

Supplementary materials

Supplementary materials for *Longitudinal associations between German and English language competencies in Grade 9 and educational success*

Table S1. Competence levels according to BT 2015

German - reading				
Level	Value range	HSA	MSA	Coding
V	≥ 660	Optimal standard	Optimal standard	above III
IV	580 to 659	Optimal standard	Normative standard plus	above III
III	500 to 579	Normative standard plus	Normative standard	III
II	420 to 499	Normative standard	Minimal standard	II
Ib	350 to 419	Minimal standard	Below minimal standard	Ib
Ia	< 350	Below minimal standard	Below minimal standard	Ia
English - reading				
Level	Value range	HSA	MSA	Coding
C1	≥ 700	Optimal standard	Optimal standard	Above B1.2
B2.2	650 to 699	Optimal standard	Optimal standard	Above B1.2
B2.1	600 to 549	Optimal standard	Normative standard plus	Above B1.2
B1.2	550 to 599	Optimal standard	Normative standard	B1.2
B1.1	500 to 549	Optimal standard	Minimal standard	B1.1
A2.2	450 to 499	Normative standard plus	Minimal standard	A2.2
A2.1	400 to 449	Normative standard	Below minimal standard	A2.1
A1.2	350 to 399	Minimal standard	Below minimal standard	A1.2
A1.1	< 350	Below minimal standard	Below minimal standard	A1.1

Note: Compare Stanat et al. 2016.

Section 1 Treatment of missing values

Table S2. List of variables in imputation and background models

Var ID	Variable	Source
wle_hoeren	WLE German listening	BT 2015
wle_lesen	WLE German reading	BT 2015
wle_ortho	WLE German orthography	BT 2015
wle_EL	WLE English listening	BT 2015
wle_ER	WLE English reading	BT 2015
BEFKIwle	Fluid intelligence - figural reasoning	BT 2015
TR_geschlecht	Sex	BT 2015
trnotedeu_r_rec	Mid-term grade German	BT 2015
trnoteeng_r_rec	Mid-term grade English	BT 2015
trnotemat_r_rec	Mid-term grade Math	BT 2015
HISEI_rec	Highest household ISEI	BT 2015
zhg_rec	Binary migration background	BT 2015
schulform	School type	BT 2015
HB_H_Abschluss_rec ^a	Highest recorded degree	administrative data
hb_zap_g_de ^a	Grade German ZAP base level	administrative data
hb_zap_e_de ^a	Grade German ZAP extended level	administrative data
hb_zap_g_eng ^a	Grade English ZAP extended level	administrative data
hb_zap_e_eng ^a	Grade English ZAP extended level	administrative data

Note: ^a Variable is not used as part of the background model for the plausible values.

Table S3. Missing values

	<i>N</i> missing	Share missing
Highest degree	23	0.02
Grade German	606	0.43
Grade English	610	0.43
Sex	0	0.00
Migration background	254	0.18
ISEI	321	0.23
School track	0	0.00

Note: The table reports missing values as share of the sample of students for which linkage is possible and which are used during the imputation process ($N = 1423$), ignoring later reduction in sample sizes.

Section 2 Predicted probabilities

Table S4. Predicted probabilities highest degree

	Degree			
	No degree	HSA ¹	MSA ²	Abitur ³
German - reading				
Ia	0.07	0.36	0.45	0.11
Ib	0.03	0.22	0.52	0.23
II	0.02	0.15	0.50	0.33
III	0.01	0.07	0.38	0.54
above III	0.01	0.05	0.33	0.61
English - reading				
A1.1	0.12	0.50	0.34	0.04
A1.2	0.04	0.32	0.52	0.12
A2.1	0.02	0.19	0.57	0.22
A2.2	0.01	0.10	0.50	0.39
B1.1	0.01	0.06	0.41	0.52
B1.2	0.00	0.03	0.29	0.67
Above B1.2	0.00	0.03	0.25	0.72

Note: Predicted probabilities are calculated with covariates at their means.

¹ lower secondary school-leaving certificate

² intermediate secondary school-leaving certificate

³ general university entrance qualification

Table S5. Predicted probabilities highest degree

# achieved HSA minimal standards	Degree			
	No degree	HSA ¹	MSA ²	Abitur ³
0	0.16	0.52	0.27	0.05
1	0.11	0.47	0.35	0.07
2	0.04	0.27	0.50	0.19
3	0.01	0.11	0.45	0.43

Note: Predicted probabilities are calculated with covariates at their means.

