 12 or both). Substitution was carried out for schools that refused to participate in HSB, but there was no substitution for students who could or would not participate. In the

Key Features

- High School and Beyond (HSB) began in 1980 to study how educational opportunities and experiences shape early adult outcomes, and includes a nationally representative probability sample of 30 030 sophomores and 28 240 seniors from 1020 randomly selected US public and private high schools.
- Data on students' educational experiences, cognitive skills, non-cognitive skills, peers, educational and occupational plans and aspirations, health and socioeconomic background were obtained via student, parent, school administrator and teacher questionnaires.
- From the initial sample of 58 270 students, a random subsample of 14 830 sophomores and 12 000 seniors were selected to participate in a longitudinal panel.
- A follow-up of the 25 370 surviving panelists is being fielded in 2021. Questionnaires gather information about: cognitive functioning and impairment; memory complaints; health conditions; work; family; finances; COVID-19 incidence and vaccination; and wellbeing. Sample members complete several cognitive tasks commonly employed in studies of ageing. Home health visits take anthropomorphic measures and collect saliva samples (for genomic analysis of the microbiome) and whole-blood samples (for human genomic analysis and for assaying markers of neurodegeneration).

end, the achieved sample in 1980 included 1020 schools and 58 270 participating students (including 30 030 sophomores and 28 240 seniors). See [Table 1](#) for more details.

From the initial sample of 58 270 students in 1980, a random subset of 26 830—including 14 830 sophomores and 12 000 seniors—has been re-interviewed on multiple occasions since 1982. By design, this longitudinal cohort sample included disproportionate numbers of base-year sample members from policy-relevant subpopulations (e.g. high-achieving racial and ethnic minorities, students from Catholic and other private high schools, high school dropouts, and students planning to pursue post-secondary schooling). Carefully constructed sampling weights allow researchers to produce statistical estimates that reflect the population from which students were selected. Nearly all panelists were born between 1962 and 1965, thus being between 56 and 59 years old in 2021.

How often have they been followed up?

All panel members were re-surveyed in 1982, 1984, and 1986; sophomores were re-surveyed in 1992 and 2014; and seniors were re-surveyed in 2015. As the time of this writing in 2021, both sophomores and seniors are being re-interviewed and invited to participate in home health visits.

The HSB surveys have had remarkably high response rates—ranging from ~90% in the 1980s to ~65% in the 2014–15 follow-ups^{2,3}; see [Tables 1](#) and [2](#) for more details. As shown in [Table 1](#), there are differentials in response rate by socioeconomic status, race/ethnicity, and nativity such that less advantaged people, the foreign-born, and members of racial/ethnic minority groups are less likely to

respond—but nonetheless response rates through 1992 were universally and remarkably high. Response rates were somewhat lower in 2014/15—although by current standards they were quite good. Several methodology reports from 1980 onward describe differential response rates and the construction of corresponding panel weights (e.g. 2, 3).

What has been measured?

As shown in [Table 3](#), the 1980 student questionnaires gathered data on sophomores' and seniors' educational experiences, educational and occupational plans and aspirations, health and disability, demographic attributes, family socioeconomic background, student friendships and more. Both cohorts completed standardized multiple-choice assessments of reading, vocabulary and mathematics. Sophomores also completed assessments in writing, science and civics, and seniors completed a paired associate test of short-term memory, mosaic recognition assessments of general cognitive ability and a spatial relations assessment.

Although not described in [Table 3](#), the 1980 survey wave also included a parent survey, a school administrator survey and a teacher survey. The parent questionnaire—administered to only a sample of participating students—primarily gathered information about family attitudes toward and financial planning for students' post-secondary educations; it also included parental reports of students' school activities and experiences. School questionnaires gathered information about enrolment, demographics, staffing, educational programmes, school control, school finances, facilities and services, dropout

