GROWUPGROWOLD
HUMAN DEVELOPMENT, BIRTH TO DEATH

Moffitt & Caspi
Genes
Environment
Health
Behavior

Concept Paper Form

Provisional Paper Title: Height, Cognition, and Education: Evidence from E-Risk
Proposing Author: Vikesh Amin
Author's Email: amin1v@cmich.edu
P.I. Sponsor: Terrie Moffit (if the proposing author is a student or colleague of an original PI)
Today's Date: 2/2/2020

Please describe your proposal in 2-3 pages with sufficient detail for helpful review.

Objective of the study:

Research in labor economics has shown that there is a positive association between height and wages (see Hübler 2016 for a review). The two most prominent explanations are related to cognitive ability and non-cognitive/social skills. Persico et al. (2004) argue that the height-wage premium is due to non-cognitive/social skills. They argue that taller children have greater access to clubs or social activities that develop productive human capital, which leads to higher wages. In contrast, Case & Paxson (2008) argue that the height-wage premium arises because height is associated with cognitive ability, which is rewarded in the labor market. They find that height at age 3 (even before schooling has started) is positively associated with cognition, and that controlling for childhood cognition reduces the height-wage premium by about 30-50 percent.

There are several studies that have estimated associations between height and cognition and non-cognitive skills (Mosca & Wright 2016; Guven & Lee 2013; Magnusson et al. 2006; Maurer 2010; Schick & Steckel 2015). However, it is difficult to estimate the causal effect of height because of confounding by unobserved genetic and early-life environmental factors. The heritability of height is estimated to be about 80% (Visscher et al. 2008), and height is also likely to be influenced early-life factors such as childhood nutrition. If unobserved genetic and early life factors that affect height are also associated with cognitive ability and non-cognitive skills, then associations are biased estimates of the causal relationship. Only a few studies that have estimated the causal effect of height on cognitive ability or non-cognitive skills. Von Hinke Kessler Scholder et al. (2013) use Mendelian randomization, where they instrument childhood height in the Avon Longitudinal Study of Parents & Children with genetic markers that predict height. The linear regression results show that contemporaneous height is associated higher IQ (at age 8), better school performance (at age 14) and lower likelihood of having behavioral problems (at age 13). In contrast, the instrumental variable results show that height is associated with higher IQ for girls (but not

boys) and a higher probability of having behavioral problems. Case & Paxson (2010) use within-sibling comparisons to estimate the effect of height on IQ in the National Longitudinal Study 1979 Children and Young Adults. Their within-sibling estimates show that height is associated with better cognitive test scores. The within-sibling estimates were also smaller in absolute magnitude than the corresponding linear regression estimates, indicating that standard associations are upward biased due to confounding from unobserved family level factors.

The objective of the project is to study the association between height and (i) cognition, (ii) non-cognitive skills, and (iii) educational attainment using the E-Risk Study. We will use data on polygenic scores (PGS) and discordant twin pair designs to help determine whether the associations reflect causal relationships by (i) conducting a reduced-form analysis of the impact of being genetically predisposed to being tall on the outcomes and (ii) how OLS associations between height and the outcomes change when controlling for unmeasured genetic and family factors (through PGSs, neighborhood and family background characteristics).

Data analysis methods:

Our goal is to assess how robust standard OLS associations are to confounding from unobserved genetic and family factors. We take two different approaches to do this. First, we use PGS, which are summary measures of an individual's genetic predisposition for a given trait, to conduct a reduced form analysis of the impact of being genetically predisposed to being tall on the outcomes. As genes are randomly inherited at conception conditional on parental genotype, the height PGS is a more exogenous measures of height than observed height. The reduced form analysis provides a test for the presence of an effect of height, which can be informative as to whether a causal relationship exists. We will perform the following regressions:

1. A linear regression of the outcomes on a PGS for height, PGS for educational attainment, and control variables for demographics, family, neighborhood SES (or neighborhood fixed effects), and principal components of the genetic data to control for population stratification.

The coefficients on the PGSs from regression (1) will not reflect a pure genetic effect, as it is confounded by family environment. For example, parental genetics will influence the family environment provided to children, which in turn affects child outcomes. The PGS therefore reflects the influence of both genes and family environment. To account for this we will perform regression (2), which utilizes variation within twin pairs, thereby eliminating the influence of shared family factors.

2. A within-twins regression for same-sex DZ twins that relates twin pair differences in outcomes to twin pair differences in the PGSs.

Second, we use the PGSs as controls for unobserved genetic factors in OLS regressions that estimate associations between height and cognition and educational attainment. We will also compare how directly controlling for genetics compare to results from within-twins regressions. The specific regressions to be performed are:

3. A linear regression of outcomes on height, a PGS for height, PGS for educational attainment, and control variables for demographics, family, neighborhood SES (or neighborhood fixed effects), and principal components of the genetic data to control for population stratification
4. A within-twins regression for MZ twins that relates twin pair differences in the outcomes to twin pair differences in the height.
5. A within-twins regression for same-sex DZ twins that relates twin pair differences in the outcomes to twin pair differences in the height, and the PGSs.

