London School of Economics

ST314 Individual Project

Candidate Number: 24155

Word Count $\approx$ 6700

15 April 2024

Table of Contents

Programme for International Student Assessment

Introduction

This report investigates factors associated with reading performance for 15-year-old students in the UK using the Programme for International Student Assessment (PISA) dataset. The dataset has a two-level hierarchical structure with students (level 1) nested within schools (level 2). PISA is a triennial survey that tries to evaluate education systems by testing the skills and knowledge of 15-year-old students. The dataset has 7610 students across 334 schools in the UK. At the student level, we consider gender, immigration status, parental education, family wealth, cultural possessions at home, and minutes spent learning per week. At the school level, we have student-teacher ratio, school size, and school ownership type. By fitting multilevel models, we aim to address three key research questions: How much do schools vary in their average reading performance, that will tell us the extent to which students' reading ability depends on the school they attend. Which student and school characteristics predict reading scores in particular gender differences in reading performance. Do the effects of student-level variables differ between schools.

Variable	Description	Data Type	Levels & Frequencies
Level 1 - Individuals			
zread	Standardised reading score	Numerical	
schoolid	School identifier	Categorical	334 schools
female	Gender	Categorical	1: Female (3552), 0: Male (4058)
immig	Immigration background	Categorical	1: Native (6708), 2: 2nd gen (456), 3: 1st gen (446)
hisced	Highest education of parents	Numerical	0–6, higher scores = better educated
wealth	Family wealth score	Numerical	
cultposs	Cultural possessions at home score	Numerical	
lmins	Minutes learning per week	Numerical	
Level 2 - Schools			
stratio	Student-teacher ratio	Numerical	
schsize	School size	Numerical	
schltype	School ownership type	Categorical	1: Private independent (515), 2: Private government-dependent (1991) 3: Public (5104)

Table 1: Categorical & Numerical Variables

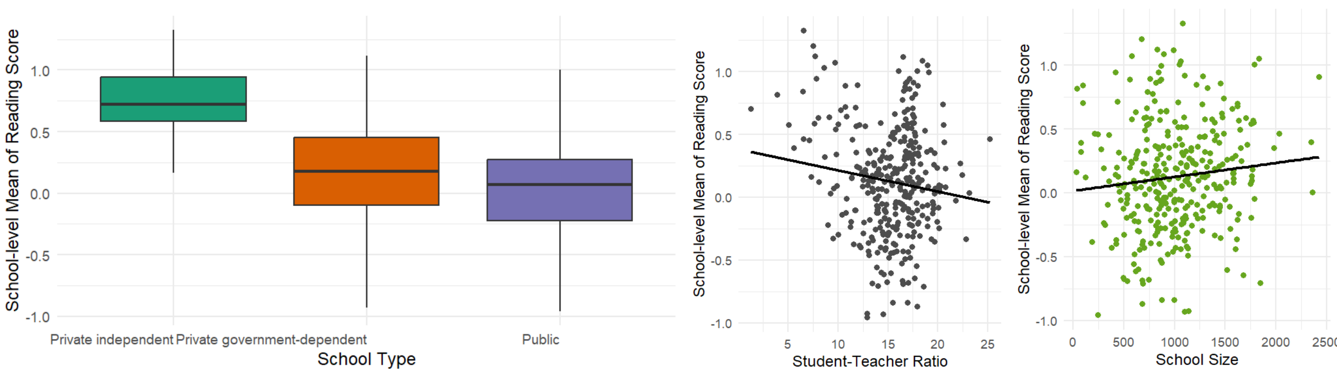
Descriptive Analysis

The response variable reading score (zread) has a mean of 0.177 and a standard deviation of 0.950, and its distribution is fairly symmetric but slightly left-skewed (Appendix 1.1). At the student level, 53.3% of students are female. Most students (88.1%) are native to the UK, while 6.0% are 2nd generation immigrants and 5.9% are 1st generation. Parental education (hisced) has a 7-point scale from 0-6, with a mean of 4.94 indicating relatively high average parental education. Family wealth and cultural possessions have means close to the standardised average of 0. Students spend 249 minutes per week on average learning, but this ranges widely from 0 to 1800 minutes.

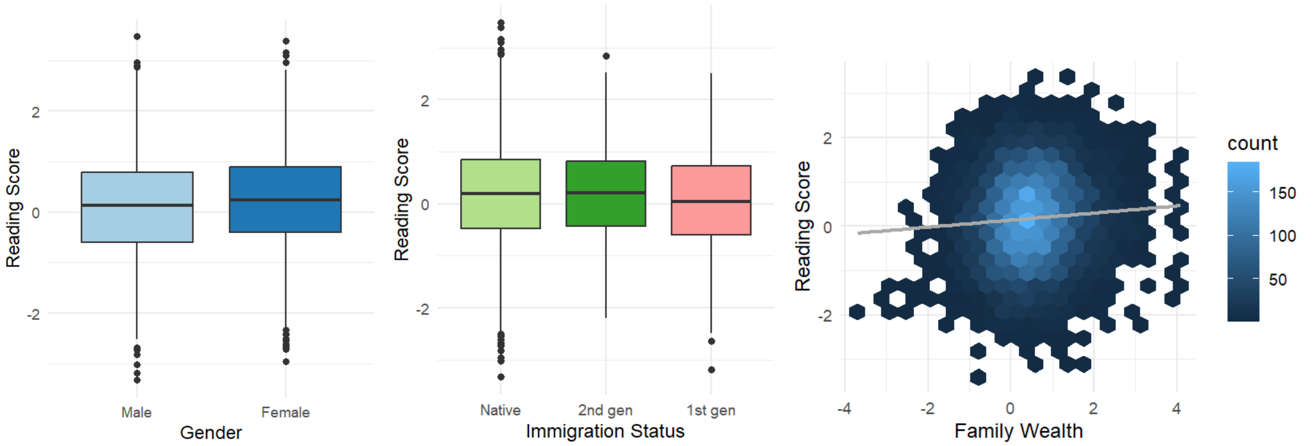
	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Individual-Level Variables						
zread	-3.3255	-0.4852	0.1902	0.1772	0.8419	3.4837
hisced	0.000	4.000	5.000	4.942	6.000	6.000
wealth	-3.7009	-0.0677	0.4616	0.5112	1.0346	4.0894
cultposs	-1.8461	-0.8559	-0.2588	-0.1696	0.4317	2.0538
lmins	0.0	200.0	240.0	249.1	280.0	1800.0
School-Level Variables						
stratio	1.25	13.74	15.94	15.44	17.49	25.19
schsize	33	735	998	1023	1285	2421

