

Problems with a Causal Interpretation of Polygenic Score Differences between Jewish and non-Jewish Respondents in the Wisconsin Longitudinal Study

Jeremy Freese, Ben Domingue, Kamil Sicinski, Sam Trejo, Pam Herd¹

Abstract

Dunkel et al. (2019) observe higher MTAG polygenic scores for educational attainment among the 53 Jewish respondents with available genomic data in the Wisconsin Longitudinal Study. They interpret their ensuing analysis as evidence that genetic differences "mediate" the association between being Jewish and higher cognitive test scores and higher educational attainment. We demonstrate instead--and perhaps counterintuitively--that the difference between Jewish and non-Jewish polygenic scores are much too *large* for their analysis to offer any evidentiary value for this conclusion. Instead, the data show clear evidence of the problems with comparing polygenic scores across ethnic groups that others have noted.

¹ Affiliations: Freese, Domingue, and Trejo are at Stanford University; Sicinski is at the University of Wisconsin-Madison; Herd is at Georgetown University. We thank Michael Nivard for comments on a previous draft.

By design, polygenic scores for educational attainment are associated with educational attainment (Lee et al. 2018), and they are also associated with cognitive test scores (Plomin and von Stumm 2018). Recently, a study using data from the Wisconsin Longitudinal Study (WLS) observed that Jewish respondents have a higher average value on a polygenic score for educational attainment than do non-Jewish respondents (Dunkel et al. 2019). Consistent with other data, Jewish respondents in WLS also have higher average cognitive test scores and educational attainment. This leads Dunkel et al. (2019) to report that the polygenic score "mediates" the relationship between being Jewish and the study's cognitive and educational outcomes.

The investigators' discussion invites the interpretation of these results as evidence of *causal* mediation: that is, that the study provides evidence that the genetic differences responsible for the polygenic score difference between Jewish and non-Jewish respondents cause the group differences in cognition and education. An obvious limitation of this study is that the sample includes only 53 Jewish respondents (less than 1% of WLS respondents). What we wish to clarify here, however, is that other differences between the Jewish and non-Jewish respondents in WLS provide further reasons to consider any causal interpretation of the study's results unfounded.

