Too early to declare a general law of social mobility and heritability for education – Supplementary Materials

Damien Morris

In Table S1 below, we provide a full breakdown of revised regression coefficients for standardized and destandandardized variance components after substituting in parent-offspring correlations from Heath et al (1985) (**1**) with respect to the Norwegian samples and birth cohorts used by Engzell and Tropf (2019) (**2**). Table S1 can be compared directly with Table 1 in Engzell and Tropf (**2**). As noted in our letter, the association between heritability and parent-offspring correlation for educational attainment disappears when using this revised data.

Table S1. Correlation of standardized and destandardized variance components with the intergenerational correlation (alternative Norwegian data).

	Standardized			Destandardized		
	h^2	c^2	e^2	$h^{2\prime}$	$c^{2\prime}$	$e^{2\prime}$
Coefficient	-0.208	0.423	-0.433	0.263	0.660	0.151
Robust SE	0.202	0.170	0.153	0.205	0.193	0.235
t statistic	-1.03	2.48	-2.83	1.28	3.41	0.64
P> |*t*|	0.320	0.026	0.013	0.221	0.004	0.530

The table shows correlations (standardized β) between r_{IG} and standardized and destandardized versions of h^2, c^2, and e^2, including a fixed effect for gender. Standard errors are clustered for each birth cohort within a country (15 clusters).

In addition, a significant negative association between parent-offspring correlation and non-shared environmental influence on educational attainment emerges. While we resist drawing any firm conclusions from these alternative analyses for reasons explained below, this association would imply that siblings tend to share fewer genetic or environmental influences in common with each other for educational attainment in samples where they are less like their parents. This is arguably more intuitive than Engzell and Tropf's results (**2**), where siblings appeared no less similar to each other in samples where they were less like their parents.[1]

Fig.1A in our letter depicts how the simple OLS regression slope for heritability (h^2) on parent-offspring correlations is altered when using these revised parent-offspring correlations. Fig.S1A and Fig.S2A show the corresponding change in the slopes for shared environmental influence (c^2) and non-shared environmental influence (e^2) respectively.

For each of these alternative analyses we used the mean of the correlations between twins and their mothers and twins and their fathers for the substituted Norwegian data. For all other samples we used the correlation with the most educated parent in the Global Database on Intergenerational Mobility (GDIM) (**3**) as used by Engzell and Tropf (**2**). Correlations with the most educated parent were not reported for the Norwegian twins in Heath et al (1985) (**1**). The twin-parent correlations obtained for each sex and each parent from Heath et al are provided in Table 1 of our letter.