Table 2: Summary Statistics for Numerical Predictors

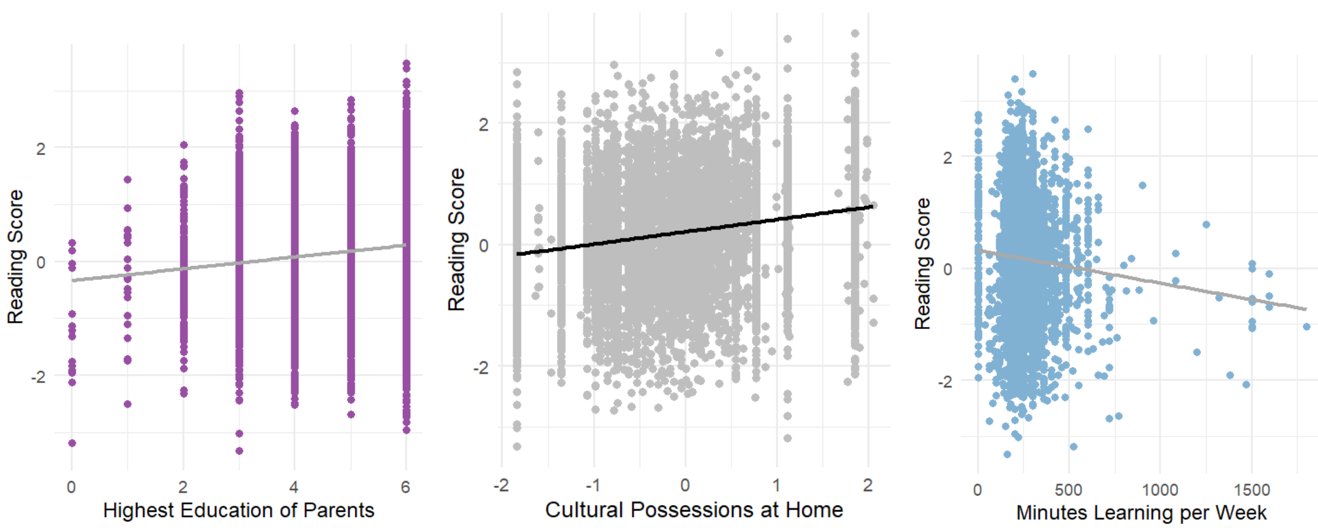
Public schools have the lowest average reading scores, while private government-dependent performs slightly better and independent schools performs significantly better. Reading scores decline slightly with increasing student-teacher ratio and school size.



Regarding relationships between student-level predictors and reading scores. Females score 0.31 standard deviations higher than males on average. Native students outperform 2nd generation immigrants by, who in turn outperform 1st generation immigrants.



There are positive associations between reading scores and parental education, family wealth, and cultural possessions. However, learning minutes has no consistent relationship with reading performance due to a few outliers' oversizing influence with the linear fitted line in the plot.



The variance components model $y_{ij} = \beta_0 + u_j + e_{ij}$ estimates that the overall mean reading score across all students and schools is 0.136. This represents the expected reading score for a student with 0 values on all predictors, which corresponds to a native male student in an independent-private school. A hypothesis test with a linear regression model showed a significant result for between-school variance. The model estimates the between-school variance in reading scores as $\sigma_u^2 = 0.1476$ and the within-school (between-student) variance as $\sigma_e^2 = 0.7489$. ICC represents the proportion of total variance in reading scores due to differences between schools:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} = \frac{0.1476}{0.1476 + 0.7489} = 0.165$$

16.5% of the variation in reading scores lies between schools, while the remaining 83.5% is within schools between students. The 95% confidence intervals for the variance components shows uncertainty in these estimates. Between-school standard deviation σ_u: (0.3493, 0.4215), within-school standard deviation σ_e: (0.8515, 0.8796). These intervals are quite narrow.