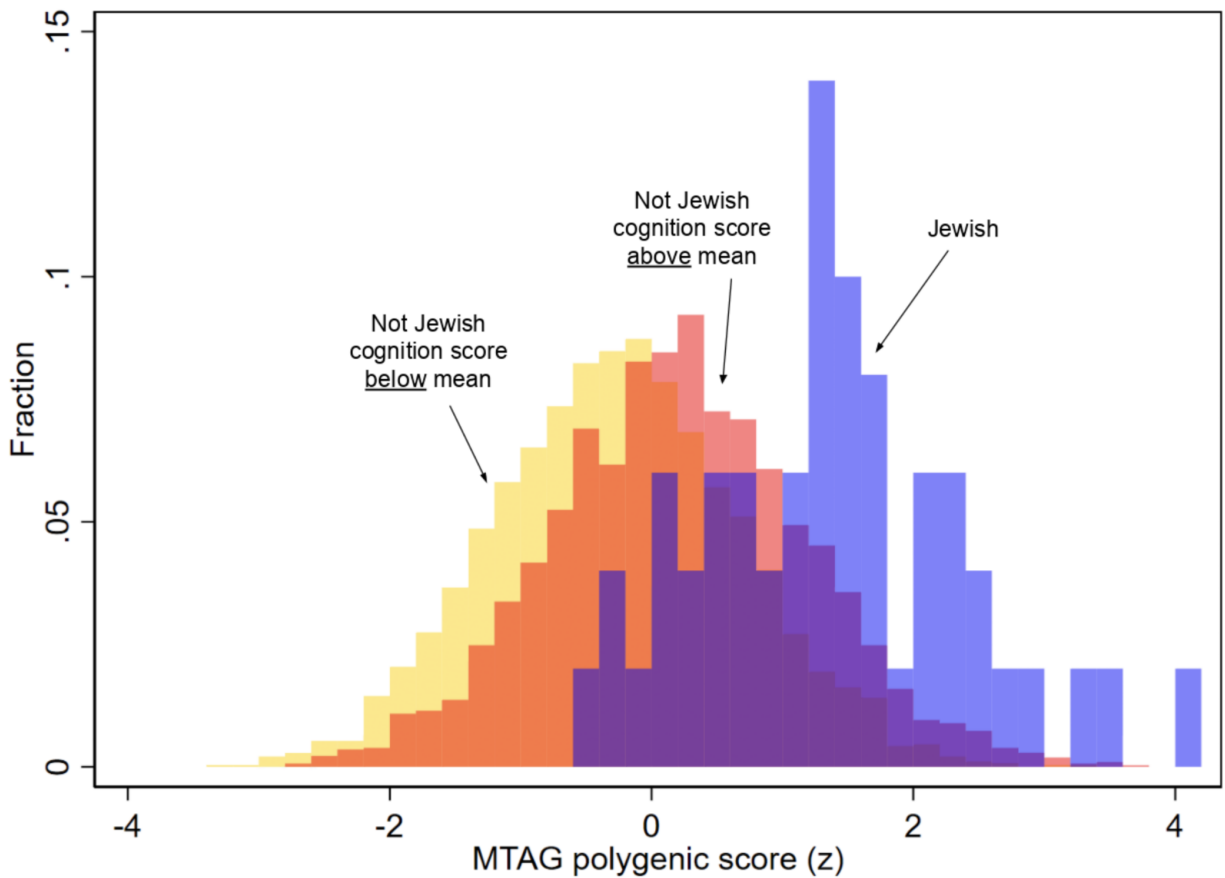
Martin et al. (2017) provide a compelling demonstration of the problems with drawing causal conclusions from polygenic score differences between ancestral groups. In one of their examples, polygenic score differences would predict that people of West African descent to be much shorter than people of European descent, even though we have robust anthropometric evidence of no such differences. The same inferential problem lurks in a more pernicious way when a phenotypic difference between groups does exist, because then the genetic differences might be mistakenly inferred to cause the phenotypic differences even when closer scrutiny reveals similar problems to those that Martin et al. (2017) highlight.

Figure 1 shows how this problem manifests for the results reported by Dunkel et al. (2019). We use the cognitive test that WLS respondents took in high school. The figure divides the non-Jewish respondents into two groups: those with test scores above the mean and those with test scores below the mean. We can see that the scores are approximately normally distributed in both groups, but that the distribution is shifted modestly upward for those with cognition scores above the mean, reflecting the plain but hardly overwhelming association between the polygenic score and cognition.

For Jewish respondents, however, the center of the polygenic score distribution is much higher. The difference between Jewish and non-Jewish respondents is several times larger than the difference in scores between those above and below the mean for non-Jewish

respondents. Notably, WLS Jewish respondents do have a higher cognitive test score than the overall mean ($z = .66$), it is lower than the mean test score among above-average WLS respondents ($z = .87$).

