Online Supplement

S1. Missing data

In PISA 2018, non-core constructs such as perceived competitiveness (~19% missing), self-efficacy (~11%), peer competitiveness (~8%), and self-perceived ability (~8%) were administered to randomly selected subsamples of students to reduce response burden (OECD, 2019). Consequently, missingness in these variables reflects planned questionnaire design rather than item non-response. Other variables had relatively low levels of missingness: immigration status (~4%), ESCS index (~2%), and grade (~0.5%). Performance measures and gender contained no missing data.

For the main analyses, we applied listwise deletion, whereby students with a missing value on at least one of the variables included in the model were excluded from the analysis. Since using listwise deletion assumes that data are missing completely at random (MCAR) we assessed the sensitivity of our results by conducting a pre-registered robustness checks by using multiple imputation by chained equations (MICE; van Buuren & Groothuis-Oudshoorn, 2011). Multiple imputation was applied to all variables, including dependent variables, with more than 5 percent missing values, which is a commonly used threshold (Schafer, 1999).The number of imputed datasets (m = 20) was determined based on the highest proportion of missing data across variables, following the recommendation that the number of imputed datasets should be at least equal to the percentage of missing cases (White et al., 2011). MICE is a well-suited method for handling missing data in large and complex datasets as it generates multiple imputed datasets, runs the analysis on each, and combines results using Rubin's (2018) combining rules, thereby properly accounting for uncertainty.

The imputed results did not differ substantially from the main analyses using listwise deletion suggesting that missing data did not meaningfully bias our main findings.

S2. Mathematics and Reading Literacy

In PISA 2018, reading literacy is defined as "understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society" (OECD, 2019, p. 28). Reading literacy focused on students' ability to locate, understand, and evaluation information from various digital sources. In these "multi-text" scenarios, students had to bring together potentially conflicting information from simulated websites, emails and forum posts. Cognitive processes were measured, such as scanning for relevant cues, integrating information from different formats, and reflecting on the reliability and quality of the text. The reading achievement test aimed to capture how students interact and use information in a digital and real-world environment.

The definition of mathematics literacy in the PISA 2018 assessment is "an individual's capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognize the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens" (OECD, 2019, p. 75). Mathematics literacy focused on students' capacity to formulate, apply, and interpret mathematical reasoning to solve problems in real-world contexts. The test consisted of four categories: algebra, geometry, arithmancy, and statistics. Students had to translate a problem into a mathematical structure, perform the necessary calculation and interpret the results back into a practical conclusion.

See https://www.oecd.org/en/about/programmes/pisa/pisa-test.html for interactive test questions used in PISA assessments (OECD, 2019).

S3. Construction of peer competitiveness

We used students' individual dispositions towards competitiveness, measured using the PISA 2018 index of attitudes towards competition (COMPETE) to create an aggregated school-level mean of peer competitiveness. “The index of attitudes towards competition (COMPETE) was constructed using students’ responses to a new question (ST181) over the extent they “strongly disagreed”, “disagreed”, “agreed” or “strongly agreed” with the following statements: “I enjoy working in situations involving competition with others”; “It is important for me to perform better than other people on a task”; and “I try harder when I’m in competition with other people”. Positive values on this scale mean that students expressed more favorable attitudes towards competition than did the average student across OECD countries” (OECD, 2019. p. 213).

To create a measure of peer competitiveness at the school level, we computed school-level averages of COMPETE using a leave-one-out procedure, whereby the score of the focal student was excluded when calculating the average for that student. This approach ensures that each student receives a peer competitiveness score that reflects the competitive dispositions of their schoolmates rather than their own disposition, thereby avoiding the problem of students contributing to their own peer environment measure. This results in a unique peer competitiveness score for each student.

The peer competitiveness measure was computed before we adjust for th5 student weight to ensure that peer competitiveness reflects as it exists in the population. When computing the peer competitiveness measure, we used the original student weight and not the adjusted senate weights. We did this to ensure that peer competitiveness reflects the actual school climate.

To ensure stable and reliable school-level aggregates, we restricted our sample to schools with at least 20 students. We conducted two robustness checks to verify that our

results were not sensitive to these analytical choices: first, restricting the sample to schools with at least 30 students, and second, computing school-level averages without the leave-one-out correction. Both robustness checks yielded substantively identical results to our main analyses, supporting the robustness of our peer competitiveness measure.

S4. Per-country internal consistency scores (Cronbach's alpha)				
Country	Perceived competitiveness	Trait competitiveness	Self-Efficacy	Self-perceived ability
Albania	0.84	0.80	0.81	0.66
Baku (Azerbaijan)	0.90	0.86	0.89	0.71
Argentina	0.75	0.73	0.75	0.65
Australia	0.85	0.78	0.77	0.87
Belgium	0.86	0.73	0.67	0.65
Bosnia and Herzegovina	0.84	0.80	0.82	0.75
Brazil	0.82	0.70	0.74	0.73
Brunei Darussalam	0.77	0.69	0.65	0.69
Bulgaria	0.85	0.79	0.85	0.80
Chile	0.84	0.79	0.82	0.73
Chinese Taipei	0.92	0.75	0.79	0.84
Colombia	0.74	0.77	0.80	0.72
Costa Rica	0.79	0.81	0.79	0.69
Croatia	0.90	0.78	0.77	0.80
Czech Republic	0.84	0.76	0.73	0.78
Denmark	0.82	0.77	0.76	0.85
Dominican Republic	0.82	0.78	0.85	0.77
Estonia	0.79	0.74	0.79	0.74
Finland	0.85	0.80	0.80	0.86
France	0.88	0.78	0.73	0.76
Georgia	0.81	0.77	0.82	0.75
Germany	0.81	0.75	0.73	0.78
Greece	0.81	0.72	0.75	0.71
Hong Kong	0.90	0.74	0.81	0.83
Hungary	0.88	0.75	0.74	0.72
Iceland	0.89	0.83	0.87	0.88
Indonesia	0.83	0.65	0.77	0.67
Ireland	0.85	0.78	0.73	0.85
Israel	0.85	0.80	0.79	0.82
Italy	0.80	0.77	0.72	0.73
Kosovo	0.82	0.69	0.81	0.75
Kazakhstan	0.90	0.84	0.78	0.65
Jordan	0.84	0.78	0.81	0.79
Korea	0.91	0.74	0.82	0.82
Latvia	0.81	0.76	0.74	0.70
Lithuania	0.87	0.76	0.78	0.79
Luxembourg	0.82	0.78	0.79	0.81
Macao	0.85	0.70	0.72	0.76
Malaysia	0.77	0.67	0.70	0.67
Malta	0.83	0.77	0.77	0.82
Mexico	0.79	0.80	0.80	0.73
Moldova	0.73	0.68	0.73	0.69
Montenegro	0.86	0.77	0.82	0.79
Morocco	0.80	0.73	0.77	0.63

S4. Per-country internal consistency scores (Cronbach's alpha)				
Country	Perceived competitiveness	Trait competitiveness	Self-Efficacy	Self-perceived ability
Netherlands	0.77	0.74	0.73	0.80
New Zealand	0.85	0.76	0.73	0.87
Panama	0.83	0.77	0.78	0.71
Peru	0.82	0.78	0.77	0.72
Philippines	0.81	0.76	0.78	0.71
Poland	0.81	0.68	0.78	0.68
Portugal	0.86	0.75	0.75	0.74
Qatar	0.86	0.81	0.76	0.80
Romania	0.78	0.74	0.74	0.67
Saudi Arabia	0.84	0.75	0.81	0.75
Serbia	0.85	0.74	0.82	0.76
Singapore	0.83	0.72	0.73	0.82
Slovak Republic	0.82	0.76	0.78	0.67
Slovenia	0.83	0.76	0.81	0.79
Spain	0.82	0.77	0.75	0.72
Sweden	0.89	0.78	0.82	0.87
Switzerland	0.83	0.75	0.73	0.79
Thailand	0.82	0.70	0.76	0.74
United Arab Emirates	0.85	0.83	0.84	0.79
Turkey	0.89	0.88	0.85	0.80
Ukraine	0.83	0.67	0.72	0.71
United Kingdom	0.86	0.80	0.77	0.86
United States	0.85	0.80	0.78	0.85
Uruguay	0.82	0.78	0.80	0.75
B-S-J-Z (China)	0.84	0.72	0.80	0.79

S5. Regressions equations

The following equations formally represent the regression analyses:

For Model A, which only includes the direct associations of SES and competitiveness:

$$Y_{ijk} = \beta_0 + \beta_1 SES_{ijk} + \beta_2 Competitiveness_{ijk} + \upsilon_{0jk} + \upsilon_{0k} + \varepsilon_{ijk}$$

For Model B1, which includes the interaction between SES and perceived competitiveness, the equation is:

$$Y_{ijk} = \beta_0 + \beta_1 SES_{ijk} + \beta_2 Perceived\ Competitiveness_{ijk} + \beta_3(SES_{ijk} * Perceived\ Competitiveness_{ijk}) + \upsilon_{0jk} + \nu_{0k} + \varepsilon_{ijk}$$

For Model B2, which includes the interaction between SES and peer competitiveness, the equation is:

$$Y_{ijk} = \beta_0 + \beta_1 SES_{ijk} + \beta_2 Peer\ Competitiveness_{ijk} + \beta_3(SES_{ijk} * Peer\ Competitiveness_{ijk}) + \upsilon_{0jk} + \upsilon_{1jk} SES_{ijk} + \nu_{0k} + \varepsilon_{ijk}$$

In Model A, we included either perceived competitiveness or peer competitiveness, together with SES, as an independent predictor to assess its direct association with the outcome variable. Model B is an extension of Model A and includes an interaction term between the competitiveness dimension and SES to examine whether the relationship between competitiveness and the outcome variable varied according to students' SES.