Table 1 Sociodemographic variables, by response status and by survey wave

	1980 (Base year)			1982			1984			1986			1992			2014-2015		
	R	NR ^a	RR ^b	R	NR ^a	RR ^b	R	NR ^a	RR ^b	R	NR ^a	RR ^b	R	NR ^a	RR ^b	R	NR ^a	RR ^b
Grade in base year																		
Sophomore	30 030	0	100%	28 120	1620	95%	13 680	1140	92%	13 430	1400	91%	12 640	2190	85%	8790	6030	59%
Senior	28 240	500	98%	11 230	770	94%	10 930	1070	91%	10 540	1460	88%	n/a			6930	5070	58%
Sex																		
Male	27 820	260	99%	17 870	1160	94%	11 250	1190	90%	10 870	1570	87%	5700	1070	84%	6850	5590	55%
Female	29 480	240	99%	19 190	850	96%	12 490	810	94%	12 230	1070	92%	6150	830	88%	8360	4950	63%
Missing	970	0	100%	2290	380	86%	860	210	80%	860	220	80%	800	280	74%	510	570	47%
Race																		
Black	7500	50	99%	6060	420	94%	4340	470	90%	4160	660	86%	1470	400	79%	2440	2370	51%
White (non-Latinx)	39 910	330	99%	22 350	990	96%	12 770	820	94%	12 560	1030	92%	7320	860	89%	8840	4750	65%
Latinx	7000	30	100%	6320	290	96%	5430	500	92%	5200	730	88%	2050	520	80%	2930	2560	53%
Other	2190	90	96%	1260	260	83%	1200	210	85%	1180	230	84%	560	130	81%	750	660	53%
Missing	1680	0	100%	3370	430	89%	860	210	80%	860	220	80%	1230	280	81%	750	760	50%
Mother's education																		
Less than high school	9640	0	100%	6410	350	95%	4430	350	93%	4290	480	90%	1990	340	85%	2640	2140	55%
High school	24 940	0	100%	15 240	710	96%	9550	690	93%	9310	920	91%	4800	700	87%	6220	4010	61%
Some college	6500	0	100%	3920	190	95%	2470	180	93%	2430	220	92%	1230	170	88%	1740	910	66%
4+ years of college	6920	0	100%	4290	180	96%	2640	200	93%	2600	240	92%	1450	140	91%	1930	920	68%
Missing	10 270	500	95%	9500	960	91%	5520	790	87%	5320	990	84%	3160	840	79%	3190	3130	50%
School location																		
Urban	13 260	120	99%	8780	590	94%	5950	630	90%	5720	860	87%	2650	590	82%	3620	2960	55%
Suburban	28 110	270	99%	17 750	990	95%	11 410	960	92%	11 110	1260	90%	6020	920	87%	7520	4850	61%
Rural	16 600	110	99%	10 570	440	96%	6380	410	94%	6270	520	92%	3190	400	89%	4060	2730	60%
Missing	290	0	100%	2240	380	85%	860	210	80%	860	220	80%	800	280	74%	510	570	47%
Born in the USA																		
Yes	54 090	0	100%	33 830	1680	95%	21 240	1630	93%	20 740	2130	91%	10 900	1620	87%	13 800	9070	60%
No	3300	0	100%	2320	200	92%	1760	230	88%	1660	330	83%	790	220	78%	1020	970	51%
Missing	880	500	64%	3200	510	86%	1610	360	82%	1560	400	80%	950	350	73%	900	1070	46%
Total	58 270	500	99%	39 350	2390	94%	24 610	2210	92%	23 960	2860	89%	12 640	2190	85%	15 720	11 100	59%

All sample sizes rounded to the nearest 10 as per the terms of our restricted data use agreement.

R, respondent; NR, non-respondent; RR, response rate.

^aNon-respondents include: (i) those eligible to respond who did not do so; and (ii) those ineligible to respond (e.g. because of death or institutionalization).

^bResponse rates include ineligible sample members as non-respondents and should be interpreted with caution. See [Table 2](#) for more informative response rates.