¹ lower secondary school-leaving certificate

² intermediate secondary school-leaving certificate

³ general university entrance qualification

Table S6. Predicted probabilities ZAP grades

	Grade				
	1	2	3	4	5/6
German - reading					
Ia	0.02	0.22	0.47	0.24	0.05
Ib	0.04	0.31	0.45	0.17	0.03
II	0.07	0.41	0.40	0.11	0.02
III	0.12	0.51	0.30	0.06	0.01
above III	0.14	0.53	0.27	0.05	0.01
German - reading					
A1.1	0.01	0.08	0.32	0.43	0.16
A1.2	0.04	0.24	0.45	0.23	0.05
A2.1	0.08	0.37	0.40	0.13	0.02
A2.2	0.15	0.48	0.29	0.07	0.01
B1.1	0.41	0.45	0.11	0.02	0.00
B1.2	0.52	0.39	0.08	0.01	0.00
Above B1.2	0.88	0.11	0.01	0.00	0.00

Note: Predicted probabilities are calculated with covariates as observed; grades are language specific grades.

Section 3 Parallel regression assumption

The proportional odds or parallel regression assumption for ordered logistic or ordered probit regressions demands that, at a given probability level, a one unit increase in a variable leads to the same increase or decrease of the probability of observing a specific outcome for each of the possible outcome categories. One way to deal with violations of the assumption is to use partial proportional odds models instead, which allow certain parameters to vary and assume parallel regression lines for the others:

$$P(Y_i > j) = \frac{\exp(\alpha_j + x_{1j}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_{3j})}{1 + \exp(\alpha_j + x_{1j}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_{3j})}.$$

In this case, the coefficients for variables x_1 and x_2 are constrained across the outcome categories, while coefficients for variable x_3 are allowed to vary (for example Williams, 2006; Yee, 2010).

We conducted Wald tests to assess the parallel regression assumption. We implemented the tests for the unadjusted models with our main variables of interest using the first imputed dataset and survey weights. In case of German reading, we find no indication, that the assumption is violated for the model with our primary outcome, highest degree ($F(8, 53) = 1.22$, p-value = 0.37). In all other models we find indications that the parallel regression assumption is violated for at least one coefficient. In order to assess the consequences of the violation for our main models, we re-estimated them using partial proportional odds models. In case of English grades as the dependent variable, this was only possible when further reducing categories on the competence levels due to empty cells. A comparison of the empty models between ordered logistic regressions and partial proportional odds models showed that despite minor changes in the predicted probabilities, the overall conclusions are not affected and the focus on the better known, simpler models seems justified.

Section 4 Results for additional domains

Table S7. Results from ordered logistic regression of highest degree on alternative domains - proficiency levels

	German				German				English			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	β	SE	β	SE	β	SE	β	SE	β	SE	β	SE
Proficiency levels												
German - listening												
Ib	1.22*	0.31	1.11*	0.30								
II	1.64*	0.31	1.34*	0.33								
III	2.52*	0.33	1.90*	0.33								
Above III	3.16*	0.36	2.21*	0.35								
German - orthography												
Ib					0.86*	0.29	0.75*	0.28				
II					1.71*	0.29	1.50*	0.30				
III					2.92*	0.33	2.42*	0.33				
Above III					4.49*	0.32	3.51*	0.37				
English - listening												
A1.2									1.43*	0.39	1.37*	0.40
A2.1									2.31*	0.35	2.20*	0.37
A2.2									3.61*	0.32	3.39*	0.34
B1.1									4.45*	0.38	4.12*	0.37
B1.2									5.67*	0.46	5.03*	0.43
Above B1.2									6.54*	0.47	5.72*	0.47
Controls												
ISEI			0.03*	0.00			0.02*	0.00			0.01*	0.00
Sex - female			0.37*	0.15			0.12	0.17			0.48*	0.15
Migration background - yes			-0.15	0.15			-0.15	0.12			-0.59*	0.16
Multi-track school			-0.22	0.37			-0.19	0.26			-0.24	0.29
Academic track school			1.65*	0.41			1.00*	0.40			0.78	0.41
Pseudo- R^2	0.09		0.18		0.19		0.23		0.24		0.28	
Δ Pseudo- R^2			0.09				0.04				0.04	
N	1183		1183		1183		1183		1183		1183	

Notes: reference categories: proficiency levels German = Ia, proficiency levels English = A1.1, sex = male, migration background = no, school type = comprehensive school; reported coefficients are on log-odds scale; results are based on $m = 15$ imputed datasets; McFadden Pseudo- R^2 is calculated as the average of all Pseudo- R^2 's across the 15 imputations assuming independence between the observations in the sample.