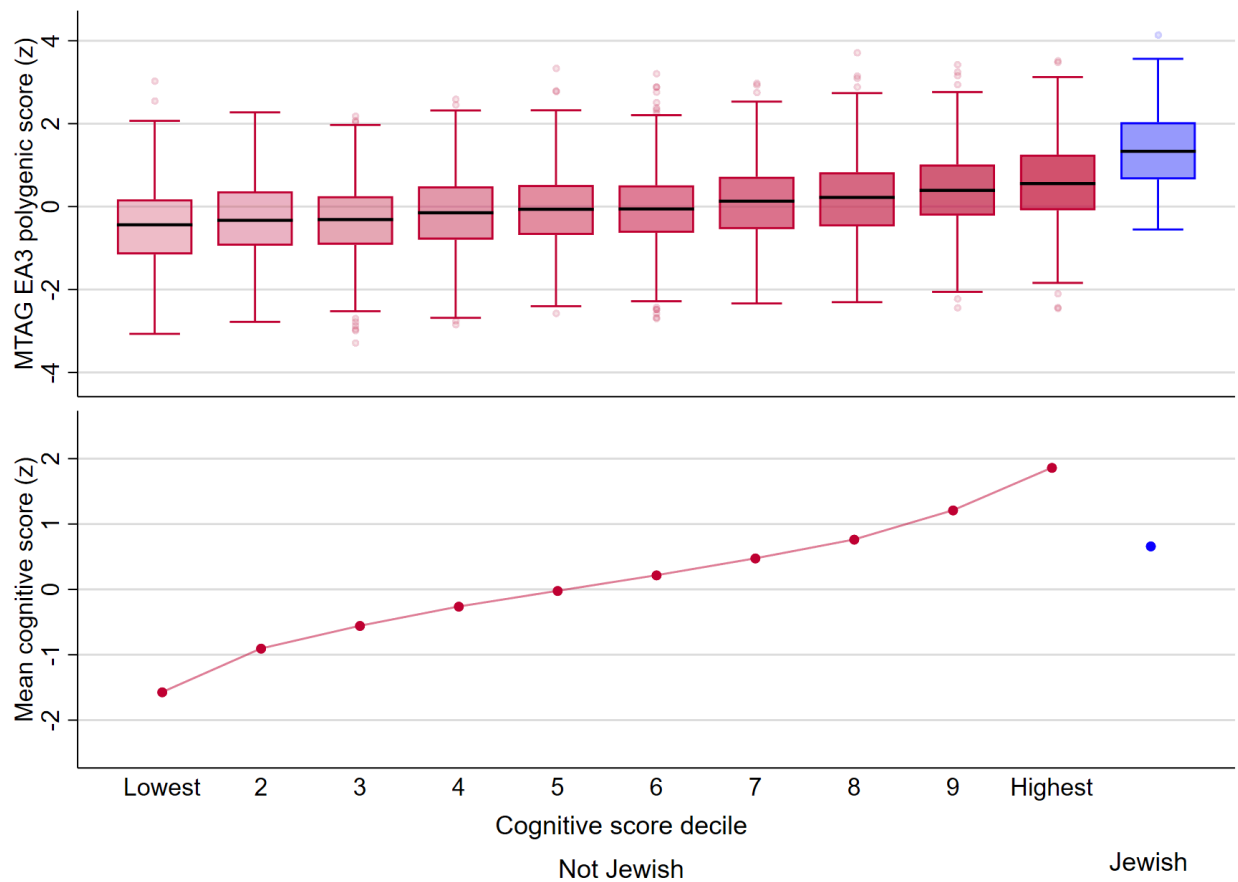
Figure 1. Histograms of MTAG polygenic score among WLS Jewish respondents and non-Jewish WLS respondents with European ancestry whose cognitive test scores are above vs. below the mean.



A different visualization of this point is provided as Figure 2. For non-Jewish respondents, the plot provides information about the distribution of MTAG scores and the mean cognitive score by the cognitive score decile. The top panel of the graph shows the monotonic increase in median MTAG score in successive deciles of the cognitive test score. The difference in the medians between the top decile and middle deciles, however, is still less than the difference between the median of the top decile for non-Jewish respondents and the median for Jewish respondents. (Indeed, one can show that Jewish respondents in WLS have a much higher average polygenic score than even the top cognitive *percentile* of

non-Jewish WLS respondents.) A naïve causal interpretation of the polygenic score difference might therefore lead one to anticipate that cognitive test scores for Jewish respondents would be not just above the mean for non-Jewish respondents, but routinely at genius levels. Instead, as noted, the average for Jewish respondents is only about two-thirds of an SD above the mean (or about 110 in the IQ metric), and even this shrinks to just below half an SD when we look at all respondents for whom we have 1975/7 survey data in WLS, as opposed to only those for whom we have genetic data.