Table 2 Summary of data collection, 1980–2021

Survey wave	Cohort (modal age)	Data collection method	Target sample size ^a	Achieved sample size ^a	Response rate	
1980	Sophomores (15)	In-school student questionnaire	33 930	28 240	83%	
		In-school achievement test	33 930	27 070	80%	
	Seniors (17)	In-school student questionnaire	36 770	30 030	82%	
		In-school achievement test	36 770	25 570	70%	
	Schools (n/a)	Mail-back principal questionnaire ^b	1020	1000	98%	
	Teachers (n/a)	Mail-back comment form ^c	70 700	35 347	50%	
1982	Sophomores (17)	In-person student questionnaire	29 740	28 120	95%	
		In-person achievement test	29 740	26 220	88%	
	Seniors (19)	Mail-back questionnaire, in-person interview, or telephone interview	12 000	11 230	94%	
		Schools (n/a)	Mail-back principal questionnaire	990	970	98%
	1984	Sophomores (19)	Mail-back questionnaire, in-person interview, or telephone interview	14 830	13 680	92%
			Seniors (21)	Same as for sophomores in 1984	12 000	10 930
1986	Sophomores (21)	Mail-back questionnaire, in-person interview, or telephone interview	14 830	13 430	91%	
		Seniors (23)	Same as for sophomores in 1986	12 000	10 540	88%
1992	Sophomores (27)	Telephone interview	14 830	12 640	85%	
2014 ^d	Sophomores (48)	Telephone interview, internet questionnaire, or mail-back questionnaire	14 070	8790	62%	
2015 ^d	Seniors (51)	Telephone interview, internet questionnaire, or mail-back questionnaire	11 300	6930	61%	
2021	Sophomores (56)	Telephone interview, internet questionnaire, or mail-back questionnaire AND in-home health visit		[Now in the field]		
	Seniors (58)	Same as for sophomores in 2021		[Now in the field]		

^aAll sample sizes are rounded to the nearest 10 as per the terms of our NCES restricted data use agreement. Response rates are calculated based on the rounded (not the more precise) sample sizes.

^bThe target sample size for schools pertains to those schools that agreed to participate. A total of 1120 schools were invited to participate; after substitution, 1020 participated.

^cFor the 1980 teacher comment form, the target and achieved sample sizes refer to the numbers of students for whom teacher comments were sought and obtained (not the numbers of teachers targeted or responding).

^dFor the 2014 and beyond survey waves, the target sample excludes deceased and institutionalized individuals.

rates, college-going rates and school and district policies. The teacher survey—designed to obtain at least one teacher report pertaining to each HSB sample member—provided teachers an opportunity to comment on HSB sample members' school performance, popularity and likelihood of success in higher education and in the labour market. A second school questionnaire was administered in 1982. As shown in [Table 3](#), follow-up surveys of students conducted in the 1980s and in 1992 gathered information about cohort members' educational, employment and family statuses, activities and transitions.

The 2014 and 2015 surveys—conducted when most sample members were in their early 50s—gathered mid-life data on health, work, family, finances and educational outcomes; see [Table 3](#) for more details. The 2021 survey—being conducted when most sample members are

approaching age 60—is gathering objective data on cognitive functioning and subjective cognitive concerns, health conditions, work, finances, family, COVID 19 and science knowledge; see [Table 4](#) for more details. The 2021 fieldwork also includes an in-person home health visit at which anthropometric measures are gathered and blood and saliva collected. Most of the blood is stored for future analysis, but assays for markers of neurodegeneration will be conducted first. The blood and saliva are also used to produce genomic measures that characterize humans and their oral microbiomes.

As shown in [Figure 1](#), HSB records have been or soon will be linked to a variety of administrative and commercial data sources, including: secondary and post-secondary school enrolments and transcripts; mortality records; consumer credit data; real-estate transaction records; voter

Table 3 Survey content, by wave and cohort, 1980 through 2014/15

Year	1980		1982		1984		1986		1992	2014/15	
	So (15)	Sr (17)	So (17)	Sr (19)	So (19)	Sr (21)	So (21)	Sr (23)	So (27)	So ^a (48)	Sr (51)
Cohort (So = 1980 sophomore; Sr = 1980 senior) (Modal age)											
Education											
Educational attainment	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Secondary curriculum and courses	✓	✓	✓								
Educational plans and aspirations	✓	✓	✓	✓	✓	✓			✓		
Secondary school grades and achievements	✓	✓	✓								
Secondary school behaviour and discipline	✓	✓	✓								
Secondary school extracurricular activities	✓	✓	✓								
Post-secondary plans	✓	✓	✓	✓	✓	✓			✓		
Post-secondary enrolment, major, courses				✓	✓	✓	✓	✓	✓	✓	✓
Significant others' educational expectations	✓	✓	✓	✓							
Friends' educational attributes	✓	✓	✓								
Cognition; cognitive and non-cognitive skills											
Reading, mathematics, vocabulary tests	✓	✓	✓								
Writing, science, civics tests	✓										
Memory, comparisons, spatial reasoning tests		✓									
Self-esteem, locus of control	✓	✓	✓				✓	✓			
Labour market; family											
Current labour market activities; jobs; income	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Occupational plans and aspirations	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Military enlistment and experiences	✓	✓	✓	✓	✓	✓					
Marital status; number of children				✓	✓	✓	✓	✓	✓	✓	✓
Health											
Health; morbidity; disabilities; height/weight	✓	✓	✓				✓	✓		✓	✓
Health risk behaviours				✓			✓	✓			
Mortality; cause of death (from the National Death Index)										✓	✓
Background/demographic											
Family socioeconomic background	✓	✓	✓	✓							
Demographic attributes; nativity; language	✓	✓	✓	✓							