Table 3: Variance Component Model

Formula

zread ~ 1 + (1 | schoolid)

Intraclass Correlation (ICC)

Adjusted	Unadjusted
0.165	0.165

Random Effects - Number of obs: 7610, groups: school, 334

	Variance	Std. Dev.
Between-School	0.1476	0.3842
Within-School	0.7489	0.8654

Fixed Effects

	Estimate	Std. Error	t value	p-value
(Intercept)	0.136	0.0235	5.787	<0.0001

Confidence Intervals

	2.50%	97.50%
Intercept	0.0898	0.182
.sig01	0.3493	0.422
.sigma	0.8515	0.88

Model Selection Strategy

Student-Level Predictors

We started the model selection process by adding student-level predictors to the model to investigate individual differences in reading performance. The potential candidates were cultural possessions (cultposs), gender (female), highest parental education level (hisced), immigration status (immig), learning minutes (lmins), and family wealth (wealth). To select the best subset of predictors, we used the dredge function from the MuMIn package, which fits all possible combinations of the specified variables and ranks the models based on their AIC and BIC values. The best model included all six student-level predictors:

$$zread_{ij} = \beta_0 + \beta_1 cultposs_{ij} + \beta_2 female_{ij} + \beta_3 hisced_{ij} + \beta_4 immig_{ij} + \beta_5 lmins_{ij} + \beta_6 wealth_{ij} + u_{0j} + e_{ij}$$

Table 4: Random Intercept Model without Interactions

Formula

zread ~ cultposs + female + hisced + immig + lmins + wealth + (1 | schoolid)

Intraclass Correlation (ICC)

Adjusted	Unadjusted
0.148	0.142

Random Effects - Number of obs: 7610, groups: school, 334

	Variance	Std. Dev.
Between-School	0.1252	0.3538
Within-School	0.7225	0.85

Fixed Effects

	Estimate	Std. Error	t value	p-value	2.5% CI	97.5% CI
(Intercept)	0.0762	0.0249	3.0595	0.0022	0.0272	0.125
Cultural Possessions	0.1472	0.0109	13.5382	<0.0001	0.1259	0.1686
Female	0.1451	0.021	6.902	<0.0001	0.1039	0.1863
Family Education	0.0495	0.0107	4.6237	<0.0001	0.0285	0.0706
2nd Gen Immigrant	-0.0699	0.0457	-1.53	0.126	-0.1594	0.0197
1st Gen Immigrant	-0.1556	0.0447	-3.4774	0.0005	-0.2433	-0.0679
learning Mins	-0.0247	0.0105	-2.3566	0.0184	-0.0453	-0.0041
Wealth	-0.0572	0.0111	-5.1607	<0.0001	-0.079	-0.0353

The inclusion of student-level predictors in the model led to a reduction in the ICC from 0.165 in the variance components model to 0.142 in the model with student-level predictors. This decrease shows that a portion of the variance in reading scores previously attributed to differences between schools can be explained by the composition of the student body within each school in terms of the included predictors. Some of the school effects on reading scores are due to the clustering of students with these six characteristics within schools.

Looking at the confidence intervals for the fixed effects coefficients in the model with student-level predictors, we observe that the intervals for all predictors exclude zero. The confidence intervals for the random effects parameters, specifically the residual variance, are narrower compared to those in the variance components model, suggesting that including student-level predictors has reduced the unexplained variability in reading scores.

There is a moderate negative association (-0.444) between the intercept and the coefficient for female, suggesting that schools with higher average reading scores tend to have a smaller gender gap in favour of girls. We also observe a moderate negative correlation between cultural possessions and the highest parental education level (-0.203), implying that the effect of cultural possessions on reading scores may be somewhat less pronounced for students with highly educated parents.

Interactions Among Student-Level Predictors

We then proceeded to test for potential interactions among level 1 predictors. To test for interactions, we used a nested loop to fit models with all possible two-way interactions among the student-level predictors and compared their AIC and BIC values to those of the model without interactions. The model comparison revealed that the interaction between cultural possessions and wealth (cultposs × wealth) and the interaction between cultural possessions and highest parental education level (cultposs × hisced) significantly improved the model fit, with p-values < 0.001 by likelihood ratio test. Unfortunately, **none** of the interactions with gender showed any significance compared with the baseline model without any interactions. The new model with these interaction terms is:

$$\begin{aligned} zread_{ij} = {} & \beta_0 + \beta_1 cultposs_{ij} + \beta_2 female_{ij} + \beta_3 hisced_{ij} + \beta_4 immig_{ij} + \beta_5 lmins_{ij} + \beta_6 wealth_{ij} \\ & + \beta_7 (cultposs_{ij} \times wealth_{ij}) + \beta_8 (cultposs_{ij} \times hisced_{ij}) + u_{0j} + e_{ij} \end{aligned}$$

The inclusion of these interaction terms led to a slight decrease in the ICC from 0.142 in the previous model to 0.148. The interactions explain a small portion of the between-school variation in reading scores. The effect of cultural possessions on reading performance varies across schools depending on the average levels of family wealth and parental education within each school.

Table 5: Random Intercept Model with Interactions

Formula

zread ~ cultposs + female + hisced + immig + lmins + wealth + (1 | schoolid) + hisced:cultposs + wealth:cultposs

Intraclass Correlation (ICC)

Adjusted	Unadjusted
0.148	0.142

Random Effects - Number of obs: 7610, groups: school, 334

	Variance	Std. Dev.
Between-School	0.1248	0.3533
Within-School	0.718	0.8474

Fixed Effects

	Estimate	Std. Error	t value	p-value	2.5% CI	97.5% CI
(Intercept)	0.0782	0.0251	3.1176	0.0018	0.0288	0.1274
Cultural Possessions	0.1437	0.011	13.0713	<0.0001	0.1221	0.1653
Female	0.1421	0.021	6.7765	<0.0001	0.101	0.1832
Family Education	0.0565	0.011	5.1355	<0.0001	0.0349	0.0781
Wealth	-0.0463	0.0112	-4.1408	0	-0.0683	-0.0243
2nd Gen Immigrant	-0.0636	0.0455	-1.3971	0.1624	-0.1529	0.0257
1st Gen Immigrant	-0.1535	0.0446	-3.441	0.0006	-0.2409	-0.0661
Learning Mins	-0.0234	0.0104	-2.2437	0.0248	-0.044	-0.0029
cultposs:hisced	0.0528	0.0112	4.7275	<0.0001	0.0309	0.0747
cultposs:wealth	-0.0572	0.0097	-5.8753	<0.0001	-0.0762	-0.0381

Looking at the confidence intervals for the fixed effects coefficients in the updated model, the intervals for the interaction terms cultposs × wealth (-0.0788, -0.0357) and cultposs × hisced (0.0309, 0.0747) do not include zero. The main effect of learning minutes becomes non-significant after the inclusion of the interaction terms.

