Are Classrooms Equalizers or Amplifiers of Inequality? A Genetically Informative Investigation of Educational Performance

**Supplementary Material**

# Appendix A. Operationalization parental education

**Table A1.** Coding of level of parental education (twin age 10, 7, 3 or 1) into ISLED scores.

|  |  |  |  |
| --- | --- | --- | --- |
| **Twin age 10 and 7** | **Twin age 3 and 1** | **ISLED**  | **Label** |
| (1) Primary school | (1) Lower education | 22.98 | Primary school or first stage of basic education |
| (2) Few years ‘mavo’ or ‘mulo’ (lower general secondary education) |  | 22.98 | Primary school or first stage of basic education |
| (4) Few years ‘LTS’ (lower technical education) |  | 22.98 | Primary school or first stage of basic education |
| (6) Few years ‘havo/VWO’ (higher general secondary education/pre-university education) |  | 22.98 | Primary school or first stage of basic education  |
| (5) Finished ‘LTS’ (lower technical education) |  | 29.34 | Lower secondary school, technical training [lts] |
|  | (2) Lower secondary school, lower general secondary education | 37.31a | Lower secondary school, technical training [lts] and lower secondary school, theoretical training [mulo, mavo] |
| (3) Finished ‘mavo’ or ‘mulo’ (lower general secondary education) |  | 45.27 | Lower secondary school, theoretical training [mulo, mavo] |
| (8) Few years ‘MBO’ (lower vocational education) |  | 45.27 | Lower secondary school, theoretical training [mulo, mavo] |
| (9) Finished ‘MBO’ (lower vocational education) |  | 52.70 | Upper secondary professional education [MBO] |
|  | (3) vocational education ‘MBO’, general secondary education ‘havo, VWO’ | 62.31b | Upper secondary professional education [MBO], higher secondary school [havo], and pre-scientific secondary school [VWO] |
| (7) Finished ‘havo/VWO’ (higher general secondary education/pre-university education) |  | 67.11c | Higher secondary school [havo] and pre-scientific secondary school [VWO] |
| (10) Few years ‘HBO’ (university of applied sciences) or university |  | 67.11c | Higher secondary school [havo] and pre-scientific secondary school [VWO] |
| (11) Finished ‘HBO’ (university of applied sciences) | (4) higher vocational education ‘HBO’ | 77.93 | Tertiary professional education [HBO] |
| (12) Finished university | (5) scientific education | 87.13 | Tertiary scientific education [WO] |
| (13) Post-doctoral education |  | 92.63d | Post-doctoral education and second stage of tertiary education, PHD |

a Average of 29.34 (lower secondary school) and 45.27 (lower general secondary education)

b Average of upper secondary professional education (52.70), higher secondary school (62.30), and pre-scientific secondary school (71.92).

c Average of 62.30 (higher secondary school) and 71.92 (pre-scientific secondary school)

d Average of 90.63 (post-doctoral education) and 94.62 (second stage of tertiary education, PhD).

# Appendix B. Robustness checks different operationalization classroom allocation

As a robustness check, we used a stricter definition of classroom allocation where we also excluded twins who went to different schools and those who were in the same school but in different grades. This leads to a sample of 3,443 twin pairs (880 MZ twin pairs in the same classroom, 596 MZ twin pairs in different classrooms, 1,444 DZ twin pairs in the same classroom, and 693 DZ twin pairs in different classrooms). Similar to the main analyses, we first checked the assumptions of equal means and variances using likelihood-ratio tests. Distinguishing a classroom component requires the variance in educational performance to be equal between the different groups (i.e., MZ and DZ twins in the same class, and MZ and DZ twins in different classes). However, in contrast to the main analyses, the assumption of equal variances between twins in the same and different classrooms cannot be met when we use the stricter operationalization of classroom allocation. Since there are different ways to deal with this, we provide three robustness analyses. First, we fitted an *ACE*-model for MZ and DZ twins in the same and different classrooms where we allow the variances to be different (see Analysis 1). This is probably the most appropriate test, but less informative as it does not allow for formal estimation of a classroom effect. Therefore, as alternative tests, we ignored that the variances are unequal and nevertheless performed analyses including a classroom component. In Analysis 2, we ignored the violation of equal variances by fitting the model with a classroom component while we constrained the variances across groups to be equal (although the LR-test showed that they cannot be equalized). Lastly, in Analysis 3, we estimated the model with a classroom component while we equalized the variances by standardizing the *Cito*-scores within each group. This standardization artificially creates equal variances between the groups.