In the multilevel model, Y_{ijk} represents the outcome variable for student *i* in school *j* within country *k*. The intercept is denoted by β_0. The coefficient $\beta_1 SES_{ijk}$ represents the estimate of socioeconomic status (SES), while $\beta_2 Competitiveness_{ijk}$ captures the estimate of either perceived or peer competitiveness, modeled separately. The interaction term $\beta_3(SES_{ijk} * Competitiveness_{ijk})$ captures the moderation of SES on the relationship between competitiveness and the outcome variable, and is included only in Model B. The model also includes random intercepts at both the school and country level, denoted by υ_{0jk} and ν_{0k}, respectively. For the interaction models including peer competitiveness, $\upsilon_{1jk} SES_{ijk}$ denotes the random slope for SES on the school level, allowing the effect of SES

on the outcome to vary across schools. Finally, ε_{ijk} represents the residual error at the student level.

S6. Descriptives analyses

Table 1 shows the full correlation matrix of all the variables included in this study. Figure 1 shows a clear SES gradient for all outcome variables and competitiveness measures: on average, students with higher SES have higher standardized test scores. The SES gradient is strongest for mathematics and reading scores, where students in the highest SES decile score more than 1 standard deviation higher than students in the lowest SES decile. For self-efficacy and self-perceived ability, the SES gradient is less steep, but the differences are still considered substantial with gaps of about 0.4 SD and 0.6 SD, respectively. For the competitiveness measures, the SES gaps are more modest. The increase across SES deciles is small but consistent for perceived and peer competitiveness. The figure also shows that SES and individual competitive dispositions (labelled as trait competitiveness in the figure) were positively related.

To better understand the SES gradient across countries, we plotted the standardized SES slopes on the four outcomes variables for each country. Figures 2 to 5 show that the SES slopes differ substantially, indicating that the strength of the SES gradient is not uniform across countries. The SES slope for reading and mathematics is the biggest in magnitude, suggesting strong social stratification in academic performance. Self-views are also stratified by SES but to a lesser extent than performance.

In figure 6, we depicted boxplots of within-school SES slopes on the outcomes and competitiveness measures. Again, mathematics and reading show the strongest and most consistent SES gradient, meaning that students with higher SES have better performance in nearly all schools. For self-views, we also observe a positive SES gradient, but with greater

dispersion. This suggests that the link between SES and self-views varies more across the sampled schools than for academic performance.

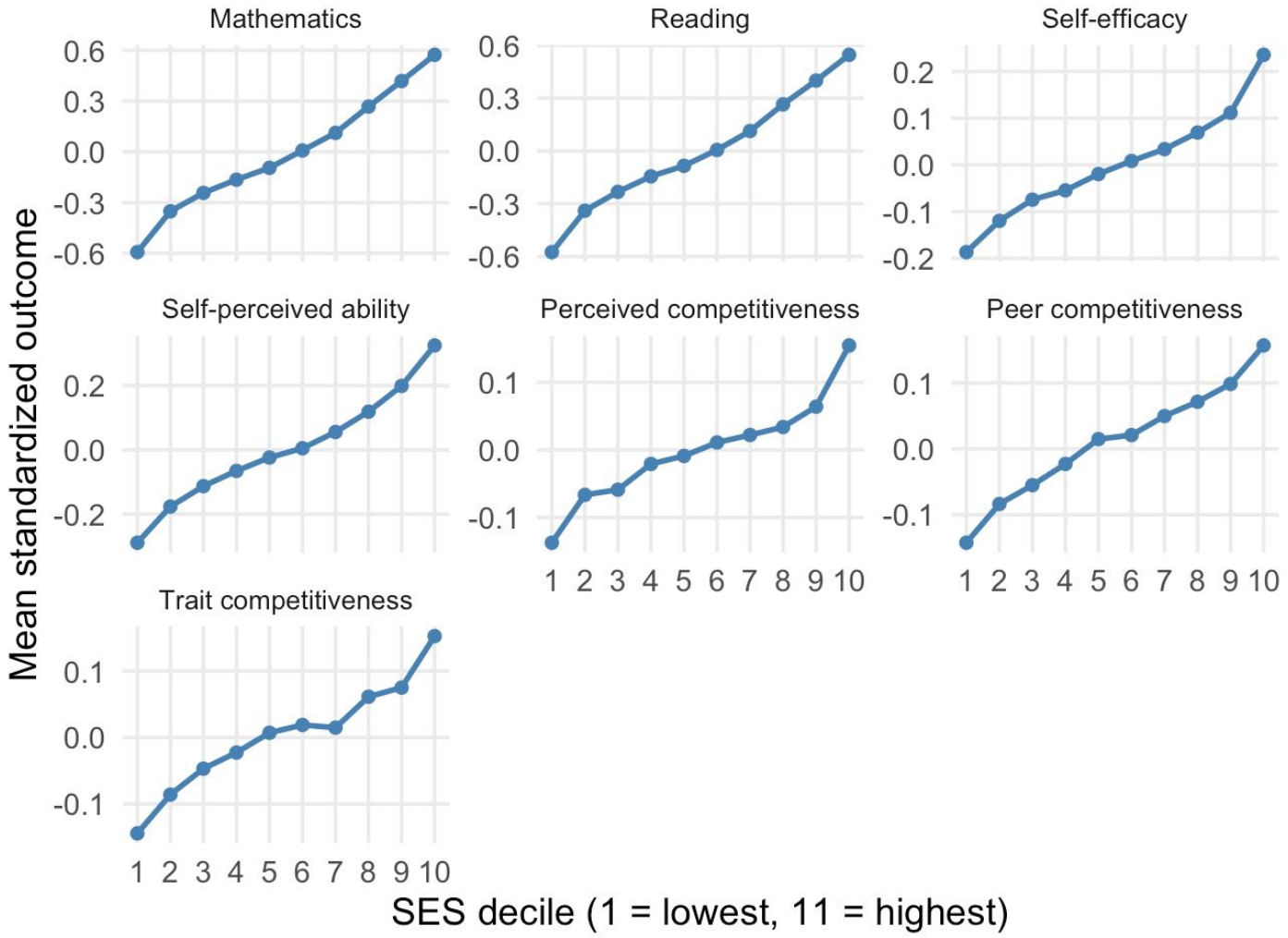

Figure 1. Standardized scores on performance, self-views and competitiveness for SES deciles.