^aSome sophomores were given a longer survey. This table only describes the content of survey items given to all sophomores and seniors.

registration and turnout data; state tumour registries; and pharmacy records. These administrative data are available for analyses, although the manner in which they can be accessed varies across source.

What has it found?

Prior to 2014, HSB survey and administrative data were extensively used for academic and policy research on issues related to education, schooling, and resulting socioeconomic outcomes in early adulthood. Several hundred academic articles, books and dissertations on issues related to education appeared in sociology, management, business, education, economics, political science, planning development, family studies, urban studies, social work, public administration, health care, health policy and other fields. A comprehensive list of publications is maintained at [<http://sites.utexas.edu/hsb/publications/>].

In the area of education, HSB data have most notably been used for studies related to the roles that schools play in: educational outcomes⁴⁻⁶; the development of students' non-cognitive skills^{7,8}; the correlates of dropping out of school⁹⁻¹¹; racial/ethnic¹²⁻¹⁷ and gender^{18,19} disparities in schooling opportunities and outcomes; the associations among extracurricular activities, paid work and academic performance in high school²⁰⁻²²; and the associations among family structure, parental involvement, related variables and children's educational outcomes.²³⁻²⁷ With respect to early adult socioeconomic outcomes, HSB has been used for research on: the role of social capital in the creation of human capital²⁸; the association between cognitive skills and wages and other labour market outcomes^{29,30}; the associations between non-cognitive skills and wages and other labour market outcomes^{31,32}; and the associations between school attributes and earnings and labour market outcomes.³³ Pre-2014 HSB data have been

Table 4 Survey content, 2021 follow-up, by mode

	Phone survey	Web survey	Paper survey	Proxy survey	Home visit
Education					
Educational attainment	✓	✓	✓	✓	
Post-secondary institution, major field	✓	✓	✓		
Cognitive functioning and impairment; biomarkers for ADRD risk					
Immediate recall (CERAD word list)	✓				
Semantic fluency (animal naming)	✓				
Phonemic fluency (F task)	✓				
Delayed recall (CERAD word list)	✓				
Working memory (digit span, forward and backward)	✓				
Memory and learning (verbal and visual paired associates)		✓			
Self-reported memory complaints (AD8)	✓	✓	✓	✓	
APOE e4 and GWAS (from saliva or blood)					✓
Markers of neuropathology (Aβ40, Aβ42, tau, NFL, p-tau 181, and GFAP from blood)					✓
Health					
Self-assessed overall adult health	✓	✓	✓	✓	
Self-assessed overall childhood health	✓	✓	✓		
COVID-19 testing, infection, vaccination	✓	✓	✓	✓	✓
Pain	✓	✓	✓		
Cervical, colon and breast cancer screening	✓	✓			
Self reported diagnoses of cancer, diabetes, hypertension, stroke, coronary heart disease, kidney disease, periodontal disease and mental health condition	✓	✓	✓	✓	
Opioid use	✓	✓	✓	✓	
Height and weight	✓	✓	✓		✓
Blood pressure, pulse					✓
Waist circumference					✓
Health risk behaviours (smoking, alcohol use)	✓	✓	✓		
Depression	✓	✓	✓		
Loneliness	✓	✓	✓		
Mortality; cause of death (from National Death Index)					
Labour market					
Labour force status, number of jobs	✓	✓	✓	✓	
Industry and occupation	✓	✓	✓	✓	
Income	✓	✓	✓		
Other measures					
Marital status	✓	✓	✓	✓	
Science knowledge	✓	✓	✓		
Oral microbiome collection and sequencing					✓

Sample members are asked to complete either the telephone or web survey; those who refuse are eventually offered the paper survey. Proxy surveys are for sample members who are unable to complete a survey themselves. All sample members completing telephone, web or paper surveys are invited to complete a home health visit for anthropometric measures and blood and saliva collection; those refusing home health visits are mailed a saliva collection kit.