All three tests lead to the same conclusions as those reported in the main text: there are classroom effects that significantly decrease with increasing parental education. The evidence for a compensation effect is thus robust against a different operationalization of classroom allocation and different ways of modeling.

# *1. Results based on ACE-model*

We decomposed the variance in *Cito*-scores (unstandardized) for MZ and DZ twins in the same and different classes into *ACE*-components, where we allowed the variances to differ between twins in the same and different classrooms. Although we do not directly estimate a classroom component, the *C*- and *E*-components for twins in the same and different classes allow us to infer whether there is a classroom effect. If twins are in the same class, the classroom is part of their shared environment. If there are classroom effects, one would thus expect twins in the same class to be more alike (larger *C*) than twins in different classes. Similarly, for twins in different classes, the classroom is part of their unique, non-shared environment. Hence, these twins can be expected to have a larger *E* than twins in the same class if classroom effects are present.

Table B1 shows larger shared environmental variance for twins in the same classroom ($V\_{C}$ = 9.170) than for twins in different classrooms ($V\_{C}$ = 6.372), indicating a classroom effect. Standardizing the variance components also shows that the shared environment explains a larger proportion of the variance in educational performance if twins are in the same classroom (12.4%) than if twins are in different classrooms (10.4%). While the absolute non-shared environmental variance is similar in both groups, the standardized results show that the non-shared environment explains more variance for twins in different classrooms (24.0%) than for twins in the same classroom (19.2%). This is also consistent with classroom influences being present.

 Table B2 shows the sources of variance for twins in the same and different classrooms moderated by parental education. The difference in shared environmental variance between twins in the same and different classrooms becomes smaller with increasing parental education (see Figure B1). The same applies to non-shared environmental variance. With lower levels of parental education, there is a larger difference in these sources of environmental variance, indicating classroom influences. These become smaller and are eventually absent with increasing parental education.

# *2. Results based on ACE-model including the classroom component with unstandardized Cito-scores*

Table B3 shows the results of the model including a classroom component where we equalized the variance in educational performance for twins in the same and different classes, even though the LR-test showed that the variances differ. Model 1 shows that there is very small and non-significant classroom variance ($V\_{CL}$ = 0.245). Only 0.3% of the variance in educational performance can be explained by classroom effects.

Classroom variance decreases significantly with increasing parental education ($b\_{cl} $= -0.024, *p* = .042), as shown in Model 2. Figure B2 plots the moderation of the unstandardized and standardized variance components. In both cases, the results show a pattern of compensation. For children with the lowest educated parents (primary education - ISLED 22.98), 5.2% of the variance in educational performance can be attributed to the classroom context. For the average parental education (higher and pre-scientific secondary school – ISLED 67.11), 2.4% can be attributed to classroom effects, and for children with the highest educated parents (postdoctoral education – ISLED 92.63) this is 0.9%.

# *3. Based on ACE-model including the classroom component with standardized Cito-scores*

Another way to deal with unequal variances is by standardizing the Cito-scores within each group. The results are presented in Table B4. There is significant classroom variance ($V\_{CL}$ = 0.044, *p* = .037), making up 4.5% of the total variance in educational performance. Classroom variance decreases significantly with increasing parental education ($b\_{cl} $= -0.003, *p* = .020), as shown in Model 2. Figure B3 plots the moderation of the unstandardized and standardized variance components. In both cases, the results show a pattern of compensation. For children with the lowest educated parents (primary education - ISLED 22.98), 10.6% of the variance in educational performance can be attributed to the classroom context. For the average parental education (higher and pre-scientific secondary school – ISLED 67.11), 6.3% can be attributed to classroom effects. For children with the highest educated parents (postdoctoral education – ISLED 92.63), this is 3.5%.

#### **Table B1**. Results of twin model for educational performance (unstandardized) for MZ twins in the same classroom (Npairs = 880), DZ twins in the same classroom (Npairs = 1,444), MZ twins in different classrooms (Npairs = 596), and DZ twins in different classrooms (Npairs = 693).

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Same classroom** |  | **Different classrooms** |
|  | **Estimate** | **se** |  | **Estimate** | **se** |
| Intercept | 551.055 | \*\*\* | (3.826) |  | 551.055 | \*\*\* | (3.826) |
| **Paths** |  |  |  |  |  |  |  |
| *a* | 7.125 | \*\*\* | (0.265) |  | 6.349 | \*\*\* | (0.379) |
| *c* | 3.028 | \*\*\* | (0.554) |  | 2.524 | \*\* | (0.831) |
| *e* | 3.777 | \*\*\* | (0.126) |  | 3.834 | \*\*\* | (0.170) |
| **Variances**  |  |  |  |  |  |  |  |
| *VA*  | 50.762 | \*\*\* | (3.773) |  | 40.306 | \*\*\* | (4.807) |
| *VC*  | 9.170 | \*\* | (3.358) |  | 6.372 |  | (4.198) |
| *VE*  | 14.269 | \*\*\* | (0.955) |  | 14.702 | \*\*\* | (1.301) |
| *Vtotal* | 74.202 | \*\*\* | (1.944) |  | 61.38 | \*\*\* | (2.145) |
| **Model fit** |  |  |  |  |  |  |  |
| # free parameters | 33 |
| *LL* | -29756.052 |
| *AIC* | 59578.105 |