A moderate negative correlation (-0.189) between cultural possessions and wealth exists, suggesting that the effect of cultural possessions on reading scores tends to be weaker in schools with higher average family wealth. Similarly, there is a weak positive correlation (0.221) between the cultposs × hisced interaction and the main effect of the highest parental education level, indicating that the effect of cultural possessions on reading scores may be more pronounced for students with highly educated parents. The positive effect of cultural possessions on reading scores might larger for students from less affluent families or those with lower parental education levels, as these cultural resources may help compensate for other potential disadvantages.

School-Level Predictors

We considered three school-level predictors: school type (schltype), student-teacher ratio (stratio), and school size (schsize). To determine which of these predictors significantly improved the model fit, we added them to the model one at a time and used likelihood ratio tests to compare the models with and without each predictor. The results showed that adding school type (schltype) significantly improved the model fit (chi-square = 62.296, df = 2, p-value < 0.001), indicating that reading scores vary across different types of schools (private independent, private government-dependent, and public schools). Furthermore, including school size (schsize) in the model also led to a significant improvement in fit (chi-square = 11.399, df = 1, p-value < 0.001), suggesting that school size plays a role in explaining variation in reading scores. We firstly selected school type, than we tried the other two variables for anova test. However, the addition of student-teacher ratio (stratio) did not significantly improve the model fit and was therefore not included in the subsequent steps. The updated model included school type and school size:

$$zread_{ij} = \beta_0 + \beta_1 cultposs_{ij} + \beta_2 female_{ij} + \beta_3 hisced_{ij} + \beta_4 immig_{ij} + \beta_5 lmins_{ij} + \beta_6 wealth_{ij} + \beta_7\left(cultposs_{ij} \times wealth_{ij}\right) + \beta_8\left(cultposs_{ij} \times hisced_{ij}\right) + \beta_9 schsize_j + \beta_{10} schltype_j + u_{0j} + e_{ij}$$

Table 6: Random Intercept Model with Interactions and School Level Variables

Formula

zread ~ cultposs + female + hisced + immig + lmins + wealth + (1 | schoolid) + hisced:cultposs + wealth:cultposs + schltype + schsize

Intraclass Correlation (ICC)

	Adjusted	Unadjusted
	0.116	0.106

Random Effects - Number of obs: 7610, groups: school, 334

	Variance	Std. Dev.
Between-School	0.09432	0.3071
Within-School	0.71776	0.8472

Fixed Effects

	Estimate	p-value	2.5% CI	97.5% CI
(Intercept)	0.7522	<0.0001	0.5908	0.9136
Cultural Possessions at Home	0.14	<0.0001	0.1184	0.1616
Gender: Female	0.1437	<0.0001	0.1027	0.1847
Highest Education of Parents	0.0536	<0.0001	0.032	0.0752
Family Wealth	-0.0519	<0.0001	-0.0739	-0.0299
Immigration Background: 2nd Generation	-0.0901	0.0473	-0.1791	-0.0011
Immigration Background: 1st Generation	-0.1738	0.0001	-0.261	-0.0866
Learning Time for English (minutes/week)	-0.0228	0.0283	-0.0432	-0.0024
School Size	0.0678	0.0007	0.0288	0.1068
School Type: Private Government-Dependent	-0.6479	<0.0001	-0.8265	-0.4693
School Type: Public	-0.7357	<0.0001	-0.9015	-0.5699
Cultural Possessions at Home × Highest Education of Parents	0.0512	<0.0001	0.0294	0.073
Cultural Possessions at Home × Family Wealth	-0.0592	<0.0001	-0.0782	-0.0402

ICC changed to 0.106. This decrease shows that a substantial portion of the between-school variation in reading scores can be attributed to differences in school type and school size. The intervals for school size (0.0287, 0.1065) and both categories of school type (private government-dependent: -0.8272, -0.4690; public: -0.9022, -0.5696) do not include zero, confirming their statistical significance at the 5% level. The effects of the student-level predictors and their interactions remain largely unchanged after the addition of the school-level variables.

The correlation matrix showed a negative correlation (-0.486) between the intercept and the coefficient for private government-dependent schools, suggesting that schools with higher average reading scores are less likely to be private government-dependent schools. Similarly, there is a moderate negative correlation (-0.483) between the intercept and the coefficient for public schools, indicating that schools with higher average reading scores are also less likely to be public schools, as expected in the data exploration phase of the analysis. Surprisingly, larger schools tend to have higher average reading performance, even after accounting for student-level factors and school type.

Regarding school type, the model results indicate that students in private government-dependent and public schools tend to have lower reading scores compared to those in private independent schools, after controlling for student-level characteristics and school size.

Random Slopes

We then tested random slopes for each student-level predictor one at a time. We compared models with and without each random slope using hand-coded likelihood ratio tests to determine whether the inclusion of the random slope significantly improved the model fit, because the default LRT in anova function overestimates P values in this scenario.