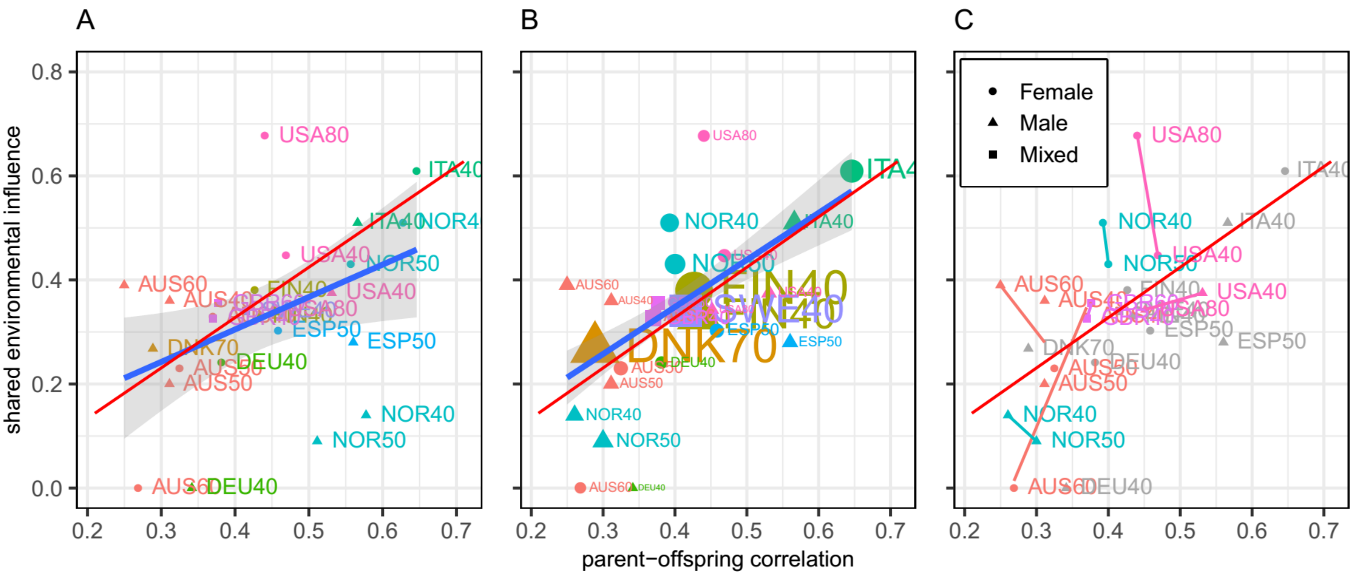

Fig.S1: Three sensitivity tests of the association between shared environmental influence and parent-offspring correlations for educational attainment as reported in Engzell and Tropf (2019): A) substituting Norwegian parent-offspring correlations from Heath et al (1985), B) precision-weighting heritability estimates, C) within-country association testing for each gender. Engzell and Tropf's original OLS regression slope is shown in red.

[1] More specifically, monozygotic twins appeared no less similar in samples where parent-offspring correlations were low (e^2 is calculated as $1\text{-}r_{MZ}$).