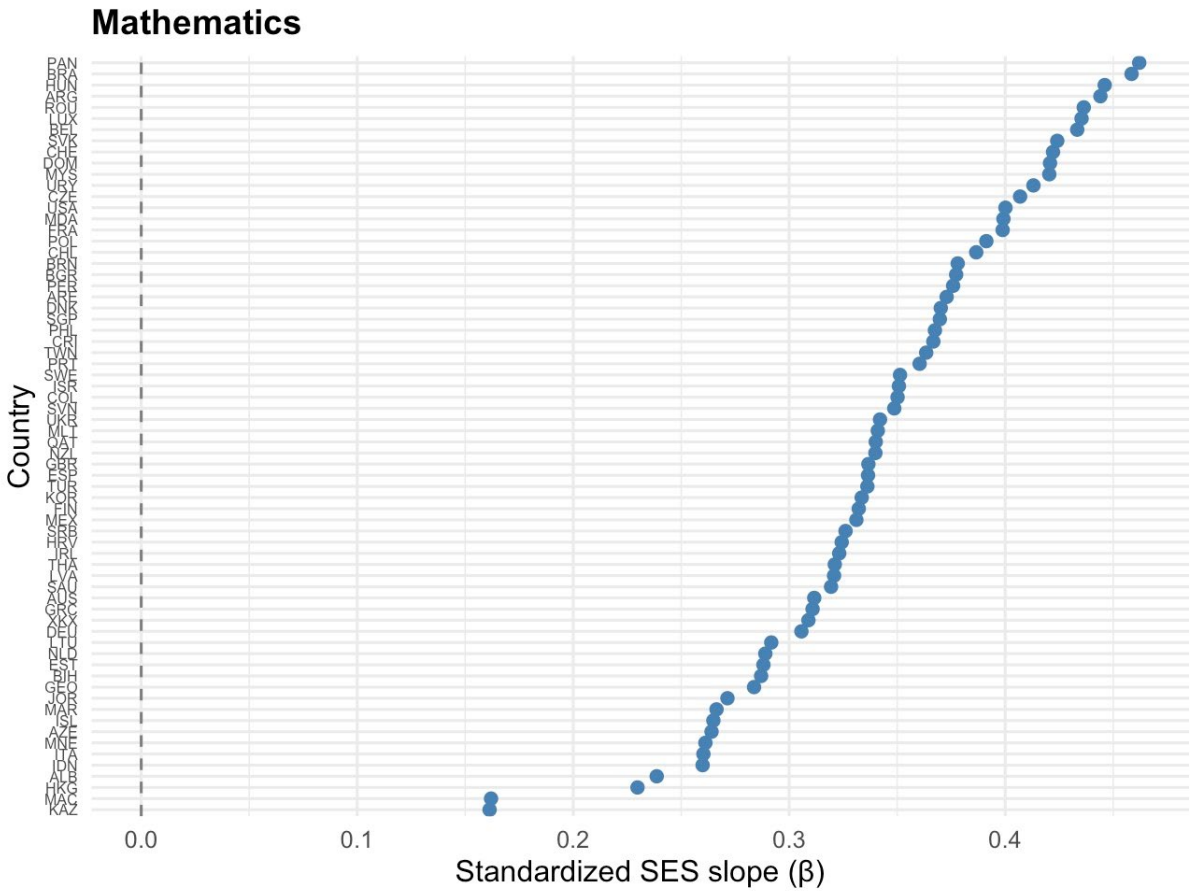

Figure 2. Standardized SES slope for mathematics for each country.

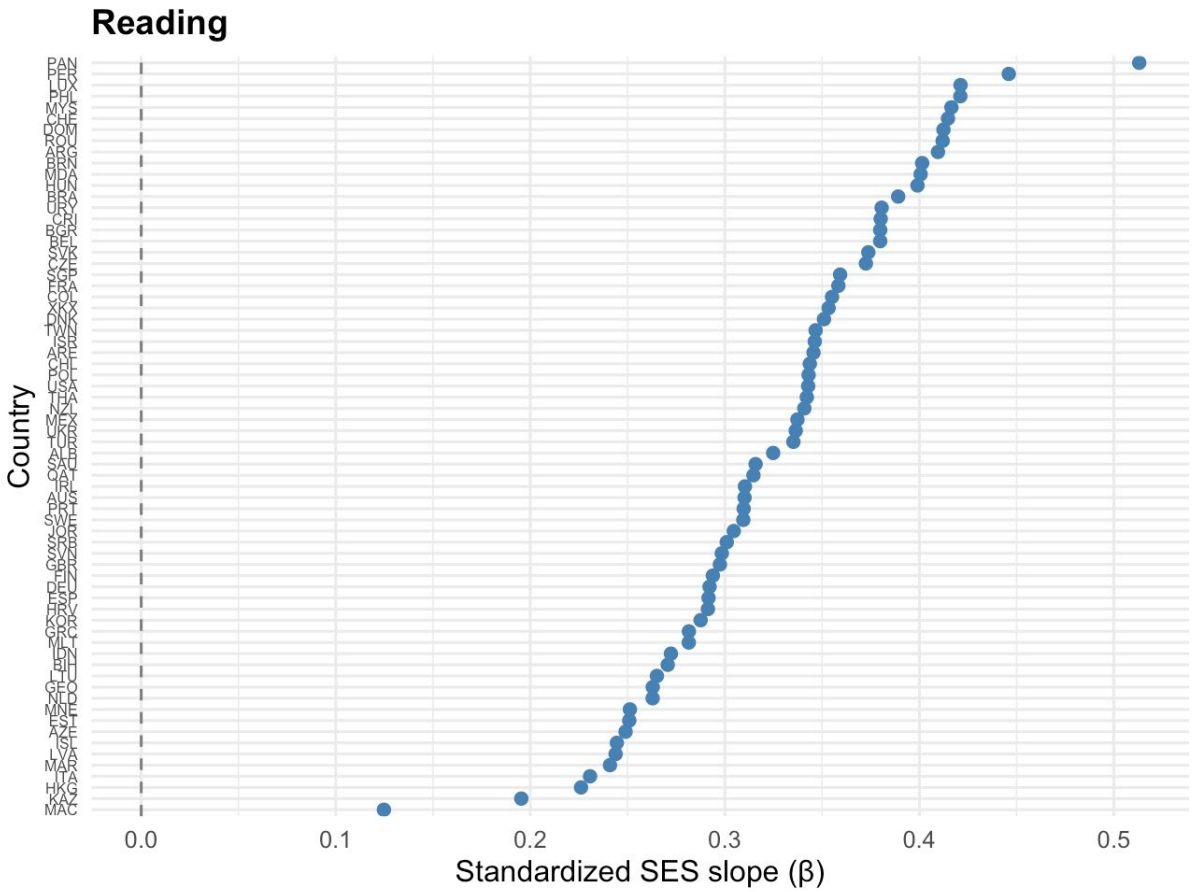

Figure 3. Standardized SES slope for reading for each country.

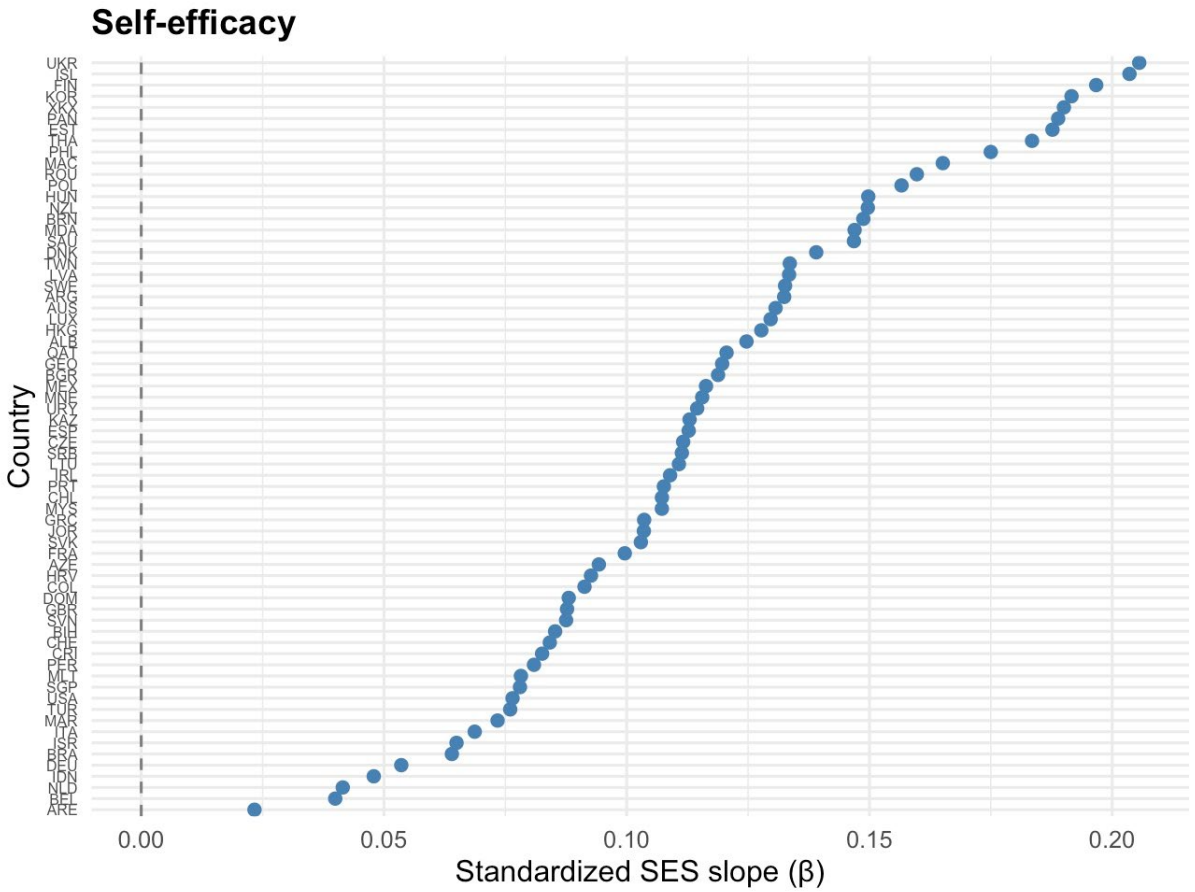

Figure 4. Standardized SES slope for self-efficacy for each country.