Table 7: Random Slope Selection

Random Slope	p-value
Cultural Possessions at Home	0.18009
Gender: Female	0.04682
Highest Education of Parents	0.05647
Family Wealth	0.30423
Immigration Background	0.00233
Learning Time for English (minutes/week)	0.04015

The results showed significant random slopes for gender (female; p-value = 0.047), immigration status (immig; p-value = 0.002), and learning minutes (lmins; p-value = 0.040). Given the significance of the random slope for gender and its relevance to the research question, we decided to include it in the model first. The updated model, incorporating the random slope for gender, can be expressed as:

$$\begin{aligned} zread_{ij} = \beta_0 &+ \beta_1 cultposs_{ij} + \beta_2 female_{ij} + \beta_3 hisced_{ij} + \beta_4 immig_{ij} + \beta_5 lmins_{ij} + \beta_6 wealth_{ij} \\ &+ \beta_7 (cultposs_{ij} \times wealth_{ij}) + \beta_8 (cultposs_{ij} \times hisced_{ij}) + \beta_9 schsize_j \\ &+ \beta_{10} schltype_j + u_{0j} + u_{1j} female_{ij} + e_{ij} \end{aligned}$$

Next, we added the random slope for learning minutes (lmins) to the model, as it was more significant than immig. The updated model incorporating both random slopes is:

$$zread_{ij} = \beta_0 + \beta_1 cultposs_{ij} + \beta_2 female_{ij} + \beta_3 hisced_{ij} + \beta_4 immig_{ij} + \beta_5 lmins_{ij} + \beta_6 wealth_{ij} + \beta_7 (cultposs_{ij} \times wealth_{ij}) + \beta_8 (cultposs_{ij} \times hisced_{ij}) + \beta_9 schsize_j + \beta_{10} schltype_j + u_{0j} + u_{1j} female_{ij} + u_{2j} lmins_{ij} + e_{ij}$$

ICC increased slightly from 0.122 in the model with only school-level predictors to 0.130. This increase suggests that allowing the effects of gender and learning minutes to vary across schools captures additional between-school variation in reading scores. Also, the variance of the random slope for gender (σ^2_{u1} = 0.016) and the variance of the random slope for learning minutes (σ^2_{u2}= 0.009) are relatively small compared to the variance of the random intercept (σ^2_{u0} = 0.113). The correlation between the random intercept and the random slope for gender (ρ_{u01} = -0.55) is negative and moderate, suggesting that schools with higher average reading scores tend to have a smaller gender gap in favor of girls. The correlation between the random intercept and the random slope for learning minutes (ρ_{u02}= 0.20) is positive but weak, indicating a slight tendency for schools with higher average reading scores to have a stronger positive effect of learning minutes on reading performance.

Table 8: Random Slope Model with Interactions and School Level Variables

Formula

zread ~ cultposs + female + hisced + immig + lmins + wealth + (1 + female + lmins | schoolid) + hisced:cultposs + wealth:cultposs + schltype + schsize

Intraclass Correlation (ICC)

Adjusted	Unadjusted
0.13	0.119

Random Effects - Number of obs: 7610, groups: school, 334

		SD	Correlation with Intercept	Correlation between Slopes
Between-School	lmins	0.09432	0.198	
	femaleFemale	0.71776	-0.549	-0.675
	Intercept	0.336336		
Within-School	Residuals	0.841395		

The correlation between the random slopes for gender and learning minutes (ρ_{u12} = -0.68) is negative and relatively strong, suggesting that schools with a larger gender gap in favor of girls tend to have a weaker effect of learning minutes on reading scores.

Cross-Level Interactions

We then proceeded to test cross-level interactions between the student-level predictors (cultposs, female, hisced, wealth, immig, and lmins) and the school-level predictors (schsize and schltype). For example, a significant interaction between gender and school type would suggest that the gender gap in reading performance differs across different types of schools.

To test for cross-level interactions, we added each interaction term to the model one at a time and compared the resulting models using likelihood ratio tests and information criteria (AIC and BIC). We

Table 9: Selection of Cross-Level Interactions

Interaction	AIC	BIC	Likelihood Ratio Test P Value
Base Model	19551	19690	-
lmins:schltype	19546	19699	0.0104
immig:schltype	19547	19713	0.0134
immig:schsize	19547	19699	0.0151
hisced:schltype	19549	19702	0.0444
wealth:schsize	19550	19695	0.0608
female:schsize	19552	19698	0.2332
wealth:schltype	19553	19705	0.268
cultposs:schltype	19553	19705	0.3042
female:schltype	19553	19706	0.3127
hisced:schsize	19552	19698	0.3466
lmins:schsize	19552	19698	0.3591
cultposs:schsize	19553	19698	0.4868

considered 12 interactions. We compared their AIC, BIC, and likelihood ratio test p-values to determine which interactions significantly improved the model fit. The results showed that three cross-level interactions were significant at the 5% level:

lmins × schltype (p-value = 0.010)

immig × schltype (p-value = 0.013)

immig × schsize (p-value = 0.015)

We added these three significant cross-level interactions to the model one at a time. The model with the lmins × schltype interaction had the lowest AIC and BIC values and was therefore selected as the best-fitting model.