*Note*: Estimates are unstandardized, controlled for year of birth and sex in all models. *Se* = standard error, *VA* = genetic variance, *VC* = shared environmental variance, *VE* = non-shared environmental variance, *Vtotal* = total variance, *LL* = loglikelihood, *AIC* = Akaike Information Criterion.

\**p* < .05, \*\**p* < .01, \*\*\**p* < .001 (two-tailed tests).

#### **Table B2**. Results of twin moderation model for educational performance (unstandardized) for MZ twins in the same classroom (Npairs = 880), DZ twins in the same classroom (Npairs = 1,444), MZ twins in different classrooms (Npairs = 596), and DZ twins in different classrooms (Npairs = 693).

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Same classroom** |  | **Different classrooms** |
|  | **Estimate** | **se** |  | **Estimate** | **se** |
| Intercept | 535.035 | \*\*\* | (3.636) |  | 535.035 | \*\*\* | (3.636) |
| Parental education | 0.155 | \*\*\* | (0.007) |  | 0.155 | \*\*\* | (0.007) |
| **Paths** |  |  |  |  |  |  |  |
| *a* | 9.674 | \*\*\* | (0.577) |  | 8.305 | \*\*\* | (1.577) |
| *c* | 1.346 |  | (1.898) |  | 5.282 |  | (3.741) |
| *e* | 4.556 | \*\*\* | (0.447) |  | 6.364 | \*\*\* | (0.733) |
| $$b\_{a}$$ | -0.040 | \*\*\* | (0.009) |  | -0.028 |  | (0.019) |
| $$b\_{c}$$ | -0.004 |  | (0.012) |  | -0.062 |  | (0.050) |
| $$b\_{e}$$ | -0.012 |  | (0.007) |  | -0.035 | \*\*\* | (0.010) |
| **Model fit** |  |  |  |  |  |  |  |
| # free parameters | 48 |
| *LL* | -44098.159 |
| *AIC* | 88292.318 |

*Note*: Estimates are unstandardized, controlled for year of birth and sex in all models. *Se* = standard error, *a* = genetic path, *c* = shared environmental path, *e* = non-shared environmental path, *b* = moderation by parental education, LL= loglikelihood, *AIC* = Akaike Information Criterion.

\*\*\**p* < .001 (two-tailed tests).



#### **Figure B1.** Difference in shared environmental (C) and non-shared environmental (E) variance in educational performance between twins in the same and different classes, moderated by parental education.

*Note*:Results based on Table B2.

|  |  |
| --- | --- |
|  |  |
| **a) unstandardized components** | **b) standardized components** |

#### **Figure B2.** Decomposition of the unstandardized (a) and standardized (b) variance in educational performance moderated by parental education.

*Note*:Sources of variance include genetic (*A*), shared environmental (*C*), classroom (*CL*), and non-shared environmental (*E*) variance. Results based on Table B3 Model 2.

#### **Table B3**. Results of twin models for educational performance for MZ twins in the same classroom (Npairs = 880), DZ twins in the same classroom (Npairs = 1,444), MZ twins in different classrooms (Npairs = 596), and DZ twins in different classrooms (Npairs = 693).