ADRD, Alzheimer's disease and related dementia; CERAD, Consortium to Establish a Registry for Alzheimer's Disease; AD8, an eight-item Alzheimer's Disease screener; APOE, apolipoprotein E; GWAS = genome-wide association study.

used for research in areas outside these key areas, such as on alcohol use,³⁴ family studies,^{35–38} obesity³⁹ and teenage pregnancy.^{40,41}

Ageing and health outcomes were included in the 2014–15 round of data collection, and the data have been used for: studies related to the relationship between high school (especially mathematics) coursework and mid-life health outcomes⁴²; the association between failing to meet

adolescent occupational expectations and early mortality via 'deaths of despair'⁴³; the association between taking science, technology, engineering and mathematics [STEM]-related courses in high school and later life occupational outcomes⁴⁴; the importance of the source of mortality information on inferences about mortality disparities⁴⁵; the association between educational attainment and inequality and mid-life noncognitive skills⁴⁶; and the relationship

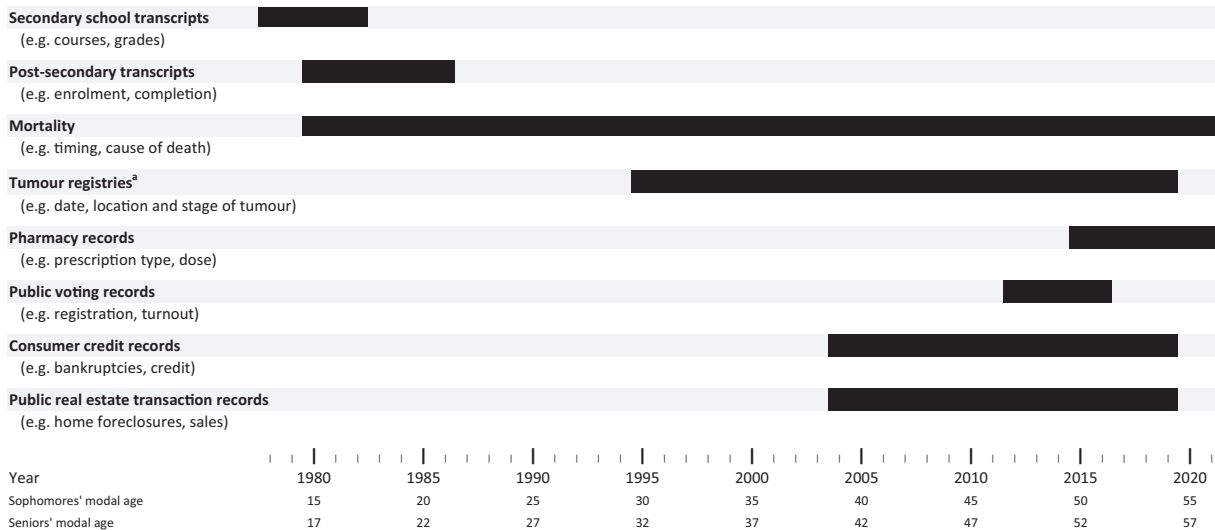


Figure 1 Administrative record data linked (or soon to be linked) to high school and beyond, by year

^aYears of availability of tumour registry data vary across states.

between dimensions of the process of schooling (beyond attainment) and mortality outcomes.⁴⁷

What are the main strengths and weaknesses?

HSB is among just a few prospective cohort studies in the USA that has followed a nationally-representative—and thus highly diverse—sample of people from adolescence through later adulthood. The design and content of the study make it invaluable for studying the ways in which the process and outcomes of education ‘get under the skin’ to shape later-life cognition, health and other outcomes. Since the sample was originally clustered in over 1000 high schools, it may be valuable for understanding place-based processes, including segregation and exposure to environmental toxicants, which can shape long-run health trajectories. Its mix of survey, administrative and biomarker data allows for careful analyses of the factors that stratify core ageing outcomes among contemporary Americans. Other strengths of the study include: its relatively high response rates; the strength of the sampling designs; and the unique historical context of the cohort included in HSB.