Interpretation of Final Model

Our final random slope model has student-level predictors (cultural possessions, gender, highest parental education, wealth, immigration status, and learning minutes), school-level predictors (school size and school type), interactions among student-level predictors (cultural possessions × highest parental education and cultural possessions × wealth), a cross-level interaction (learning minutes × school type), and random slopes for learning minutes and gender across schools. The formula of the final model is:

$$\begin{aligned} zread_{ij} = \beta_0 &+ \beta_1 cultposs_{ij} + \beta_2 female_{ij} + \beta_3 hisced_{ij} + \beta_4 immig_{ij} + \beta_5 lmins_{ij} + \beta_6 wealth_{ij} \\ &+ \beta_7\left(cultposs_{ij} \times wealth_{ij}\right) + \beta_8\left(cultposs_{ij} \times hisced_{ij}\right) + \beta_9 schsize_j \\ &+ \beta_{10} schltype_j + \beta_{11}\left(lmins_{ij} \times schltype_j\right) + u_{0j} + u_{1j} female_{ij} + u_{2j} lmins_{ij} + u_{0j} \\ &+ e_{ij} \end{aligned}$$

$zread_{ij}$ is the standardized reading score for student i in school j, β_0 is the overall intercept, β_1 to β_{11} are the fixed effects coefficients, u_{0j} is the random intercept for school j, u_{1j} is the random slope for learning minutes in school j, u_{2j} is the random slope for gender in school j, e_{ij} is the student-level residual. The random part of the model has school-level random effects and the student-level residual variance.

The variance estimates are: $\sigma_{u0}^2(sig2u0) = 0.112$ is the variance of the school-level random intercepts, $\sigma_{u1}^2(sig2u1) = 0.008$ is the variance of the school-level random slopes for learning minutes, $\sigma_{u2}^2(sig2u2)$ = 0.016 is the variance of the school-level random slopes for gender $\sigma_e^2(sigma)$ = 0.708 is the student-level residual variance. The random intercept variance (σ_{u0}^2) shows the variability in the average reading scores across schools, after accounting for the fixed effects. The random slope variances (σ_{u1}^2 *and* σ_{u2}^2) indicate the extent to which the effects of learning minutes and gender on reading scores vary across schools.

We also found out that the correlations among the random effects are: $\rho_{u01} = 0.22$ is the correlation between the random intercepts and the random slopes for learning minutes. ρ_{u02} = -0.54 is the correlation between the random intercepts and the random slopes for gender. ρ_{u12} = -0.69 is the correlation between the random slopes for learning minutes and gender, which is relatively large.

In addition, positive correlation between the random intercepts and the random slopes for learning minutes (ρ_{u01}) shows a slight "fanning out" pattern, meaning that schools with higher average reading scores tend to have a stronger positive effect of learning minutes on reading performance. The negative correlation between the random intercepts and the random slopes for gender (ρ_{u02}) indicates a "fanning in" pattern, where schools with higher average reading scores tend to have a smaller gender gap in favor of girls. The strong negative correlation between the random slopes for learning minutes and gender (ρ_{u12}) implies that schools with a stronger positive effect of learning minutes tend to have a smaller gender gap in reading scores.

The fixed part of the model estimates the average effects of the predictors on reading scores, holding the other predictors constant.

β_1 = 0.141 ($p < 0.001$): one SD increase in cultural possessions is associated with a 0.141 SD increase in reading scores, on average, because all numerical predictors have already been standardised.

β_2 = 0.145 ($p < 0.001$): female students (coded as 1) are expected to score 0.145 SD higher than male students (coded as 0).

β_3 = 0.053 ($p < 0.001$): a one SD increase in the highest parental education level is associated with a 0.053 SD increase in reading scores on average.

β_4 = -0.052 ($p < 0.001$): a SD increase in wealth is associated with a 0.052 SD decrease in reading scores on average.

β_5 (immig2nd gen) = -0.089 ($p = 0.048$): second-generation immigrant students are expected to score 0.089 SD lower than native students (the baseline category) on average.

β_5 (immig1st gen) = -0.167 ($p < 0.001$): first-generation immigrant students are expected to score 0.167 SD lower than native students on average.

β_6 = -0.135 ($p = 0.009$): one SD increase in learning minutes is associated with a 0.135 SD decrease in reading scores for students in private independent schools (the baseline category) on average.

β_9 = 0.067 ($p < 0.001$): one SD increase in school size is associated with a 0.067 SD increase in reading scores on average.

β_{10}(Private government-dependent) = -0.621 ($p < 0.001$): students in private government-dependent schools are expected to score 0.621 SD lower than those in private independent schools.

β_{10} (Public) = -0.705 ($p < 0.001$): students in public schools are expected to score 0.705 SD lower than those in private independent schools.

$\beta_7 = 0.050$ ($p < 0.001$): the positive interaction between cultural possessions and the highest parental education level suggests that the effect of cultural possessions on reading scores is stronger for students with higher parental education.

β_8 = -0.059 ($p < 0.001$): the negative interaction between cultural possessions and wealth indicates that the effect of cultural possessions on reading scores is weaker for students from wealthier families.

β_{11}(Private government-dependent) = 0.097 ($p = 0.092$): the effect of learning minutes on reading scores is 0.097 SD higher for students in private government-dependent schools compared to those in private independent schools, although this difference is not statistically significant at the 5% level, we decided to avoid further interpreting the result.

β_{11}(Public) = 0.148 ($p = 0.006$): the effect of learning minutes on reading scores is 0.148 SD higher for students in public schools compared to those in private independent schools.

The confidence intervals for the random effects and residual variance are from the output of confint function. .sig01 is the standard deviation of the random intercepts $\left(\sqrt{\sigma_{u0}^2}\right)$ .sig02 is the standard deviation of the random slopes for learning minutes $\left(\sqrt{\sigma_{u1}^2}\right)$ .sig03 is the standard deviation of the random slopes for gender $\left(\sqrt{\sigma_{u2}^2}\right)$ (Appendix 2.2)

.sig04: the correlation between the random intercepts and the random slopes for learning minutes (ρ_{u01})

.sig05: the correlation between the random intercepts and the random slopes for gender (ρ_{u02})

.sig06: the correlation between the random slopes for learning minutes and gender (ρ_{u12})

.sigma: the standard deviation of the student-level residuals $\left(\sqrt{\sigma_e^2}\right)$

The 95% confidence intervals for these parameters shows the range within which the true parameter values are expected to lie with 95% probability. For example, the 95% confidence interval for .sig01 is [0.294, 0.378], meaning that we are 95% confident that the true value of $\sqrt{\sigma_{u0}^2}$ lies between 0.294 and 0.378.

