Title: Using Polygenic Scores for Educational Attainment to Probe Genetic Effects on Schooling Outcomes among Fragile Families

Authors: Adrienne D. Woods, Arianna Gard, Erin B. Ware, Colter Mitchell & Daniel Notterman

To Be Considered For:

1) 404. Gene-Environment Interaction and Health (Organizer: Jessica D. Faul)

Shorter Abstract (150 words):

Understanding how DNA influences life outcomes is important to researchers parsing the variability of social science phenomena. We linked the polygenic risk score for educational attainment (Belsky et al., 2016; Rietveld et al., 2013; Lee et al., 2018; Trejo et al., 2018) to schooling outcomes in the Fragile Families and Child Wellbeing Study, a project following an economically and racially diverse group of children from birth through age 15. Polygenic scores for children with European ancestry were significantly and positively associated with teacher-reported grades at age 5 and 9, and with child-reported grades at age 15. However, and most importantly, results differ by ancestry. Though the polygenic scores significantly predicted academic and cognitive outcomes for students of European ancestry, they predicted in-school behavioral outcomes for students with African or Hispanic ancestry. Overall, our results reveal a need for further exploration of this polygenic score, particularly among diverse ancestry groups.

Longer Abstract (4 pages):

Background. Genome-wide association studies (GWAS) analyze millions of genetic variants in hundreds of thousands of individuals to identify alleles associated with particular outcomes of interest. These alleles function together in a continuum that varies across individuals and can be used to predict social science outcomes (referred to as a polygenic score; Chabris, Lee, Cesarini, Benjamin, & Laibson, 2015; Belsky et al., 2016). Several recent studies have both created and reported polygenic scores that have been successfully linked to educational attainment (Belsky et al., 2016; Rietveld et al., 2013; Lee et al., 2018; Trejo et al., 2018). For instance, Belsky et al. (2016) demonstrated that an individual's polygenic score can predict lifespan development ranging from early academic skill acquisition through adult economic and social outcomes beyond educational attainment. They also noted that although children with higher polygenic scores were born into better-off homes and were more upwardly mobile through social classes, their polygenic scores still predicted adult outcomes when controlling for childhood social class.

Overall, understanding how DNA influences life outcomes is important to researchers parsing the variability of social science phenomena, and may one day be used to more precisely control for the causal effects of the environment in order to invest in maximally effective interventions (Harden, 2018). But, the mechanism(s) by which this link emerges require further exploration. Belsky and colleagues utilized the Dunedin Study, which followed about 1,000 New Zealanders of mainly European ancestry for four decades. Thus, we particularly need to understand how polygenic scores for educational attainment operate in a more diverse sample, and in a U.S. setting. To that end, the present study replicated Belsky et al.'s analysis using the Fragile Families and Child Wellbeing Study, which obtained data from an economically and racially diverse group of children and their families from birth through age 15. Ours is among the first studies that examines polygenic scores for educational attainment among such a sample, and is therefore important toward validating and refining this score for future research.

Method. We analyzed educational, cognitive, and behavioral data available from the age 5, age 9, and age 15 waves of data collection. Items are drawn from a variety of sources, including survey reports by children, parents, and teachers, and direct assessment of both cognitive and academic skill. Continuous outcomes were assessed with OLS regression to estimate standardized beta coefficients, and dichotomous outcomes with logistic regression to estimate odds ratios. The polygenic score was included as a predictor in analyses after being standardized to have a mean of 0 and a standard deviation of 1 within each ancestry group. All analyses were conducted separately by race/ethnicity (White $n = 455$, Black $n = 1264$, Hispanic $n = 664$) and analyses controlled for ancestry-specific principal components computed from GWAS data.