* $p < 0.05$ (two tailed tests).

Table S8. Results from ordered logistic regression of highest degree on alternative domains - continuous scores

	German				German				English			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	β	SE	β	SE	β	SE	β	SE	β	SE	β	SE
Proficiency score												
German - listening												
	0.01*	0.00	0.01*	0.00								
German - orthography												
					0.01*	0.00	0.01*	0.00				
English - listening												
									0.02*	0.00	0.01*	0.00
Controls												
ISEI			0.03*	0.00			0.02*	0.00			0.01*	0.00
Sex - female			0.36*	0.15			0.10	0.17			0.48*	0.14
Migration background - yes			-0.14	0.15			-0.20	0.13			-0.60*	0.16
Multi-track school			-0.26	0.38			-0.22	0.27			-0.25	0.28
Academic track			1.62*	0.41			1.04*	0.39			0.65	0.39
Pseudo- R^2	0.09		0.18		0.18		0.22		0.25		0.28	
Δ Pseudo- R^2			0.09				0.04				0.03	
N	1183		1183		1183		1183		1183		1183	

Notes: reference categories: sex = male, migration background = no, school type = comprehensive school; reported coefficients are on log-odds scale; results are based on $m = 15$ imputed datasets; McFadden Pseudo- R^2 is calculated as the average of all Pseudo- R^2 's across the 15 imputations assuming independence between the observations in the sample.

* $p < 0.05$ (two tailed tests).