The final model provides much information on the factors influencing students' reading performance and their magnitudes, it has both student-level and school-level predictors, also their interactions and random effects. The magnitudes of its coefficients show the importance of cultural possessions, parental education, and school characteristics in explaining differences in reading scores, while also showing significant variation in the effects of learning minutes and gender across schools. The presence of "fanning out" and "fanning in" patterns in the random effects shows that the influence of learning minutes and gender on reading scores varies depending on the overall performance level of the school.

Limitations

Causal inferences about the relationships between the predictors and reading achievement is limited. For example, while we observe a significant association between cultural possessions and reading scores, we cannot conclude that increasing cultural possessions will lead to improved reading performance. Confounder may play a role here, such as parental involvement or students' motivation, that influence both cultural possessions and reading achievement.

The PISA data also relies on self-reported information from students and school administrators, which may be subject to biases or inconsistencies. For instance, students may not accurately report their number of learning minutes they receive in school.

The model assumptions of linearity, normality, and homoscedasticity may also be violated. There may be non-linear relationships or threshold effects that are not fitted by the current model specification. The effect of learning minutes on reading scores may not be constant across the entire range of values, and

2^{nd} gen immigrants showed no significant difference with native student, merging the two levels may be a justified decision. Diminishing returns at high levels of learning time may also be investigated.

The normality assumption requires that the residuals at both the student and school levels follow a normal distribution. But students from disadvantaged backgrounds elite schools may have residuals that deviate substantially from a normal distribution.

Given the focus on gender differences in the research questions, it is important to consider the limitations of the binary gender variable used in the analysis. The PISA data only allow for a binary classification of students as male or female, which may not fully include the modern complexity of gender identity and its relationship to reading performance. Moreover, the interpretation of the gender effect as a gap in favour of girls may oversimplify the nature of gender differences in reading achievement, as it assumes a uniform advantage for girls across all schools and subgroups of students.

Furthermore, the model assumes that the random effects, particularly the random slopes for learning minutes and gender, follow a multivariate normal distribution with a constant covariance structure across schools. However, the correlation between the random slopes for learning minutes and gender may be stronger in schools with a particular instructional approach or curriculum emphasis.

Non-Technical Summary

Have you ever wondered what makes some students excel in reading while others struggle? As it turns out, the answer is not as simple as one might think. The study aimed to uncover the extent to which schools impact students' reading ability, identify the key predictors of reading success at both the individual and school levels, and explore whether the effects of these predictors vary across different schools.

To address these questions, we employed a multilevel modelling approach, which allows for the simultaneous examination of both student-level and school-level factors while accounting for the hierarchical structure of the data. This approach is an extension of linear regression that recognizes the potential dependence among observations within the same school and enables the estimation of both the average relationships between predictors and the outcome and the variation in these relationships across schools.

The analysis revealed that schools play a significant role in shaping students' reading performance, with approximately 12-16% of the total variation in reading scores attributable to differences between schools.

At the student level, several predictors were found to be significantly associated with reading performance. Cultural possessions, parental education, and being female were positively related to reading scores, while wealth and immigrant status (both first and second generation) were negatively associated with reading achievement.

The analysis also showed significant gender differences in reading performance, with female students outperforming their male counterparts.

Students in larger schools tended to have higher reading scores, while those in private government-dependent and public schools had lower scores compared to students in private independent schools.

Importantly, the analysis also revealed that the effects of some individual-level predictors, specifically learning minutes and gender, varied significantly across schools. This means that the impact of these factors on reading performance was not uniform across all schools, with some schools showing stronger relationships between these predictors and reading scores.

In addition, the positive effect of cultural possessions on reading scores was found to be stronger for students with higher parental education but weaker for students from wealthier families. Additionally, the impact of learning minutes on reading performance was found to differ across school types, with public schools showing a stronger positive relationship compared to private independent schools.

This study reminds us that every student is unique, and that there is no one-size-fits-all solution to improving reading performance. So the next time you see a student struggling with reading, remember that there may be a complex web of factors at play - from their family background to their school environment to their individual characteristics. And by unravelling these mysteries and provide the right support and resources, we can help every student become a confident and successful reader.

UK Household Longitudinal Study

Introduction

Physical functioning is a important part of health in older adults. The Understanding Society: UK Household Longitudinal Study (UKHLS) provides a dataset to investigate physical functioning over time and the factors that influence it. The dataset has 28,468 observations from 5,602 individuals. This statistical report aims to address three key research questions using the dataset: (1) What is the nature of change in physical functioning with age? (2) To what extent do individuals vary in their physical functioning trajectories? (3) What are the predictors of physical functioning? The survey conducted through 11 waves that collects data on various aspects of life, including health, employment, education, ethnicity, housing and social relationships, helps to find the relationship to physical functioning. The survey was conducted among all adult members in households annually. For this report, we focus on a subsample of individuals aged 60 and older who provided self-reported measures of physical functioning using the SF-12 questionnaire. The response variable assesses both physical and mental health status, with higher scores indicating worse health. By examining the SF-12 physical functioning scores across multiple waves of the UKHLS, we can explore the trajectories of physical functioning in later life and identify the key determinants of these trajectories.