Results. Analyses are ongoing, but preliminary results reveal two important findings. First, our results replicate those of Belsky et al. (2016). Polygenic scores for children with European ancestry (hereafter, EA) are significantly and positively associated with teacher-reported grades at age 5 and 9, and with child-reported grades at age 15. We also find evidence of a gene-environment correlation among EA students, such that those with a higher PGS are

more likely to attend private school. Standardized betas range from .18 to .26 (see Tables 1-3). Second, and most important, results differ by ancestry. Though the polygenic scores significantly predict academic and cognitive outcomes for EA students, they also appear to predict in-school behavioral outcomes for students with African or Hispanic ancestry (hereafter, AA and HA, respectively). For instance, though the score significantly predicts educational outcomes at age 5 for EA children, there were no significant associations among these same measures for AA or HA children (see Table 1). At age 15, the score most strongly predicted grades for EA students, and the type of classes EA students report taking (whether honors or remedial). However, these associations were not evident for AA or HA students. Instead, there was a significant link to disciplinary behavior and grade retention among AA and HA students, suggesting that the "educational attainment" score may operate differently within non-EA samples. Overall, our results reveal a need for further exploration of the cognitive, academic, and behavioral outcomes that relate to this polygenic score, particularly among diverse ancestry groups. Researchers should take caution when generalizing this score, as it may not predict the same outcomes among people of non-EA ancestry.

Table 1.

Associations between polygenic score and cognitive and academic outcomes at age 5.

	European Ancestry			**African Ancestry**			**Hispanic Ancestry**		
	n	**Beta**	**OR**	**n**	**Beta**	**OR**	**n**	**Beta**	**OR**
PPVT [1]	236	.183**		782	-.050		255	.036	
WJ Test 22 [1]	238	.163*		789	.003		256	.148	
Academic Ratings [2]	158	.225**		303	.150		120	.058	
Works to best of ability [2]	158	.227**		301	.072		120	.124	
Special Education [2]	151		.45*	279		1.07	113		.37

Note: 1 = Direct Assessment, 2 = Teacher-Report, 3 = Child-Report, 4 = Parent-Report
***$p < .001$, **$p < .01$, *$p < .05$; Multiple comparisons corrected using the Benjamini-Hochberg method.

Table 2.

Associations between polygenic score and cognitive and academic outcomes at age 9.

	European Ancestry			African Ancestry			Hispanic Ancestry		
	n	**Beta**	**OR**	**n**	**Beta**	**OR**	**n**	**Beta**	**OR**
Digit Span [1]	366	.143**		1140	.081**		443	.075	
PPVT [1]	364	.250***		1138	.095**		443	.136**	
WJ Test 9 [1]	363	.261***		1132	.064		433	.119*	
Academic Ratings [2]	292	.163**		734	.076		292	.172**	
Works to best of ability [2]	291	.160*		731	.083		289	.072	
Special Education [2]	292		.89	729		1.07	290		1.01
Remedial Services [2]	293		.74	726		.89	289		.66**
Gifted/Talented Services [2]	290		1.04	722		1.64**	287		3.20**
Cheated on test [4]	366		1.03	1136		.87	442		1.24
Skipped School [4]	366		1.20	1137		.92	442		.86
Suspended/expelled [3]	328		.63	1003		.88	358		.61
Repeated Grades [3]	376		.66*	1152		.85	454		.51***

Note: 1 = Direct Assessment, 2 = Teacher-Report, 3 = Child-Report, 4 = Parent-Report
***$p < .001$, **$p < .01$, *$p < .05$; Multiple comparisons corrected using the Benjamini-Hochberg method.

Table 3.

Associations between polygenic score and cognitive and academic outcomes at age 15.

	European Ancestry			African Ancestry			Hispanic Ancestry		
	n	**Beta**	**OR**	**n**	**Beta**	**OR**	**n**	**Beta**	**OR**
English Grade [1]	368	.207***		1141	-.013		441	.078	
Mathematics Grade [1]	370	.122*		1149	.011		451	.126**	
History/Social Studies Grade [1]	361	.168***		1129	.027		429	.084	
Science Grade [1]	367	.141**		1138	.103***		445	.079	
Special Education [2]	371		.97	1147		.90	449		.55**
Failed a class [1]	375		.64***	1165		.89	452		.65**
Taken honors classes [1]	389		1.63***	1165		1.02	457		1.34
Remedial Classes [2]	366		.53**	1118		.86	433		.76
Gifted/Talented [2]	369		1.34	1139		1.11	441		1.35
Skipped School [1]	392		.75	1173		.91	457		.89
Suspended/expelled [2]	393		.74	1169		.76***	456		.63*
Repeated Grades since age nine [2]	392		.79	1169		.83	455		.50**

Note: 1 = Child-Report, 2 = Parent-Report

***$p < .001$, **$p < .01$, *$p < .05$; Multiple comparisons corrected using the Benjamini-Hochberg method.