Variables needed at which ages:

Concept	**Variable(s)**
Demographic controls	Age, gender, race
Genetics	PGSs for height and educational attainment, and principal components to control for population stratification.
Outcome variables	Educational Attainment: measures at age 18
	Cognition: measures at ages 5, 12, and 18
	Social/non-cognitive skills: are there any measures available at ages 5, 12, and 18?
Explanatory variable	Height: at ages 5, 12, and 18
Socioeconomic controls	Composite measures of family/neighborhood SES

Significance of the Study (for theory, research methods or clinical practice):

While height is associated with better cognition, non-cognitive skills, and education it is difficult to estimate the causal effect. By making use of genetic data and discordant twin pair designs, the project will provide new evidence on the extent to which associations in the literature are likely confounded by unobserved genetic and family factors.

References cited:

Schick, A., & Steckel, R. H. (2015). Height, human capital, and earnings: The contributions of cognitive and noncognitive ability. *Journal of Human Capital*, *9*(1), 94-115.

Case, A., & Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of political Economy*, *116*(3), 499-532.

Case, A., & Paxson, C. (2010). Causes and consequences of early-life health. *Demography*, *47*(1), S65-S85.

Guven, C., & Lee, W. S. (2013). Height and cognitive function at older ages: is height a useful summary measure of early childhood experiences?. *Health Economics*, *22*(2), 224-233.

Hübler, O. (2016). Height and wages. *Oxford Handbook of Economics and Human Biology. Oxford University Press, Oxford*.

Magnusson, P. K., Rasmussen, F., & Gyllensten, U. B. (2006). Height at age 18 years is a strong predictor of attained education later in life: cohort study of over 950 000 Swedish men. *International Journal of Epidemiology*, *35*(3), 658-663.

Maurer, J. (2010). Height, education and later-life cognition in Latin America and the Caribbean. *Economics & Human Biology*, *8*(2), 168-176.

Mosca, I., & Wright, R. E. (2016). Height and cognition at older ages: Irish evidence. *Economics Letters*, *149*, 98-101.

Persico, N., Postlewaite, A., & Silverman, D. (2004). The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy*, *112*(5), 1019-1053.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255-266.

Scholder, S. V. H. K., Smith, G. D., Lawlor, D. A., Propper, C., & Windmeijer, F. (2013). Child height, health and human capital: evidence using genetic markers. *European Economic Review*, *57*, 1-22.

Data Security Agreement

Provisional Paper Title: Height, Cognition, and Education: Evidence from E-Risk
Proposing Author: Vikesh Amin
Today's Date: 2/2/2020

☒	I am current on Human Subjects Training (CITI (www.citiprogram.org) or equivalent)
☒	My project is covered by the Duke ethics committee OR I have /will obtain ethical approval from my home institution.
☒	I will treat all data as "restricted" and store in a secure fashion. My computer or laptop is: a) encrypted (recommended programs are FileVault2 for Macs, and Bitlocker for Windows machines) b) password-protected c) configured to lock-out after 15 minutes of inactivity AND d) has an antivirus client installed as well as being patched regularly.
☒	I will not "sync" the data to a mobile device.
☒	In the event that my laptop with data on it is lost, stolen or hacked, I will immediately contact Moffitt or Caspi.
☒	I will not share the data with anyone, including my students or other collaborators not specifically listed on this concept paper.
☒	I will not post data online or submit the data file to a journal for them to post. *Some journals are now requesting the data file as part of the manuscript submission process. Study participants have not given informed consent for unrestricted open access, so we have a managed-access process. Speak to Temi or Avshalom for strategies for achieving compliance with data-sharing policies of journals.*
☒	I will delete all data files from my computer after the project is complete. Collaborators and trainees may not take a data file away from the office. This data remains the property of the Study and cannot be used for further analyses without an approved concept paper for new analyses.

Signature: Vikesh Amin

CONCEPT PAPER RESPONSE FORM

A

Provisional Paper Title	Height, Cognition, and Education: Evidence from E-Risk
Proposing Author	Vikesh Amin
Other Contributors	Click or tap here to enter text.
Potential Journals	Click or tap here to enter text.
Today's Date: Click here to enter date	
Intended Submission Date: Click here to enter date	

Please keep one copy for your records and return one to the proposing author

B. To be completed by potential co-authors:

☐	Approved
☐	Not Approved
☐	Let's discuss, I have concerns

Comments: Click here to enter text

Please check your contribution(s) for authorship:

☐	Conceptualizing and designing the longitudinal cohort study
☐	Conceptualizing data collection protocols and creating variables
☐	Data collection
☐	Conceptualizing and designing this specific paper project
☐	Statistical analyses and interpretation (or reproducibility check)
☐	Writing
☐	Reviewing manuscript drafts
☐	Final approval before submission for publication
☐	Agreement to be accountable for the work
☐	Acknowledgment only, I will not be a co-author

Signature: Click here to enter text