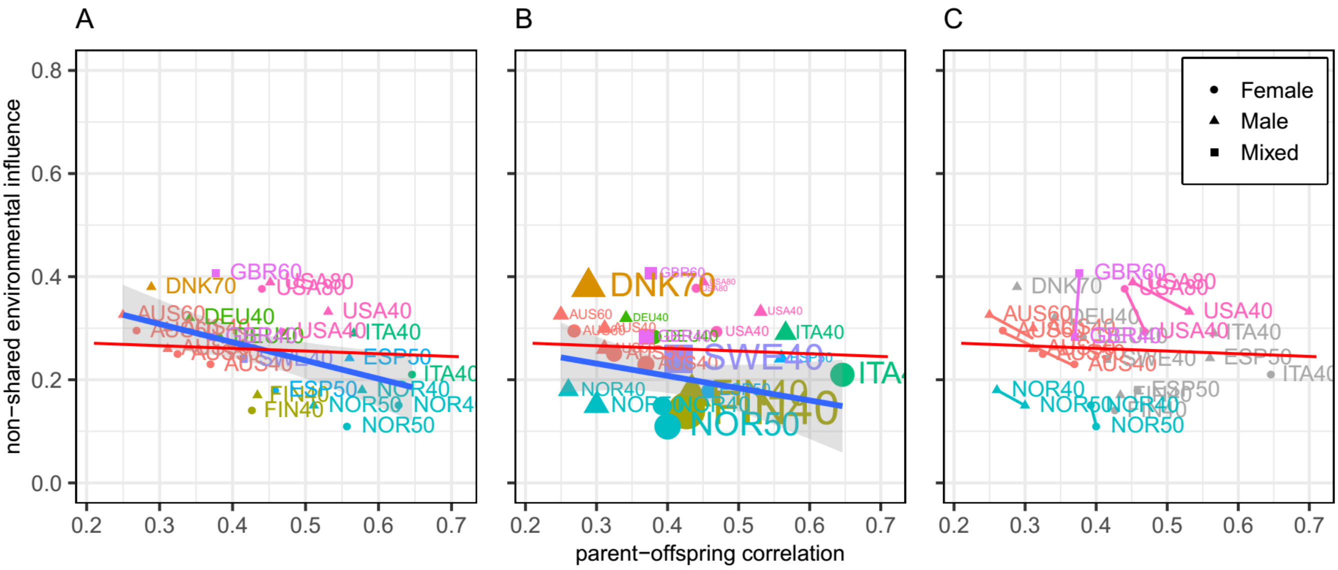

Fig.S2: Three sensitivity tests of the association between non-shared environmental influence and parent-offspring correlations for educational attainment as reported in Engzell and Tropf (2019): A) substituting Norwegian parent-offspring correlations from Heath et al (1985), B) precision-weighting heritability estimates, C) within-country association testing for each gender. Engzell and Tropf's original OLS regression slope is shown in red.

The analysis described above was not pre-registered. However, the recalculation of the heritability slopes originally presented in Fig.1B and Table 1 of Engzell and Tropf (2) was performed at the specific request of a reviewer after our original submission had highlighted the discrepancies between the parent-offspring correlations in the GDIM (3) and those provided for Norwegian twins in Heath et al (1). Moreover, the parent-offspring correlations from Heath et al were not arbitrarily selected. They were obtained following an exhaustive search of the original twin studies comprising Engzell and Tropf's meta-analytic data set. No other twin studies included in the authors' main analyses provided correlations between twins and their parents. Researcher degrees of freedom were therefore extremely limited.[2]

[2] A previous version of this online supplement and a previous version of our online R-script had mistakenly linked the parent-offspring correlations in Lykken et al (1990) to the MIDUS sample and therefore included some exploratory analysis with revised parent-offspring correlations from both Norway and MIDUS. As Minnesota was excluded from Engzell and Tropf's main calculations, this analysis has not been updated.

We stress that the purpose of this analysis was not to establish the true relationship between intergenerational mobility and heritability for educational attainment, but to investigate how confident we can be of results based on matching twin heritability data to parent-offspring correlation data in the GDIM (3). If the results are susceptible to change when intergenerational mobility data from the original twin studies is used, as we have shown, then we cannot be any more confident of these alternative results, which still rely heavily on matched GDIM data in the majority of cases where data on correlations between twins and their parents remains unavailable.[3]

For the above reason, we have not sought to re-run the analysis using parent-offspring correlations from an earlier 1915-1939 Norwegian twin birth-cohort from Heath et al (1985). This birth cohort predated the earliest parent-offspring correlations in the GDIM and was excluded from Engzell and Tropf's analysis. Including it would not provide a sensitivity test of the matched GDIM data.

There are additional tests of the GDIM data that we did not have the opportunity to perform, but which bear on the question of the accuracy of the data matching exercise. For example, several twin studies provide endogenous data on *variances* in educational attainment that we did not test against the GDIM variances used by Engzell and Tropf. In addition, Branigan et al (2013) (**6**) appear to have ascribed inaccurate birth-year ranges to some of their twin birth cohorts when summarising their meta-analytic dataset, leading Engzell and Tropf to inadvertently match some twin samples to GDIM offspring from the wrong decade, e.g. the Danish twins in Bingley et al (2005) (**5**) were born 1925-1977 not 1954-2004.[4] We did not test Engzell and Tropf's results against this corrected twin data.

[3] As performed, the analysis is a lenient test of the GDIM data, insofar as the revised parent-offspring correlations were submitted to the same smoothing process that Engzell and Tropf applied to the GDIM parent-offspring correlations. This means they have been brought more in line with GDIM values for the previous and/or subsequent decade of Norwegian data for each sex.

[4] These issues with Branigan et al (2013) have since been flagged with the communicating author.

As we have concerns as to whether valid inferences can be drawn from the merged twin and GDIM data, we propose that further tests for associations between (standardized or destandardized) variance components and intergenerational mobility should be deferred until more samples with both twin and parent data on educational attainment are identified. By the same logic, we do not incorporate the revised data on parent-offspring correlations for Norwegian twins into our two additional analyses: our precision-weighted regression analysis, or our within-country association tests. These are separate, methodological, robustness checks of Engzell and Tropf's headline finding and, again, are not intended to indicate the true association (or lack thereof) between heritability and social mobility for educational attainment.