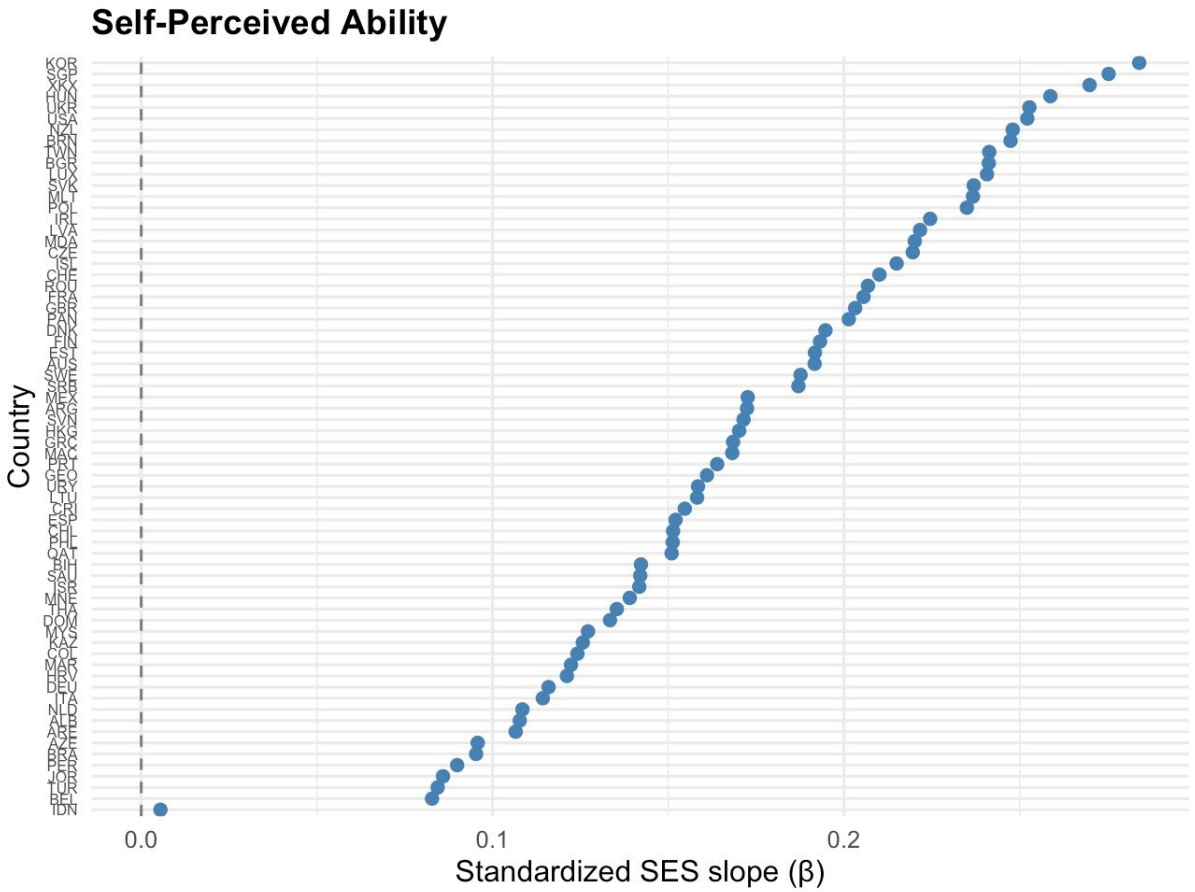

Figure 5. Standardized SES slope for self-perceived ability for each country

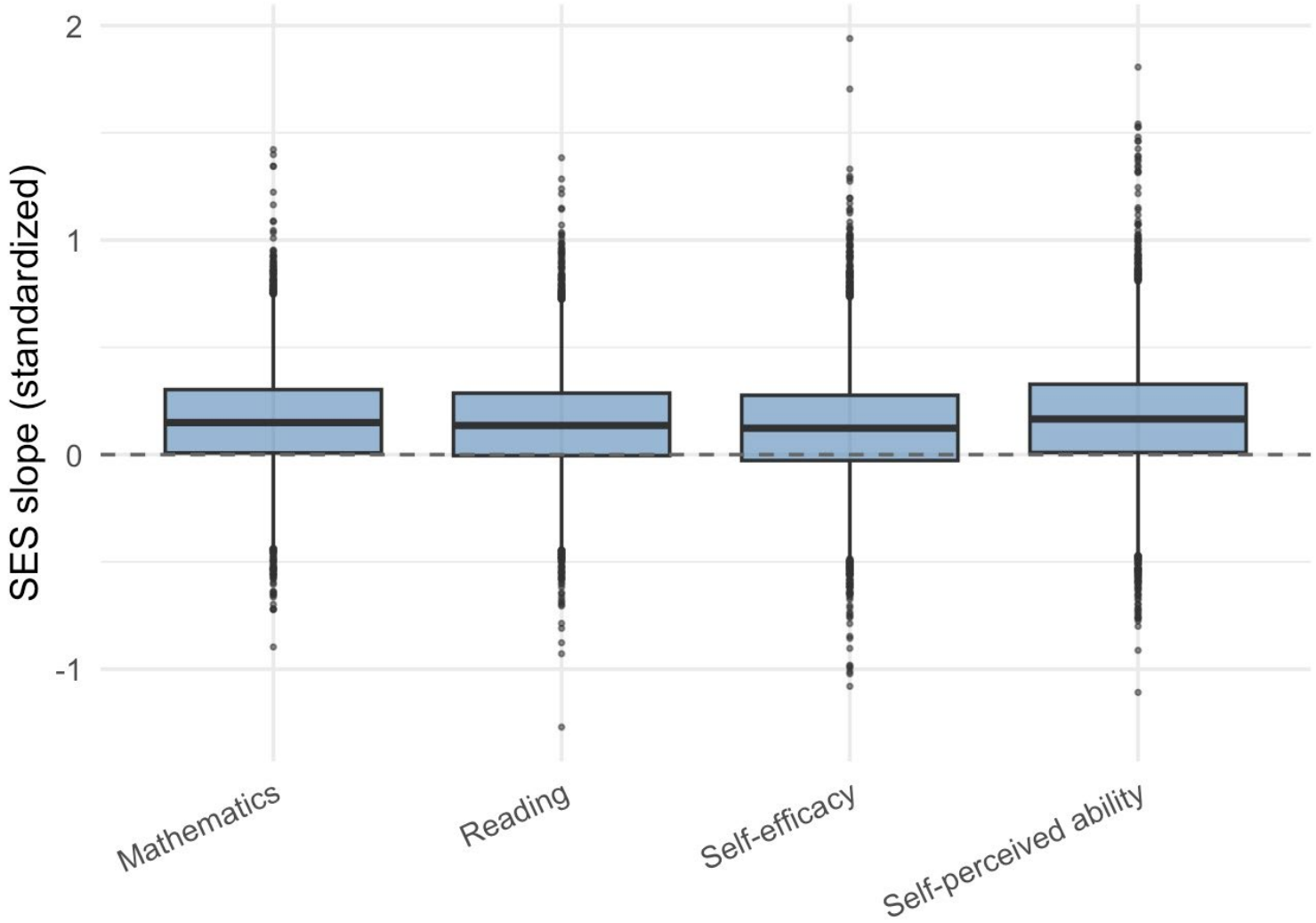

Figure 6. Boxplots for the standardized within-school SES slope.