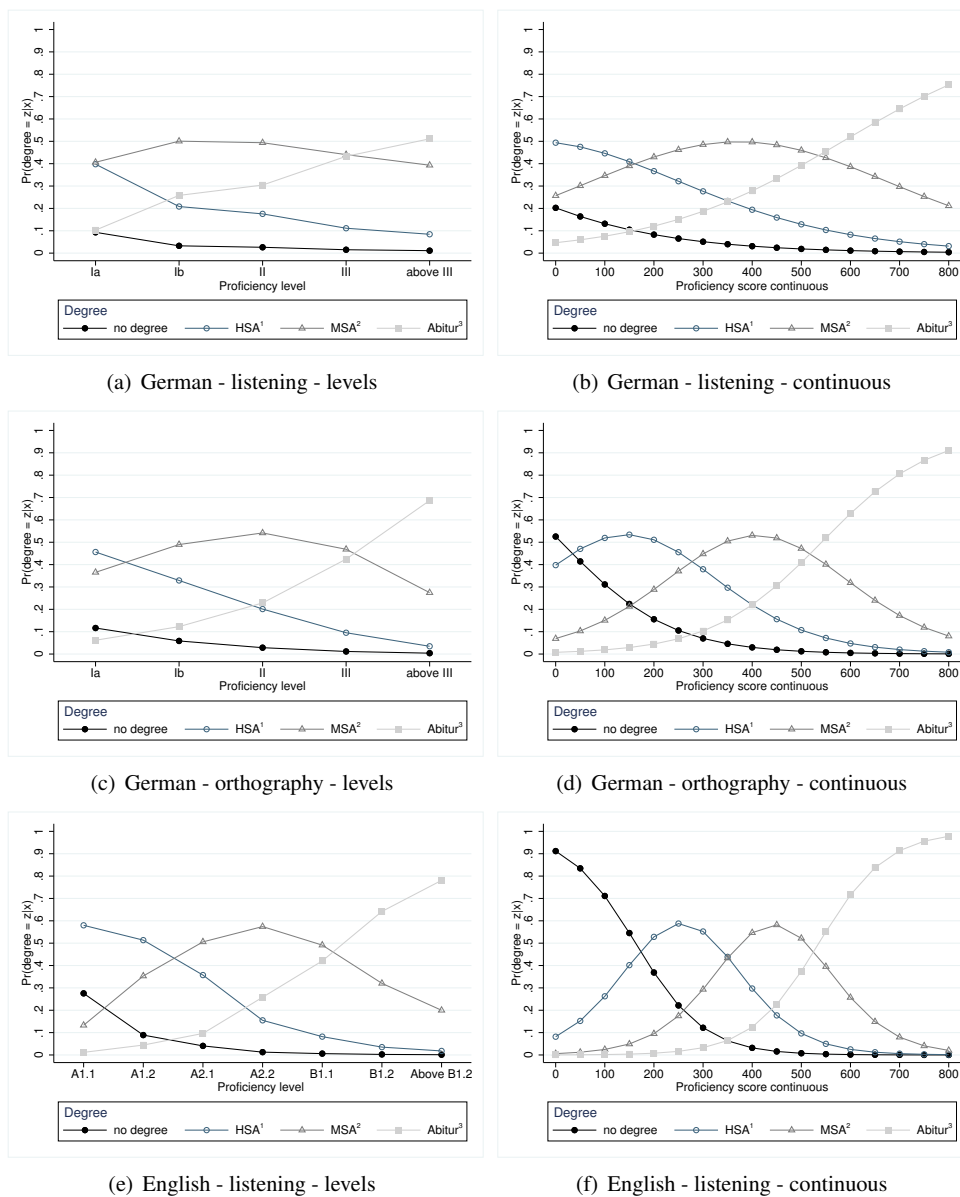


Figure S1. Predicted probabilities of achieving a specific degree conditional on the alternative domains (German listening, German orthography and English listening) from adjusted models, controlling for sex, migration background, household socio-economic status, and school track.

- ¹ lower secondary school-leaving certificate
² intermediate secondary school-leaving certificate
³ general university entrance qualification

Table S9. Results from ordered logistic regressions of grades on alternative domains - proficiency levels

	Grade German				Grade German				Grade English			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	β	SE	β	SE	β	SE	β	SE	β	SE	β	SE
Proficiency levels												
German - listening												
Ib	-0.24	0.37	-0.14	0.34								
II	-0.66*	0.27	-0.47	0.25								
III	-0.84*	0.31	-0.58	0.29								
Above III	-1.56*	0.33	-1.24*	0.30								
German - orthography												
Ib					-0.68	0.36	-0.62	0.34				
II					-1.16*	0.32	-1.03*	0.32				
III					-1.77*	0.36	-1.56*	0.36				
Above III					-2.85*	0.43	-2.55*	0.42				
English - listening												
A1.2									-1.31*	0.53	-1.31*	0.53
A2.1									-2.55*	0.55	-2.53*	0.55
A2.2									-4.12*	0.50	-4.10*	0.51
B1.1									-5.39*	0.52	-5.39*	0.53
B1.2									-6.56*	0.57	-6.61*	0.57
Above B1.2									-7.87*	0.83	-7.85*	0.87
Controls												
ISEI			-0.01*	0.00			-0.01	0.00			-0.00	0.00
Sex - female			-0.82*	0.19			-0.68*	0.20			-0.02	0.19
Migration background - yes			0.21	0.15			0.20	0.17			0.26	0.18
Multi-track school			-0.59*	0.25			-0.67*	0.31			0.07	0.30
Academic track school			-0.21	0.23			-0.01	0.23			0.44	0.26
Pseudo- R^2	0.03		0.05		0.06		0.08		0.23		0.23	
Δ Pseudo- R^2			0.03				0.02				0.00	
N	735		735		735		735		707		707	

Notes: reference categories: proficiency levels German = Ia, proficiency levels English = A1.1, sex = male, migration background = no, school type = comprehensive school; reported coefficients are on log-odds scale; results are based on $m = 15$ imputed datasets; McFadden Pseudo- R^2 is calculated as the average of all Pseudo- R^2 's across the 15 imputations assuming independence between the observations in the sample.
* $p < 0.05$ (two tailed tests).

Table S10. Results from ordered logistic regressions of grades on alternative domains - continuous scores

	Grade German				Grade German				Grade English			
	Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
	β	SE	β	SE	β	SE	β	SE	β	SE	β	SE
Proficiency score												
German - listening												
	-0.00*	0.00	-0.00*	0.00								
German - orthography												
					-0.01*	0.00	-0.01*	0.00				
English - listening												
									-0.02*	0.00	-0.02*	0.00
Controls												
ISEI			-0.01*	0.00			-0.01	0.01			-0.00	0.00
Sex - female			-0.81*	0.18			-0.65*	0.19			-0.02	0.19
Migration background - yes			0.19	0.15			0.20	0.17			0.27	0.18
Multi-track school			-0.60*	0.26			-0.68*	0.31			0.11	0.29
Academic track school			-0.20	0.23			0.06	0.23			0.54	0.28
Pseudo- R^2	0.02		0.05		0.06		0.08		0.24		0.25	
Δ Pseudo- R^2			0.03				0.02				0.00	
N	735		735		735		735		707		707	

Notes: reference categories: sex = male, migration background = no; school type = comprehensive school; reported coefficients are on log-odds scale; results are based on $m = 15$ imputed datasets; McFadden Pseudo- R^2 is calculated as the average of all Pseudo- R^2 's across the 15 imputations assuming independence between the observations in the sample.
* $p < 0.05$ (two tailed tests).