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Model 1** |  | **Model 2** |
|  | **Estimate** | **se** |  | **Estimate** | **se** |
| Intercept | 550.311 | \*\*\* | (3.822) |  | 534.852 | \*\*\* | (3.630) |
| Parental education |  |  |  |  | 0.155 | \*\*\* | (0.007) |
| **Paths** |  |  |  |  |  |  |  |
| *a* | 6.888 | \*\*\* | (0.219) |  | 9.533 | \*\*\* | (0.478) |
| *c* | 2.840 | \*\*\* | (0.219) |  | 0.000 |  | (0.000) |
| *cl* | 0.495 |  | (1.425) |  | 2.809 | \* | (1.100) |
| *e* | 3.786 | \*\*\* | (0.122) |  | 4.795 | \*\*\* | (0.400) |
| $$b\_{a}$$ |  |  |  |  | -0.040 | \*\*\* | (0.007) |
| $$b\_{c}$$ |  |  |  |  | 0.000 |  | (0.000) |
| $$b\_{cl}$$ |  |  |  |  | -0.024 | \* | (0.012) |
| $$b\_{e}$$ |  |  |  |  | -0.016 | \*\* | (0.006) |
| **Variances**  |  |  |  |  |  |  |  |
| *VA*  | 47.449 | \*\*\* | (3.016) |  | a |  |  |
| *VC*  | 8.066 | \*\* | (2.678) |  | a |  |  |
| *VCL* | 0.245 |  | (1.411) |  | a |  |  |
| *VE*  | 14.336 | \*\*\* | (0.926) |  | a |  |  |
| *Vtotal* | 70.096 | \*\*\* | (1.500) |  | a |  |  |
| **Model fit** |  |  |  |  |  |  |  |
| # free parameters | 31 |  | 44 |
| *LL* | -29765.895 |  | -44102.470 |
| *AIC* | 59593.790 |  | 88292.939 |

*Note*: Estimates are unstandardized, controlled for age and sex in all models. *Se* = standard error, *VA* = genetic variance, *VC* = shared environmental variance, *VCL* = classroom variance*, VE* = non-shared environmental variance, *Vtotal* = total variance, *LL* = loglikelihood, *AIC* = Akaike Information Criterion.

a Not applicable, because in the moderation model the size of a variance component depends on the level of parental education.

\**p* < .05, \*\**p* < .01, \*\*\**p* < .001 (two-tailed tests)

|  |  |
| --- | --- |
|  |  |
| **a) unstandardized components** | **b) standardized components** |

#### **Figure B3.** Decomposition of the unstandardized (a) and standardized (b) variance in educational performance moderated by parental education.

*Note*:Sources of variance include genetic (*A*), shared environmental (*C*), classroom (*CL*), and non-shared environmental (*E*) variance. Based on Table B4 Model 2.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Model 1** |  | **Model 2** |
|  | **Estimate** | **se** |  | **Estimate** | **se** |
| Intercept | 1.505 | \*\*\* | (0.459) |  | -0.237 |  | (0.444) |
| Parental education |  |  |  |  | 0.017 | \*\*\* | (0.001) |
| **Paths** |  |  |  |  |  |  |  |
| *a* | 0.808 | \*\*\* | (0.027) |  | 1.063 | \*\*\* | (0.058) |
| *c* | 0.312 | \*\*\* | (0.065) |  | 0.000 |  | (0.005) |
| *cl* | 0.211 | \*\*\* | (0.050) |  | 0.455 | \*\*\* | (0.105) |
| *e* | 0.211 | \*\*\* | (0.050) |  | 0.563 | \*\*\* | (0.049) |
| $$b\_{a}$$ |  |  |  |  | -0.004 | \*\*\* | (0.001) |
| $$b\_{c}$$ |  |  |  |  | 0.000 |  | (0.000) |
| $$b\_{cl}$$ |  |  |  |  | -0.003 | \* | (0.001) |
| $$b\_{e}$$ |  |  |  |  | 0.563 | \*\*\* | (0.049) |
| **Variances**  |  |  |  |  |  |  |  |
| *VA*  | 0.652 | \*\*\* | (0.043) |  | a |  |  |
| *VC*  | 0.097 | \* | (0.041) |  | a |  |  |
| *VCL* | 0.044 | \* | (0.021) |  | a |  |  |
| *VE*  | 0.199 | \*\*\* | (0.013) |  | a |  |  |
| *Vtotal* | 0.994 | \*\*\* | (0.021) |  | a |  |  |
| **Model fit** |  |  |  |  |  |  |  |
| # free parameters | 31 |  | 44 |
| *LL* | -15532.938 |  | -29905.319 |
| *AIC* | 31127.876 |  | 59898.637 |

#### **Table B4.** Results of twin models for educational performance for MZ twins in the same classroom (Npairs = 880), DZ twins in the same classroom (Npairs = 1,444), MZ twins in different classrooms (Npairs = 596), and DZ twins in different classrooms (Npairs = 693).

*Note*: Estimates are unstandardized, controlled for age and sex in all models. *Se* = standard error, *VA* = genetic variance, *VC* = shared environmental variance, *VCL* = classroom variance*, VE* = non-shared environmental variance, *Vtotal* = total variance, *LL* = loglikelihood, *AIC* = Akaike Information Criterion.

a Not applicable, because in the moderation model the size of a variance component depends on the level of parental education.

\**p* < .05, \*\**p* < .01, \*\*\**p* < .001 (two-tailed tests).