Figure 2. MTAG values and cognitive score values by cognitive score decile among non-Jewish respondents, with comparison to Jewish respondents.



In other words, while the large difference in the polygenic score for educational attainment between Jewish and non-Jewish respondents might look like evidence for genetic differences driving the higher test scores and educational attainment of Jews, it is better understood as exemplifying the sort of wonky, implausible result that is why researchers

have been repeatedly warned against interpreting group differences in polygenic scores in the first place.

Dunkel et al. (2019) try to dismiss the plausibility of such differences by citing work indicating that Ashkenazi Jews and non-Jewish persons with European ancestry have "relatively low levels of genetic differentiation." Yet other studies, notably Price et al. (2008), have identified Ashkenazi Jewish ancestry as a main source of systematic genomic variation among those with European ancestry (see also Need et al. 2009).

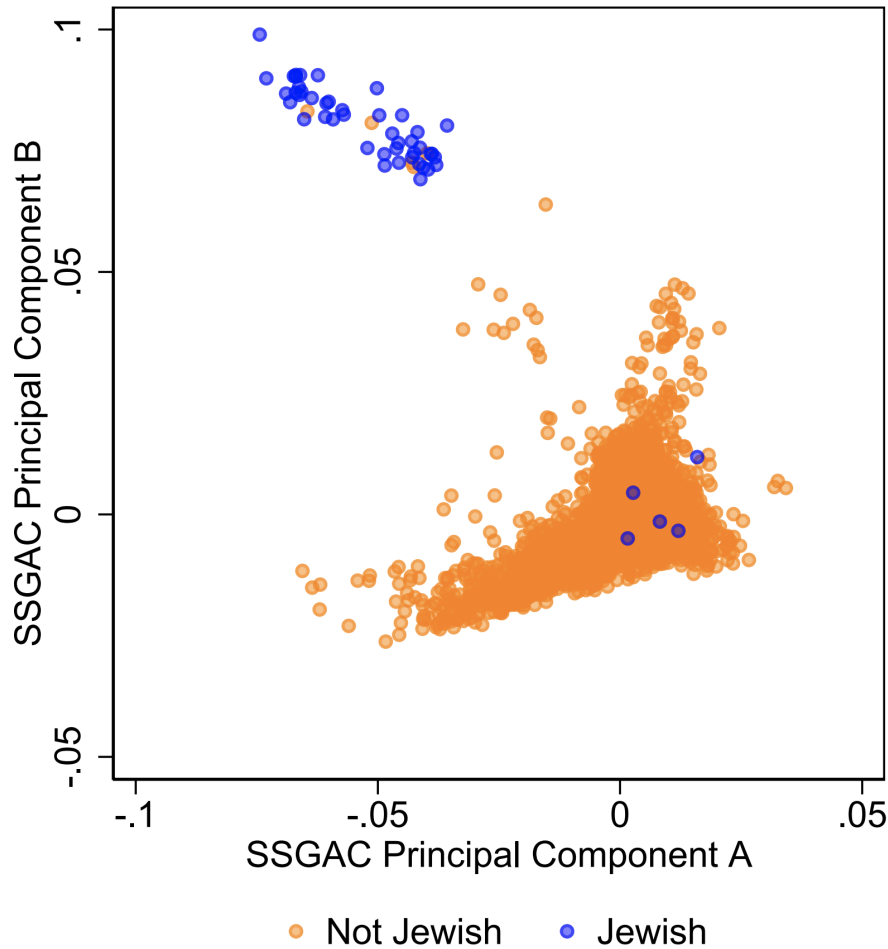
In any event, whatever might be said about subpopulations in general, the few dozen Jewish respondents in WLS can be readily distinguished using genomic information from the rest of the WLS sample. WLS provides information on the first ten principal components from the genome-data, which are intended to reflect the primary ancestry-related ways in which members of the sample differ.² Jewish and non-Jewish respondents differ significantly on nine of the ten principal components ($p < .05$; seven of ten at $p < .001$). Several combinations of just two components are sufficient to produce near-complete separation of these two groups in the data (for one example see Figure 3).³