Table 1. Correlation Matrix

Variable	1	2	3	4	5	6	7	8	9
1. Mathematics									
2. Reading	.81***								
3. Self-efficacy	.01***	.03***							
4. Self-perceived reading ability	.15***	.26***	.31***						
5. Perceived competitiveness	.03***	.01***	.20***	.12***					
6. Peer competitiveness	-.04***	-.03***	.09***	.08***	.15***				
7. SES (ESCS)	.40***	.39***	.07***	.18***	.04***	-.04***			
8. Grade	.01***	.01***	-.03***	.02***	-.00	-.06***	.04***		
9. Female	-.06***	.11***	-.02***	.08***	-.05***	-.01***	-.01***	-.02***	
10. Immigration background	.06***	.04***	-.01***	.01***	.01***	.04***	.03***	.05***	-.01**

*Note. N = 315095. Correlations are weighted using PISA final student weights. For mathematics and reading, correlations are pooled across 10 plausible values following Rubin's rules. ** p < .01, *** p < .001.*

S7. Robustness Checks

We carried out several sensitivity analyses to evaluate the robustness of the results. Table 2 and 3 show results with imputed data (Van Ginkel et al., 2020). Missing data were handled using multiple chained equations (MICE; Buuren & Groothuis-Oudshoorn, 2011) in R. We first applied listwise deletion to all variables with less than 5% missingness. Variables with more than 5% missing data (perceived competitiveness, peer competitiveness, self-efficacy, self-perceived ability) were subsequently imputed. The imputation model included all substantive variables of interest along with country and the final student weight as predictors. We used predictive mean matching since our imputed variables are all continuous with $m = 20$ imputed datasets and 20 iterations per imputation.

After imputation, we merged each imputed datasets with the plausible values for reading and mathematics. Within each imputed dataset, we then recomputed school-level peer competitiveness and senate-adjusted weights. We estimated all multilevel analyses separately on each imputed dataset and we pooled the estimates across imputations using Rubin's Rules (2018).[1]

Across MICE analyses, results for perceived competitiveness were essentially unchanged compared to the main models, both for the direct effect and for the negative interaction with SES. The effect of peer competitiveness was somewhat more sensitive to the handling of missingness. For self-views, the null main effect in the listwise models became small but significantly positive under MICE. For performance, the negative main effect remained negative under MICE but with a smaller magnitude. Additional analyses indicated

[1] Unlike our main analyses, which combined multiple imputation with PISA's BRR replication weights to account for the multi-stage sampling design, the robustness analyses reported here rely on model-based standard errors derived from the multi-level models. Combining MICE with BRR was computationally infeasible, as it would have required fitting each model 80 times per imputed dataset plus 10 more time for the model with academic performance (e.g., 1,620 fits for self-views and 16,2000 fits for performance). It is therefore possible that the standard errors are somewhat underestimated in comparison to the analyses with BRR. However, the purpose of this analysis is to assess the robustness of the direction and magnitude of the effects in the main analyses rather than providing inferences that are corrected for sampling design.

that this attenuation was partly driven by country-level missingness bias, as countries with high missingness on peer competitiveness systematically differed in average SES and achievement levels. The SES × peer competitiveness interaction remained consistently negative in direction across both modelling approaches and was statistically significant for mathematics in both listwise and MICE model, though it was no longer significant under MICE for reading or for self-views.

Second, we estimated models separately for boys and girls (Echazarra, 2020; Niederle & Vesterlund, 2007; Stoet & Geary, 2013; see Tables 4–7). The main effects of competitiveness were largely consistent across genders: peer competitiveness was negatively associated with achievement in both groups (although not statistically significant for boys in the models with SES interactions), and perceived competitiveness was positively associated with self-views. The moderating role of SES, however, differed by gender and was most pronounced for girls' self-views. For girls, higher perceived competitiveness was associated with a smaller gap in self-views between students from higher and lower SES backgrounds whereas for boys the same interaction was small and non-significant. Peer competitiveness showed no comparable buffering pattern for self-views in either group.

Third, we added individual trait competitiveness to the models with peer competitiveness (Table 8). Once individual competitiveness was controlled for, the previously negative effect of peer competitiveness on achievement became positive, and the positive effect on self-views became stronger. Inspection of the missing data showed that lower-SES students, and students in lower-SES, lower-performing schools, had disproportionately high rates of missing data on individual trait competitiveness, the item used to construct peer competitiveness at the school level. The SES × peer competitiveness interaction remained negative for mathematics but was null for reading and for self-views, suggesting that the moderating role of SES for peer-level competitiveness is for mathematics.

Fourth, we modelled peer competitiveness at the school level, excluding each student's own score when computing the school mean, to better reflect the broader competitive climate (Table 9). Under this specification, peer competitiveness was positively associated with both achievement and self-views, in contrast to the negative association observed in the main analyses. Taken together with the results from Table 8, this pattern is consistent with a suppressor effect, in which individual competitiveness obscures the contextual influence of the competitive peer environment. Once individual differences in competitiveness are accounted for, more competitive school environments appear to be associated with better outcomes rather than worse.

Finally, we raised the minimum school size from 20 to 30 students (Tables 10 and 11). Results for perceived competitiveness and its interaction with SES remained stable. The negative effect of peer competitiveness on achievement became somewhat stronger, and the positive association between peer competitiveness and self-perceived reading ability, visible in other specifications, was no longer significant. The SES × peer competitiveness interaction was no longer significant in this sample. Because smaller, lower-performing, and lower-SES schools were disproportionately excluded by the higher threshold, this pattern most likely reflects sample selection rather than a change in the underlying relation.

In summary, the main analyses seem robust to various sensitivity checks. The robustness checks revealed a suppressor effect and selective bias in the relationship between peer competitiveness and academic performance. The initially negative association between peer competitiveness and performance turned positive when individual competitiveness was controlled for, suggesting that individual trait competitiveness suppressed the positive contextual influence of competitive peer environments. Moreover, the positive relationship that emerged in the imputed models indicates that selective missingness, particularly among lower SES students and schools, affected the findings using listwise deletion, although only

marginally. While most of the interaction effects proved to be robust against various alternative model specifications, especially for academic performance, the direct effects were more sensitive to how the data were handled and which model specifications were used. Overall, the results appear consistent for academic performance, but the estimates for self-views were more dependent on modeling choice

Table 2. Robustness check: pooled estimates using MICE of the multilevel models for academic performance.

	Mathematics				Reading			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.167**(.063)	.146*(.062)	.166**(.063)	.166**(.063)	.027(.057)	.011(.056)	.027(.057)	.027(.057)
Socioeconomic Status	.150***(.003)	.158***(.003)	.149***(.003)	.149***(.003)	.143***(.001)	.151***(.002)	.142***(.001)	.142***(.001)
Peer Competitiveness	-.025***(.005)	-.025***(.005)			-.012*(.005)	-.010*(.005)		
SES * Peer		-.011***(.002)				-.002(.002)		
Perceived Competitiveness			.022***(.001)	.022***(.001)			.017***(.001)	.017***(.001)
SES * Perceived				-.006***(.001)				-.007***(.001)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.403	.399	.397	.397	.410	.406	.408	.408
Country	.520	.515	.526	.525	.472	.464	.475	.475
Residual	.677	.670	.677	.677	.698	.692	.698	.698
School SES Slope		.122				.125		
Fit Statistics								
AIC	1100822.6	1098036.8	1100441.0	1100419.3	1128535.4	1125914.8	1128315.7	1128280.0
BIC	1100976.8	1098224.1	1100595.3	1100584.6	1128689.7	1126102.1	1128470.0	1128445.3
LL	-550397.3	-549001.4	-550206.5	-550194.7	-564253.7	-562940.4	-564143.9	-564125.0
ICC Countries	.303	.304	.310	.310	.253	.251	.257	.257
ICC Schools	.182	.182	.177	.177	.192	.192	.189	.189
R^2 Marginal	.081	.084	.083	.083	.086	.089	.087	.087
R^2 Conditional	.527	.537	.529	.529	.493	.502	.494	.494
N Observations	449,602	449,602	449,602	449,602	449,602	449,602	449,602	449,602
N Schools	13,412	13,412	13,412	13,412	13,412	13,412	13,412	13,412
N Countries	69	69	69	69	69	69	69	69

Note. Pooled estimates with imputed data (m = 20). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Controls include immigration status, gender, and grade.

Table 3. Robustness check: pooled estimates using MCIE of the multilevel models for self-views.