Another virtue of HSB is the potential for cross-cohort comparisons: it is one of several cohort studies in NCES’s Secondary Longitudinal Studies Program. Its predecessor—the National Longitudinal Study of the High School Class of 1972—will (pending funding) be followed up soon using similar protocols as HSB. Its successors—the National Educational Longitudinal Study of 1988, the Educational

Longitudinal Study of 2002, the High School Longitudinal Study of 2009 and the new High School and Beyond 2020 cohort—all could potentially be followed up in comparable ways in the future.

HSB also has weaknesses for some purposes. Sample members come from one narrow range of birth cohorts, so analyses of inter-cohort trends are not generally possible using HSB data alone. All sample members completed at least 10th grade, so people who left formal schooling earlier—e.g. through high school dropout, institutionalization or early death—are not represented. Likewise, people who immigrated to the USA after completing secondary school are not represented. Because of the original purpose of the study, information about some early life cognitive and health domains—e.g. childhood disease, adolescent mental health—is somewhat limited. The surveys conducted in the early 1980s include limited information about health risk behaviours (e.g. smoking, drug use, exercise). Finally, since the survey was not conducted for many years, there is a large gap in temporal coverage between when sample members were in their early (seniors) to late (sophomores) twenties and when they were in mid-life (although as shown in [Figure 1](#), some data are available in intervening years).

Can I get hold of the data? Where can I find out more?

Survey data and documentation through the 2014–15 follow-ups are available to qualified researchers for research

purposes at no cost through NCES's Restricted-use License Program; see [<https://nces.ed.gov/pubsearch/licenses.asp>]. The procedure to obtain a licence includes submitting: (i) a formal request document signed by the principal researchers and the senior official of the organization; (ii) a statement listing the requested database, the research goals and data use, the sectors of the community that will be served and assurance that the data will not be used for administrative or regulatory purposes; (iii) a data security plan which typically requires that data be stored on a stand-alone computer with no internet connections; (iv) a signed and notarized affidavit of non-disclosure for each individual who will have access to the data; and (v) the estimated period during which the data will be used (typically 5 years and renewable).

Genomic data from the 2021 HSB survey will be available from the National Institute on Aging (NIA) Genetics of Alzheimer's Disease Data Storage Site (NIAGADS); see [<https://www.niagads.org/home>]. Blood-based biomarker data (e.g. A β 40, A β 42, total tau, NfL, p-tau 181, GFAP, DNA methylation) will be available through the Global Alzheimer's Association Interactive Network (GAAIN); see [<http://gaaain.org>]. In both cases, a unique and anonymous code will be used to link sample members' genomic data or blood-based biomarker data to their survey information at NCES; only NCES will hold the cross-walk between these data files. To access linked survey and genomic and/or blood-based biomarker records, researchers will have to: apply for access to restricted survey data from NCES; apply for access to biomarker data from NIAGADS and/or GAAIN; and analyse the linked records in a secure facility.

Ethics approval

The protocols for the 2021 wave of HSB data collection were reviewed and approved by the Institutional Review Boards of the University of Minnesota (approval STUDY00009650) and NORC at the University of Chicago (approval 21–02-148).

Author contributions

All four authors were co-leaders of the project to design and implement the 2021 wave of HSB fieldwork and are recipients of funding support for that effort. All four contributed to drafting and editing this manuscript. J.R.W. acts as guarantor for the article.

Funding

This work and the 2021 wave of HSB data collection were supported by the National Institute on Aging of the U.S. National Institutes of Health (grant number R01AG058719) and by the Alzheimer's Association (grant number SG-20717567). The 2014–

15 wave of HSB data collection was supported by: the Alfred P. Sloan Foundation (grant number 2012–10-27); the U.S. National Science Foundation (grant number HRD 1348527 to C.M.; grant number HRD 1348557 to J.W.; grant number DRL 1420691 to C.M.; grant number DRL 1420330 to E.G.; and grant number DRL 1420572 to J.W.); the U.S. Institute on Education Sciences (grant number R305U140001 to C.M. and grant number R305U180002 to C.M.); and the Spencer Foundation (grant number 201500075 to C.M. and grant number 20160116 to J.W.).