Separation like this between groups fundamentally contradicts any premise the polygenic scores between groups can be directly compared. The ancestry-related differences in gene frequencies effectively overwhelm the causal differences between individuals that polygenic scores hope to capture. That Jewish respondents have an average polygenic score so different from even those non-Jewish respondents with the highest test scores or most education is evidence of that problem.

² Typically studies of individual variation would statistically adjust for these components, but the combination of the clear separation of groups and the small sample of Jewish respondents involved would make that approach unviable in WLS even apart from other issues with comparing polygenic scores across groups.

³ The numbering of principal components in the released WLS polygenic score data are shuffled to allay identifiability of respondents; here we refer to the two components we use simply as A and B.

Figure 3. Distinctness of Jewish and non-Jewish WLS respondents on two of the first ten principal components of genome-wide data.



In sum, researchers have been warning against using polygenic scores for comparisons across race/ethnic groups for some time now, and a closer look at the data behind Dunkel et al.'s (2019) results provides another illustration of why. The obvious danger of naive efforts to use polygenic scores is that the influence of various sorts of social differences will be misidentified as genetic influences. For example, even though graduates from over 400 high schools are included in WLS, most Jewish respondents in WLS graduated from just two high schools. The non-Jewish respondents from those two schools were also much more likely to attend college than the overall proportion in WLS (89% and 57% vs. 42%). Similarly, even though more than 20% of WLS respondents grew up on farms, virtually no

Jewish respondents did (0%).⁴ Respondents from farms in WLS are less likely to have attended college than other non-Jewish respondents, net of either polygenic score or cognitive test scores in adolescence. In these and other ways, the Jewish respondents in WLS are not just distinct from the bulk of the study's non-Jewish respondents, but these differences are plainly germane to understanding differences in educational attainment.

References

Dunkel, Curtis S., Michael A. Woodley of Menie, Jonatan Pallesen, Emil O. W. Kirkegaard. 2019. Polygenic scores mediate the Jewish phenotypic advantage in educational attainment and cognitive ability compared with Catholics and Lutherans. *Evolutionary Behavioral Sciences*.

Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghziyan, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Linnér, R.K. and Fontana, M.A., 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, 50(8), p.1112.

Martin, Alicia R., Christopher R. Gignoux, Raymond K. Walters, Genevieve L. Wojcik, Benjamin M. Neal, Simon Gravel, Mark J. Daly, Carlos D. Bustamante, Eimear E. Kenny. 2017. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *American Journal of Human Genetics* 100: 635-649.

Need Anna C., Dalia Kasperavičiūtė, Elizabeth T. Cirulli E.T. & David B. Goldstein. 2009. A Genome-Wide Genetic Signature of Jewish Ancestry Perfectly Separates Individuals with and without Full Jewish Ancestry in a Large Random Sample of European Americans. *Genome Biology* 10(1): R7.

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature reviews. Genetics*, 19(3), 148-159.

Price A.L., Butler J., Patterson N., Capelli C., Pascali V.L., Scarnicci F., RuizLinares A., Groop L., Saetta A.A., Korkolopoulou P., Seligsohn U., Waliszewska A., Schirmer C., Ardlie K., Ramos A., Nemesh J., Arbeitman L., Goldstein D.B., Reich D., Hirschhorn J.N. 2009. Discerning the ancestry of European Americans in genetic association studies. *PLoS Genetics* 4:e236.

⁴ Given that the small sample size of Jewish respondents prompts concerns about identifiability, percentages for Jewish respondents are rounded to the nearest multiple of 5.