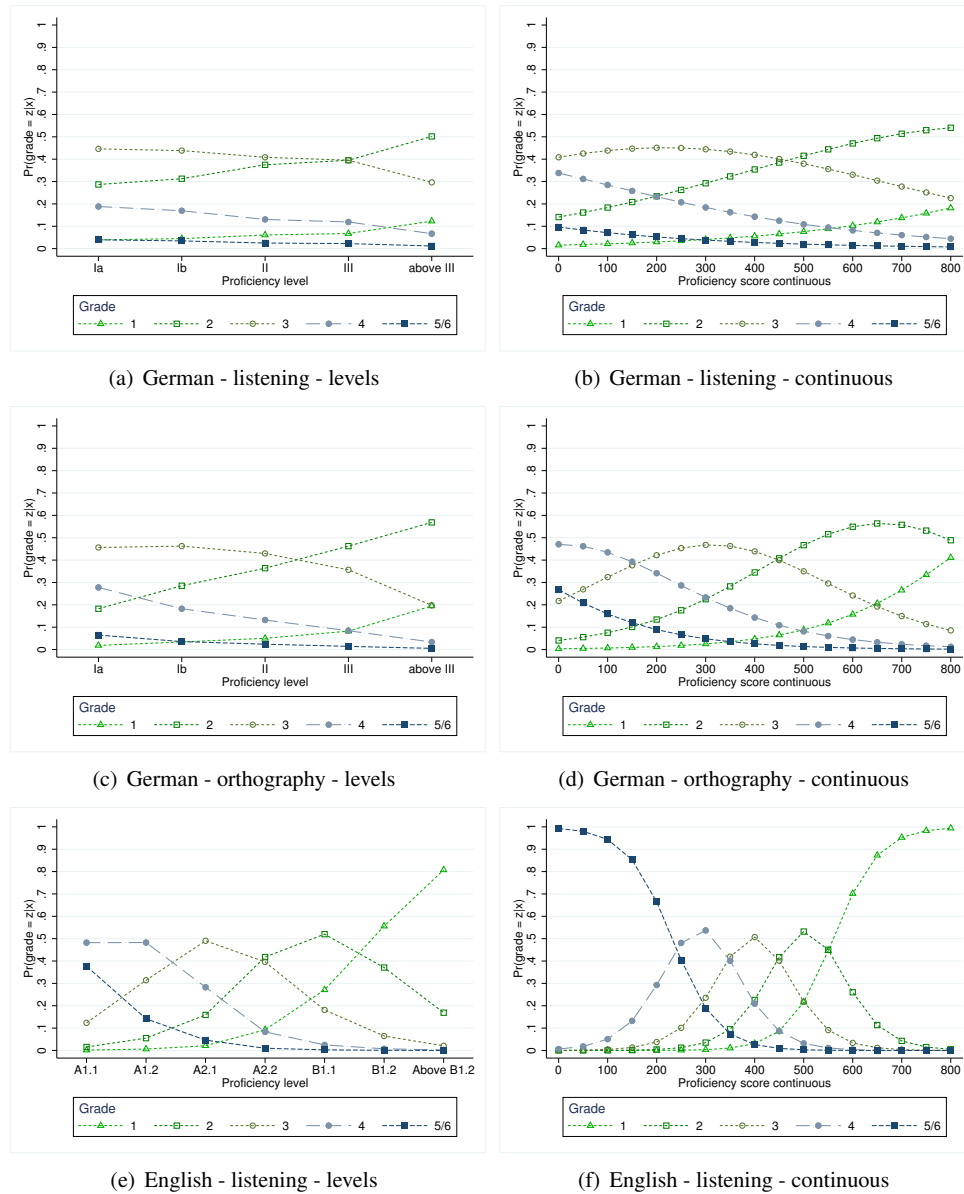


Figure S2. Predicted probabilities of achieving different grade in specific language conditional on the alternative domains (German listening, German orthography and English listening) from adjusted models, controlling for sex, migration background, household socio-economic status, and school track.