Acknowledgements

The order of authorship is alphabetical. We would like to thank: Nicole Schmidt, Elizabeth Gibson, and Isabella Stade for their tireless and careful work on the 2021 wave of HSB data collection; Elizabeth Johnson and Robert Reynolds for their outstanding work on HSB since the 2014–15 wave of HSB data collection; Elia Boschetti for careful research assistance; and the survey team at NORC at the University of Chicago for their intellectual contributions.

We also appreciate support provided by¹ the Eunice Kennedy Shriver National Institute for Child Health and Human Development of the U.S. National Institutes for Health to the University of Texas at Austin's Population Research Center (grant number P2CHD042849), the University of Wisconsin-Madison's Center for Demography and Ecology (grant number P2CHD047873), and the University of Minnesota's Minnesota Population Center (grant number P2CHD041023) and² the National Institute on Aging of the U.S. National Institutes of Health to the University of Texas at Austin's Center on Aging and Population Sciences (grant number P30AG066614), the University of Wisconsin's Center for Demography of Health and Aging (grant number P30AG017266), and the University of Minnesota's Life Course Center (grant number P30AG066613).

Conflict of interest

None declared.

References

1. National Center for Education Statistics. *High School and Beyond, 1980: A Longitudinal Survey of Students in the United States*. Washington, DC: National Center for Education Statistics, U.S. Department of Education, 1981.
2. Muller C, Black SE, Grodsky E, Warren JR. *High School and Beyond Sophomore Cohort Fifth Follow-up Documentation. NCES 2019–138*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2019.
3. National Center for Education Statistics. *High School and Beyond, 1980: Senior Cohort Third Follow-up (1986)*. Washington, DC: U.S. Dept. of Education, National Center for Education Statistics, 1988.
4. Gamoran A. The stratification of high school learning opportunities. *Sociol Educ* 1987;60:135–55.
5. Lee VE, Bryk AS. A multilevel model of the social distribution of high school achievement. *Sociol Educ* 1989;62:172–92.
6. Raudenbush SW, Rowan B, Kang SJ. A multilevel, multivariate model for studying school climate with estimation via the EM