	Self-efficacy				Self-perceived ability			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.024(.027)	.019(.027)	.015(.026)	.015(.026)	-.053(.027)	-.053(.027)	-.058*(.027)	-.058*(.027)
Socioeconomic Status	.124***(.002)	.124***(.003)	.109***(.002)	.109***(.002)	.181***(.002)	.181***(.002)	.173***(.002)	.173***(.002)
Peer Competitiveness	.013***(.003)	.023(.019)			.016***(.003)	.016***(.003)		
SES * Peer		-.004(.002)				.000(.002)		
Perceived Competitiveness			.197***(.002)	.197***(.002)			.115***(.002)	.115***(.002)
SES * Perceived				-.010***(.002)				-.005**(.002)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.135	.090	.131	.130	.154	.154	.152	.152
Country	.222	.221	.216	.216	.222	.222	.219	.219
Residual	.961	.959	.943	.943	.940	.940	.934	.934
School SES Slope		.101						
Fit Statistics								
AIC	1391666.5	1391461.4	1374311.5	1374266.8	1373663.4	1373665.2	1367589.5	1367579.1
BIC	1391820.7	1391648.6	1374465.7	1374432.0	1373817.6	1373830.5	1367743.8	1367744.3
Log-Likelihood	-695819.2	-695713.7	-687141.7	-687118.4	-686817.7	-686817.6	-683780.8	-683774.5
ICC Countries	.050	.050	.049	.049	.052	.052	.051	.051
ICC Schools	.018	.012	.018	.018	.025	.025	.025	.025
R^2 Marginal	.019	.020	.057	.057	.045	.045	.057	.057
R^2 Conditional	.086	.094	.120	.120	.118	.118	.128	.128
N Observations	449,602	449,602	449,602	449,602	449,602	449,602	449,602	449,602
N Schools	13,412	13,412	13,412	13,412	13,412	13,412	13,412	13,412
N Countries	69	69	69	69	69	69	69	69

Note. Multiple imputation (m = 20). * p < .05, ** p < .01, *** p < .001. Controls include immigration status, gender, and grade.

Table 4. Robustness check: academic performance (boys only).

	Mathematics				Reading			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.064***(.005)	.051***(.005)	.064***(.005)	.064***(.005)	.097***(.004)	.086***(.004)	.097***(.004)	.097***(.004)
Socioeconomic Status	.158***(.004)	.169***(.004)	.156***(.004)	.157***(.004)	.150***(.004)	.161***(.004)	.149***(.004)	.149***(.004)
Peer Competitiveness	.003(.006)	.002(.006)			.005(.006)	.005(.006)		
SES * Peer		-.011**(.004)				-.002(.004)		
Perceived Competitiveness			.022***(.003)	.022***(.003)			.011***(.003)	.011***(.003)
SES * Perceived				-.010**(.003)				-.010***(.003)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.392	.387	.392	.392	.401	.394	.401	.401
Country	.502	.500	.504	.503	.451	.447	.451	.451
Residual	.685	.676	.685	.685	.711	.701	.711	.711
School SES Slope		.147				.149		
Fit Statistics								
AIC	392827.6	391540.0	392706.5	392671.2	403700.6	402459.1	403663.9	403637.9
BIC	392956.7	391698.9	392835.6	392810.2	403829.6	402618.0	403793.0	403777.0
Log-Likelihood	-196400.8	-195754.0	-196340.2	-196321.6	-201837.3	-201213.6	-201819.0	-201805.0
ICC Countries	.288	.292	.289	.289	.234	.236	.234	.234
ICC Schools	.176	.175	.175	.175	.185	.184	.185	.185
R^2 Marginal	.081	.084	.081	.081	.073	.077	.073	.073
R^2 Conditional	.507	.523	.508	.508	.461	.478	.461	.461
N Observations	151,750	151,750	151,750	151,750	151,750	151,750	151,750	151,750
N Schools	9,340	9,340	9,340	9,340	9,340	9,340	9,340	9,340
N Countries	69	69	69	69	69	69	69	69

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status and grade.

Table 5. Robustness check: academic performance (girls only).

	Mathematics				Reading			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.039***(.005)	.026***(.004)	.041***(.005)	.041***(.005)	.075***(.005)	.064***(.005)	.076***(.005)	.076***(.005)
Socioeconomic Status	.182***(.005)	.196***(.005)	.181***(.005)	.181***(.005)	.178***(.004)	.192***(.005)	.177***(.004)	.176***(.004)
Peer Competitiveness	-.006(.007)	-.007(.008)			.003(.006)	.001(.007)		
SES * Peer		-.012***(.003)				-.006*(.003)		
Perceived Competitiveness			.023***(.003)	.023***(.003)			.015***(.003)	.015***(.003)
SES * Perceived				-.008**(.003)				-.008**(.003)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.379	.373	.377	.377	.383	.376	.383	.383
Country	.513	.515	.517	.517	.470	.467	.471	.471
Residual	.688	.679	.687	.687	.708	.699	.708	.708
School SES Slope		.141				.144		
Fit Statistics								
AIC	422897.4	421561.0	422736.0	422722.6	432298.4	431058.1	432244.3	432227.0
BIC	423027.5	421721.1	422866.1	422862.6	432428.4	431218.2	432374.4	432367.1
Log-Likelihood	-211435.7	-210764.5	-211355.0	-211347.3	-216136.2	-215513.1	-216109.2	-216099.5
ICC Countries	.300	.307	.303	.303	.254	.257	.255	.255
ICC Schools	.163	.161	.161	.161	.169	.167	.169	.169
R^2 Marginal	.086	.091	.088	.087	.085	.090	.085	.085
R^2 Conditional	.509	.527	.511	.511	.472	.488	.473	.473
N Observations	163,345	163,345	163,345	163,345	163,345	163,345	163,345	163,345
N Schools	9,410	9,410	9,410	9,410	9,410	9,410	9,410	9,410
N Countries	69	69	69	69	69	69	69	69

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status and grade.

Table 6. Robustness check: self-views (boys only).

	Self-efficacy				Self-perceived ability			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.035***(.005)	.031***(.005)	.031***(.005)	.031***(.005)	-.069***(.006)	-.081***(.011)	-.071***(.006)	-.071***(.006)
Socioeconomic Status	.137***(.004)	.136***(.003)	.117***(.004)	.118***(.004)	.179***(.004)	.182***(.003)	.168***(.004)	.169***(.004)
Peer Competitiveness	.002(.029)	-.004(.034)			.007(.015)	.003(.045)		
SES * Peer		-.007*(.003)				-.001(.004)		
Perceived Competitiveness			.232***(.004)	.232***(.004)			.126***(.004)	.126***(.004)
SES * Perceived				-.005(.004)				-.002(.004)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.181	.176	.172	.172	.188	.175	.184	.184
Country	.212	.212	.207	.207	.240	.241	.236	.236
Residual	.985	.975	.961	.961	.957	.947	.950	.950
School SES Slope		.155				.165		
Fit Statistics								
AIC	491665.3	490898.1	483952.4	483950.7	483442.8	482471.7	481088.7	481089.8
BIC	491794.4	491056.9	484081.5	484089.7	483571.9	482630.6	481217.8	481228.8
Log-Likelihood	-245819.6	-245433.0	-241963.2	-241961.3	-241708.4	-241219.9	-240531.4	-240530.9
ICC Countries	.043	.044	.043	.043	.057	.059	.056	.056
ICC Schools	.031	.030	.030	.030	.035	.031	.034	.034
R^2 Marginal	.021	.021	.070	.070	.035	.036	.049	.049
R^2 Conditional	.093	.113	.138	.138	.124	.146	.135	.135
N Observations	151,750	151,750	151,750	151,750	151,750	151,750	151,750	151,750
N Schools	9,340	9,340	9,340	9,340	9,340	9,340	9,340	9,340
N Countries	69	69	69	69	69	69	69	69

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status and grade.

Table 7. Robustness check: self-views (girls only).

	Self-efficacy				Self-perceived ability			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.003(.006)	-.002(.006)	.013*(.006)	.013*(.006)	.098***(.004)	.097***(.004)	.104***(.004)	.104***(.004)
Socioeconomic Status	.115***(.005)	.119***(.004)	.105***(.005)	.103***(.005)	.184***(.005)	.185***(.005)	.179***(.005)	.178***(.005)
Peer Competitiveness	.016(.018)	.008(.022)			.029**(.010)	.029**(.010)		
SES * Peer		-.006*(.003)				-.013**(.004)		
Perceived Competitiveness			.151***(.004)	.151***(.004)			.097***(.004)	.097***(.004)
SES * Perceived				-.029***(.004)				-.019***(.003)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.166	.158	.163	.163	.177	.177	.177	.177
Country	.269	.271	.264	.263	.229	.229	.231	.230
Residual	.927	.919	.916	.916	.906	.906	.901	.901
School SES Slope		.132						
Fit Statistics								
AIC	509394.9	508578.4	505628.1	505474.9	502348.2	502329.2	500770.8	500707.3
BIC	509524.9	508738.5	505758.1	505615.0	502478.2	502469.3	500900.9	500847.4
Log-Likelihood	-254684.4	-254273.2	-252801.0	-252723.5	-251161.1	-251150.6	-250372.4	-250339.7
ICC Countries	.075	.078	.074	.074	.058	.058	.059	.059
ICC Schools	.029	.027	.028	.029	.035	.035	.035	.035
R^2 Marginal	.016	.017	.039	.039	.042	.042	.051	.051
R^2 Conditional	.119	.136	.138	.138	.131	.131	.140	.140
N Observations	163,345	163,345	163,345	163,345	163,345	163,345	163,345	163,345
N Schools	9,410	9,410	9,410	9,410	9,410	9,410	9,410	9,410
N Countries	69	69	69	69	69	69	69	69

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status and grade.