Table S2. Correlation of standardized and destandardized variance components with the intergenerational correlation (twin studies precision-weighted).

	Standardized			Destandardized		
	h^2	c^2	e^2	$h^{2\prime}$	$c^{2\prime}$	$e^{2\prime}$
Coefficient	-0.301	0.488	-0.116	NA	NA	NA
Robust SE	0.255	0.093	0.334	NA	NA	NA
t statistic	-1.18	5.27	-0.35	NA	NA	NA
P> |*t*|	0.259	0.000	0.734	NA	NA	NA

The table shows correlations (standardized β) between r_{IG} and standardized and destandardized versions of h^2, c^2, and e^2, including a fixed effect for gender. Standard errors are clustered for each birth cohort within a country (15 clusters).

In Table S2 we provide the full breakdown of revised regression coefficients for the standardized variance components when estimates for those components are precision-weighted. As noted in our letter, this precision-weighting causes the association between parent-offspring correlations and heritability for educational attainment to become non-significant. By contrast, the direction and significance of the associations for c^2 and e^2 are unchanged from Table 1 of Engzell and Tropf (2019) (**2**).[5] Fig.1B in our letter illustrates how the regression slope for heritability on parent-offspring correlations is altered when using precision-weighted h^2 estimates. Fig.S1B and Fig.S2B show the corresponding change in the slopes for shared environmental influence and non-shared environmental influence respectively. This analysis was pre-registered.[6]

We do not report alternative standardised regression coefficients for our within-country analysis as the data is extremely sparse, but Fig.S1C and Fig.S2C show how the within-country regression slopes for each gender compare with the regression slope for the aggregated data for c^2 and e^2 respectively. This complements Fig.1C from our letter which explored within-country regression slopes for h^2. Four of the seven available regression slopes for c^2 fall in the opposite direction to the aggregate association plotted by Engzell and Tropf (2019), again suggesting that Simpson's Paradox may potentially be at work (7). For e^2, six out of the seven within-country regression slopes fall in a negative direction, suggesting the null effect in the aggregate data may potentially be concealing a positive association between intergenerational mobility and non-shared environmental influence within countries.[7] This analysis was also pre-registered. [8]

Supplementary References

1. A. C. Heath, *et al.*, Education policy and the heritability of educational attainment. *Nature* **314**, 734–736 (1985).

2. P. Engzell, F. C. Tropf, Heritability of education rises with intergenerational mobility. *Proc. Natl. Acad. Sci.* **116**, 25386–25388 (2019).

3. What is the Global Database on Intergenerational Mobility (GDIM)? *World Bank* (May 31, 2020).

[5] The inverse variance weights calculated for the standardized ACE estimates could not be assumed to hold for the destandardized estimates (as these multiply the ACE estimates from Heath et al (1985) by the variances in the GDIM). Therefore, we cannot provide revised values for the destandardized ACE estimates.

[6] Timestamped pre-registration document available at https://osf.io/wrjzh/

[7] i.e. a negative association between parent-offspring correlations and e^2.

[8] Timestamped pre-registration document available at https://osf.io/wrjzh/

4. D. T. Lykken, T. J. Bouchard, M. McGue, A. Tellegen, The Minnesota Twin Family Registry: Some Initial Findings. *Acta Genet. Medicae Gemellol. Twin Res.* **39**, 35–70 (1990).

5. P. Bingley, K. Christensen, I. Walker, Twin-based Estimates of the Returns to Education: Evidence from the Population of Danish Twins. 30.

6. A. R. Branigan, K. J. McCallum, J. Freese, Variation in the Heritability of Educational Attainment: An International Meta-Analysis. *Soc. Forces* **92**, 109–140 (2013).

7. R. Kievit, W. E. Frankenhuis, L. Waldorp, D. Borsboom, Simpson's paradox in psychological science: a practical guide. *Front. Psychol.* **4** (2013).