- algorithm and application to US high-school data. *J Educ Stat* 1991;16:295–330.
7. Bryk AS, Lee VE, Holland PB. *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press, 1993.
 8. Coleman JS, Hoffer T. *Public and Private High Schools: The Impact of Communities*. New York, NY: Basic Books, 1987.
 9. McCaul EJ, Donaldson GA Jr, Coladarci T, Davis WE. Consequences of dropping out of school: findings from high school and beyond. *J Educ Res* 1992;85:198–207.
 10. McNeal RB Jr. High school dropouts: a closer examination of school effects. *Soc Sci Quart* 1997;78:209–22.
 11. Pittman RB. Social factors, enrollment in vocational/technical courses, and high school dropout rates. *J Educ Res* 1991;84:288–95.
 12. Lucas SR, Good AD. Race, class, and tournament track mobility. *Sociol Educ* 2001;74:139–56.
 13. Madhere S. Convergence and divergence in the process of academic development for Black, White, and Hispanic high school students. *J Educ Stud Placed Risk* 1997;2:137–60.
 14. Morgan SL. Trends in black-white differences in educational expectations: 1980–92. *Sociol Educ* 1996;69:308–19.
 15. Rivkin SG. Black/white differences in schooling and employment. *J Hum Resour* 1995;30:826–52.
 16. Tauber KE, James DR. Racial segregation among public and private schools. *Sociol Educ* 1982;55:133–43.
 17. Tauber KE, James DR. Racial segregation among public and private schools: a response. *Sociol Educ* 1983;56:204–07.
 18. Marsh HW. Sex differences in the development of verbal and mathematics constructs: the High School and Beyond study. *Am Educ Res J* 1989;26:191–225.
 19. Riegle-Crumb C, King B, Grodsky E, Muller C. The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *Am Educ Res J* 2012;49:1048–73.
 20. Marsh HW. Employment during high school: character building or a subversion of academic goals? *Sociol Educ* 1991;64:172–89.
 21. Marsh HW. Extracurricular activities: beneficial extension of the traditional curriculum or subversion of academic goals? *J Educ Psychol* 1992;84:553–62.
 22. Marsh HW, Kleitman S. Consequences of employment during high school: character building, subversion of academic goals, or a threshold? *Am Educ Res J* 2005;42:331–69.
 23. Astone NM, McLanahan SS. Family structure, parental practices and high school completion. *Am Sociol Rev* 1991;56:309–20.
 24. Astone NM, McLanahan SS. Family structure, residential mobility, and school dropout: a research note. *Demography* 1994;31:575–84.
 25. Marsh HW. Two-parent, stepparent, and single-parent families: changes in achievement, attitudes, and behaviors during the last two years of high school. *J Educ Psychol* 1990;82:327–40.
 26. Milne AM, Myers DE, Rosenthal AS, Ginsburg A. Single parents, working mothers, and the educational achievement of school children. *Sociol Educ* 1986;59:125–39.
 27. Zimiles H, Lee VE. Adolescent family structure and educational progress. *Dev Psychol* 1991;27:314–20.
 28. Coleman JS. Social capital in the creation of human capital. *Am J Sociol* 1988;94:595–120.
 29. Murnane RJ, Willett JB, Levy F. The growing importance of cognitive skills in wage determination. *Rev Econ Stat* 1995;lxxvii:251–66.
 30. Murnane RJ, Willett JB, Tyler JH. Who benefits from obtaining a GED? Evidence from high school and beyond. *Rev Econ Stat* 2000;82:23–37.
 31. Deluca S, Rosenbaum JE. Individual agency and the life course: do low-SES students get less long-term payoff for their school efforts? *Sociological Focus* 2001;34:357–76.
 32. Rosenbaum J. Are noncognitive behaviors in school related to later life outcomes? In: Rosenbaum J, ed. *Beyond College for All: Career Paths for the Forgotten Half*. New York, NY: Russell Sage, 2001.
 33. Eide ER, Showalter MH, Sims DP. The effects of secondary school quality on the distribution of earnings. *Contempor Econ Policy* 2002;20:160–70.
 34. Martin MJ, Pritchard ME. Factors associated with alcohol use in later adolescence. *J Stud Alcohol* 1991;52:5–9.
 35. Goldscheider F, Goldscheider C. Whose nest? A two-generational view of leaving home during the 1980s. *J Marriage Fam* 1993;55:851–62.
 36. Goldscheider FK, Goldscheider C. Family structure and conflict: nest-leaving expectations of young adults and their parents. *J Marriage Fam* 1989;51:87–97.
 37. Goldscheider FK, Goldscheider C. The intergenerational flow of income: family structure and the status of Black Americans. *J Marriage Fam* 1991;53:499–508.
 38. O'hare WP. Gonna get married: family formation in rural areas. *Popul Today* 1991;19:6–9.
 39. Faith MS, Manibay E, Kravitz M, Griffith J, Allison DB. Relative body weight and self-esteem among African Americans in four nationally representative samples. *Obes Res* 1998;6:430–37.
 40. Hanson SL, Myers DE, Ginsburg AL. The role of responsibility and knowledge in reducing teenage out-of-wedlock childbearing. *J Marriage Fam* 1987;49:241–56.
 41. Parnell AM, Swicegood G, Stevens G. Nonmarital pregnancies and marriage in the United States. *Soc Forces* 1994;73:263–87.
 42. Carroll JM, Muller C, Grodsky E, Warren JR. Tracking health inequalities from high school to midlife. *Soc Forces* 2017;96:591–628.
 43. Muller C, Duncombe A, Carroll JM, Mueller AS, Warren JR, Grodsky E. Association of job expectations among high school students with early death during adulthood. *JAMA Netw Open* 2020;3:e2027958.
 44. Black SE, Muller C, Spitz-Oener A, He Z, Hung K, Warren JR. The importance of STEM: High school knowledge, skills and occupations in an era of growing inequality. *Res Policy* 2021;50:104249.
 45. Warren JR, Milesi C, Grigorian K, Humphries M, Muller C, Grodsky E. Do inferences about mortality rates and disparities vary by source of mortality information? *Ann Epidemiol* 2017;27:121–27.
 46. Smith CM, Grodsky E, Warren JR. Late-stage educational inequality: can selection on noncognitive skills explain waning social background effects? *Res Soc Strat Mobil* 2019;63:100424.
 47. Warren JR, Muller C, Hummer RA, Grodsky E, Humphries M. Which aspects of education matter for early adult mortality? Evidence from the high school and beyond cohort. *Socius* 2020;6:237802312091808.