Table 8. Robustness check: academic performance and self-views including individual trait competitiveness.

	Mathematics		Reading		Self-efficacy		Self-perceived ability	
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.118***(.005)	.105***(.005)	-.010*(.004)	-.020***(.004)	.014***(.004)	.014***(.004)	-.081***(.005)	-.089***(.005)
Socioeconomic Status	.152***(.004)	.162***(.004)	.145***(.003)	.154***(.003)	.099***(.003)	.099***(.003)	.167***(.003)	.170***(.003)
Peer Competitiveness	.061***(.003)	.061***(.003)	.058***(.002)	.058***(.002)	.028***(.003)	.028***(.003)	.031***(.002)	.029***(.003)
SES * Peer		-.008***(.002)		.000(.002)		-.005(.003)		-.003(.003)
Individual Competitiveness	.068***(.003)	.067***(.003)	.058***(.002)	.057***(.002)	.297***(.003)	.297***(.003)	.168***(.003)	.167***(.003)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.387	.384	.393	.389	.130	.130	.156	.146
Country	.530	.530	.477	.472	.215	.215	.218	.220
Residual	.687	.680	.707	.700	.920	.920	.924	.918
School SES Slope		.122		.124				.120
Fit Statistics								
AIC	806972.7	804861.0	825116.4	823146.4	974979.5	974975.1	978485.2	977273.7
BIC	807132.7	805052.8	825276.3	823338.3	975139.4	975145.7	978645.1	977465.6
Log-Likelihood	-403471.4	-402412.5	-412543.2	-411555.2	-487474.8	-487471.6	-489227.6	-488618.8
ICC Countries	.312	.316	.258	.258	.051	.051	.052	.053
ICC Schools	.166	.166	.175	.175	.019	.019	.026	.024
R^2 Marginal	.094	.097	.092	.095	.106	.106	.074	.074
R^2 Conditional	.527	.539	.485	.496	.168	.168	.146	.158
N Observations	315,095	315,095	315,095	315,095	315,095	315,095	315,095	315,095
N Schools	10,021	10,021	10,021	10,021	10,021	10,021	10,021	10,021
N Countries	69	69	69	69	69	69	69	69

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status and grade.

Table 9. Robustness check: academic performance and self-views with school-level competitiveness.

	Mathematics		Reading		Self-efficacy		Self-perceived ability	
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.126***(.005)	.112***(.005)	-.004(.004)	-.014***(.004)	.047***(.005)	.043***(.007)	-.062***(.005)	-.071***(.007)
Socioeconomic Status	.157***(.004)	.167***(.004)	.149***(.003)	.159***(.003)	.120***(.003)	.122***(.004)	.179***(.003)	.182***(.002)
Peer Competitiveness	.083***(.003)	.082***(.003)	.077***(.002)	.076***(.002)	.121***(.003)	.119***(.004)	.084***(.002)	.082***(.003)
SES * Peer		-.008***(.002)		-.002(.002)		-.006*(.003)		-.005(.003)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.386	.383	.392	.388	.127	.122	.155	.145
Country	.530	.530	.477	.472	.221	.218	.219	.220
Residual	.689	.682	.709	.702	.962	.957	.937	.931
School SES Slope		.124		.125		.107		.123
Fit Statistics								
AIC	809385.3	807221.2	826600.2	824592.2	1002444.0	1001581.8	987285.9	986021.8
BIC	809534.6	807402.4	826749.5	824773.5	1002593.2	1001763.1	987435.2	986203.0
Log-Likelihood	-404678.7	-403593.6	-413286.1	-412279.1	-501208.0	-500773.9	-493629.0	-492993.9
ICC Countries	.310	.314	.257	.257	.049	.049	.050	.052
ICC Schools	.165	.164	.174	.174	.016	.015	.025	.022
R^2 Marginal	.093	.096	.090	.094	.033	.033	.051	.052
R^2 Conditional	.524	.536	.483	.493	.096	.105	.123	.136
N Observations	315,095	315,095	315,095	315,095	315,095	315,095	315,095	315,095
N Schools	10,021	10,021	10,021	10,021	10,021	10,021	10,021	10,021
N Countries	69	69	69	69	69	69	69	69

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status, gender, and grade.

Table 10. Robustness check: academic performance (schools with ≥30 students).

	Mathematics				Reading			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.143***(.008)	.125***(.008)	.149***(.008)	.149***(.008)	.042***(.008)	.026**(.008)	.046***(.008)	.047***(.008)
Socioeconomic Status	.164***(.007)	.164***(.004)	.163***(.007)	.162***(.007)	.158***(.006)	.160***(.003)	.157***(.006)	.156***(.006)
Peer Competitiveness	-.081***(.013)	-.081***(.013)			-.053***(.010)	-.052***(.010)		
SES * Peer		-.006(.004)				.002(.003)		
Perceived Competitiveness			.027***(.005)	.028***(.005)			.020***(.004)	.022***(.004)
SES * Perceived				-.013**(.004)				-.016***(.004)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.386	.388	.375	.376	.385	.384	.379	.379
Country	.482	.483	.500	.500	.451	.448	.464	.464
Residual	.670	.662	.670	.670	.692	.684	.692	.692
School SES Slope		.123				.125		
Fit Statistics								
AIC	507535.5	505110.0	507325.3	507259.1	519118.0	516883.3	518966.1	518898.1
BIC	507677.4	505282.2	507467.2	507411.1	519259.9	517055.6	519108.0	519050.1
Log-Likelihood	-253753.8	-252538.0	-253648.6	-253614.5	-259545.0	-258424.7	-259469.0	-259434.0
ICC Countries	.280	.284	.298	.298	.245	.246	.257	.257
ICC Schools	.180	.183	.168	.168	.178	.180	.171	.172
R^2 Marginal	.087	.086	.087	.087	.085	.085	.086	.086
R^2 Conditional	.506	.522	.512	.512	.472	.485	.477	.477
N Observations	186,141	186,141	186,141	186,141	186,141	186,141	186,141	186,141
N Schools	4,762	4,762	4,762	4,762	4,762	4,762	4,762	4,762
N Countries	64	64	64	64	64	64	64	64

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status, gender, and grade.

Table 11. Robustness check: self-views (schools with ≥30 students).

	Self-efficacy				Self-perceived ability			
	Model 1a	**Model 1b**	**Model 2a**	**Model 2b**	**Model 3a**	**Model 3b**	**Model 4a**	**Model 4b**
Intercept	.074***(.011)	.073***(.011)	.075***(.010)	.076***(.010)	-.049***(.010)	-.061***(.011)	-.049***(.009)	-.048***(.009)
Socioeconomic Status	.125***(.007)	.125***(.007)	.112***(.007)	.110***(.007)	.187***(.007)	.189***(.003)	.179***(.007)	.178***(.007)
Peer Competitiveness	-.005(.027)	-.005(.027)			.008(.013)	.006(.014)		
SES * Peer		-.007(.008)				-.002(.004)		
Perceived Competitiveness			.184***(.007)	.186***(.007)			.112***(.006)	.113***(.006)
SES * Perceived				-.022***(.006)				-.014**(.005)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Random effects								
School	.135	.135	.125	.125	.148	.143	.145	.145
Country	.252	.252	.242	.242	.244	.244	.239	.239
Residual	.964	.964	.949	.948	.946	.939	.941	.941
School SES Slope						.124		
Fit Statistics								
AIC	634643.6	634639.8	628538.0	628449.4	628224.5	626910.7	625935.4	625901.8
BIC	634785.4	634791.8	628679.9	628601.4	628366.4	627083.0	626077.2	626053.8
Log-Likelihood	-317307.8	-317304.9	-314255.0	-314209.7	-314098.3	-313438.3	-312953.7	-312935.9
ICC Countries	.063	.063	.060	.060	.061	.062	.059	.059
ICC Schools	.018	.018	.016	.016	.022	.021	.022	.022
R^2 Marginal	.018	.018	.050	.051	.045	.045	.056	.057
R^2 Conditional	.097	.097	.123	.123	.124	.138	.133	.133
N Observations	186,141	186,141	186,141	186,141	186,141	186,141	186,141	186,141
N Schools	4,762	4,762	4,762	4,762	4,762	4,762	4,762	4,762
N Countries	64	64	64	64	64	64	64	64

Note. * p < .05, ** p < .01, *** p < .001. Controls include immigration status, gender, and grade.

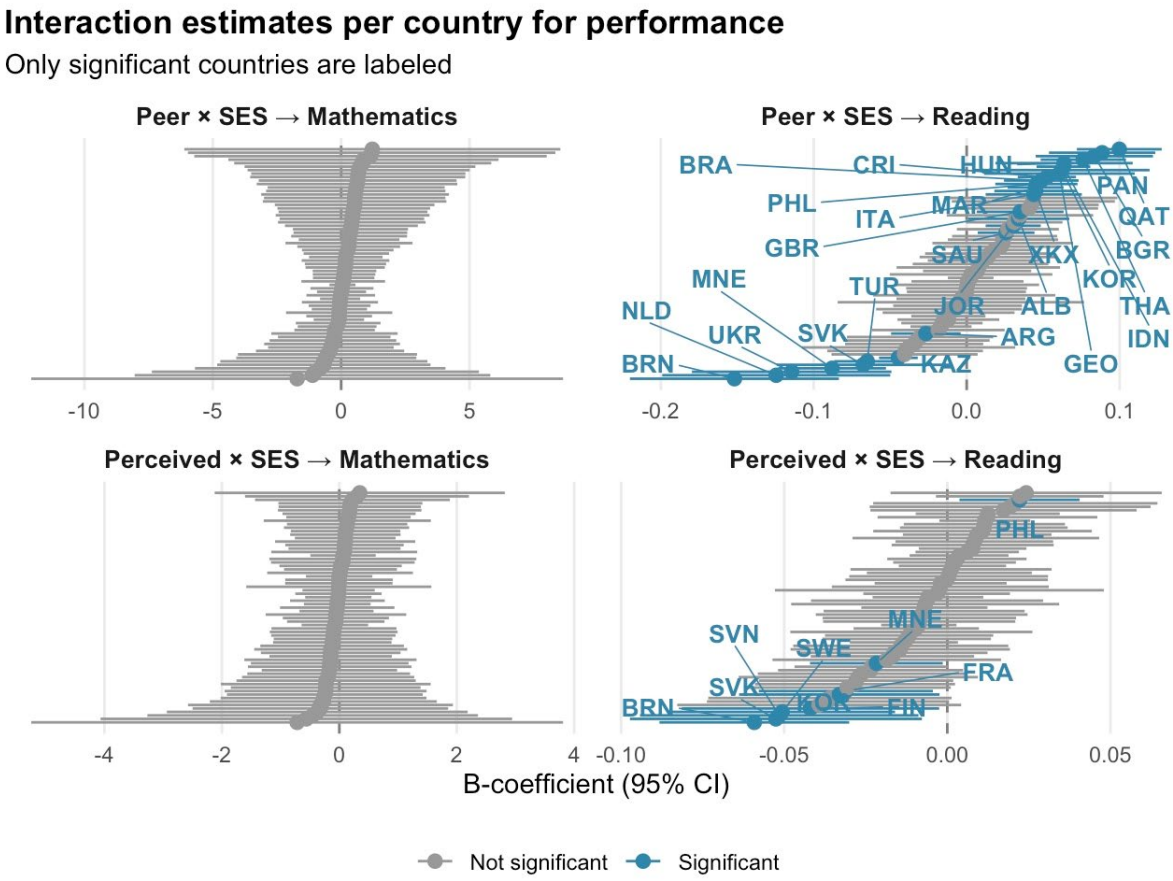

Figure 7. Interaction estimates between SES and competitiveness on performance for each country.

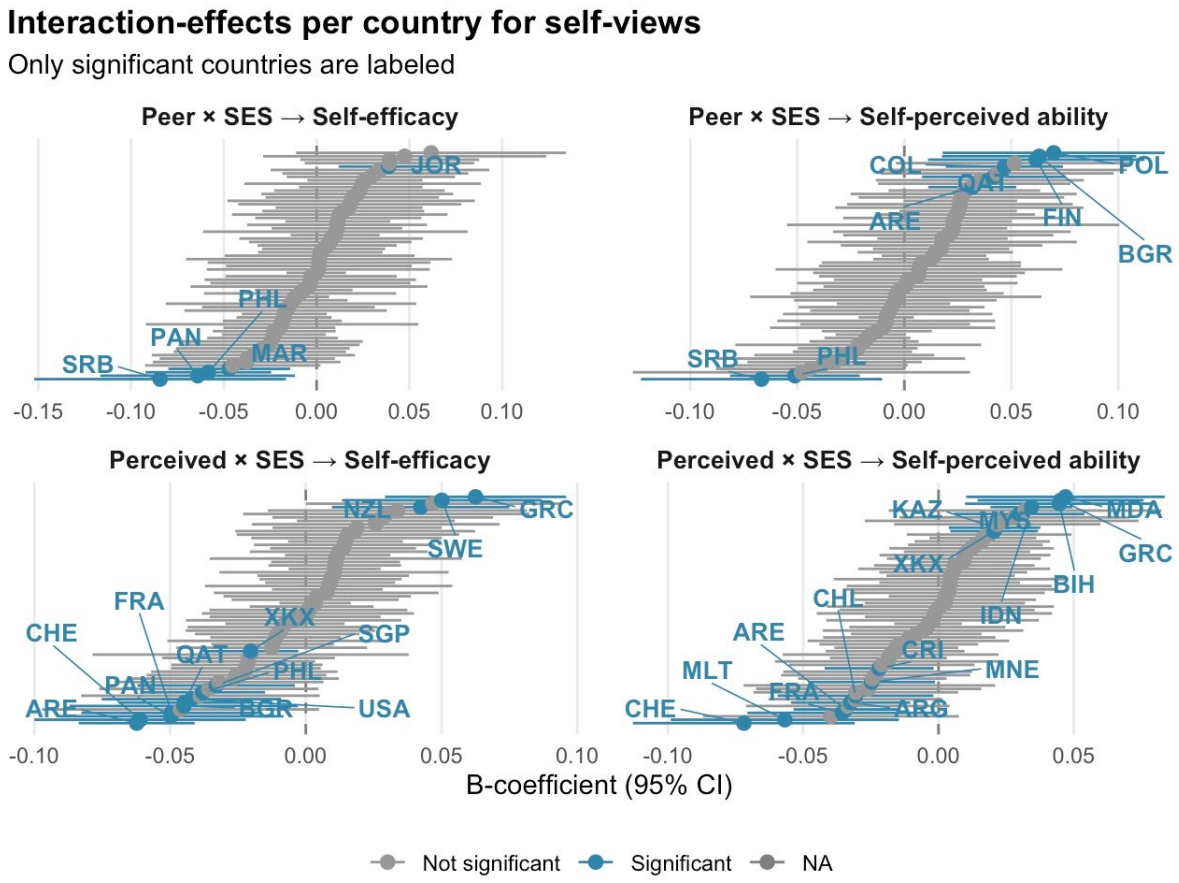

Figure 8. Interaction estimates between SES and competitiveness on self-views for each country.

References

Buuren, S. V., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in *R*. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03

Echazarra, A. (2020). *Do students learn in co-operative or competitive environments* (PISA in Focus 107; PISA in Focus, Vol. 107). https://doi.org/10.1787/e7a5f3e4-en

Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, *122*(3), 1067–1101. https://doi.org/10.1162/qjec.122.3.1067

OECD (2019), *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/acd78851-en

Rubin, D. B. (2018). Multiple imputation. In Flexible imputation of missing data, second edition (pp. 29-62). Chapman and Hall/CRC.

Van Ginkel, J. R., Linting, M., Rippe, R. C. A., & Van Der Voort, A. (2020). Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *Journal of Personality Assessment*, *102*(3), 297–308. https://doi.org/10.1080/00223891.2018.1530680

Stoet, G., & Geary, D. C. (2013). Sex Differences in Mathematics and Reading Achievement Are Inversely Related: Within- and Across-Nation Assessment of 10 Years of PISA Data. *PLoS ONE*, *8*(3), e57988. https://doi.org/10.1371/journal.pone.0057988

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical methods in medical research*, *8*(1), 3-15.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in medicine*, *30*(4), 377-399.