Descriptive Analysis

The dataset has a two-level structure, level one is for each wave of study, level two is individuals. The survey has an average of 5.08 observations per person. The log standardised response variable, zpcs, represents the standardised SF-12 physical functioning score, which has a mean of -0.0015 and a standard deviation of 1.028. The age of participants ranges from 60 to 90 years, with a mean of 70.15 years. The mean age of participants for each wave is close to 70, indicating age is a better measure of time than wave because physical functioning is closely related to age.

The dataset has two time-invariant variables gender (female), and ethnicity (Table 2.1), four time-variant variables highest level of educational qualifications, whether respondent has coresident partner, housing tenure, and log income. Individuals in the dataset has a slightly higher proportion of female (53%) than male and the sample is predominantly White (92%), with smaller proportions of Asian (4%), Black (3%), and other ethnicities (1%). Approximately 30% of the sample has a degree or higher education, while the remaining 70% have lower educational attainment. Most samples are homeowners (82%), with 14% in social rental and 4% in private rental housing. The average log-income is 9.78.

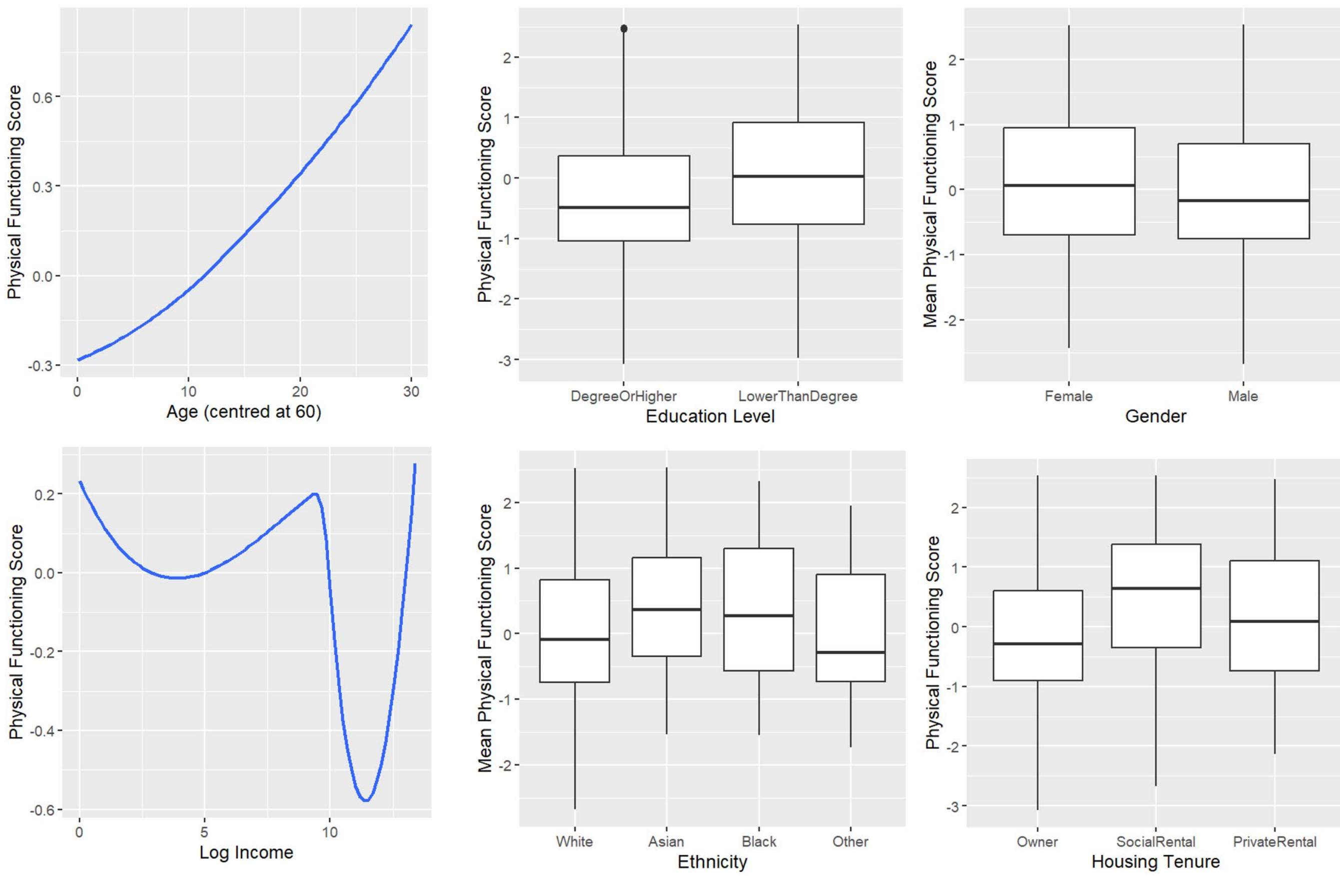
Table 2.1: Categorical & Numerical Variables

Variable	Description	Data Type	Levels
Level 1 - Observations			
ind	Individual identifier	Categorical	
wave	Wave of measurement (TV)	Categorical	1 to 11
zpcs	Standardised physical functioning score (TV)	Numerical	Higher scores: poorer functioning
age	Age in years (TV)	Numerical	
highed	Highest level of educational qualifications (TV)	Categorical	1=degree or other higher degree, 0=lower than degree
partner	Respondent has coresident partner (TV)	Categorical	1: Yes, 0: No
tenurehh	Housing tenure (TV)	Categorical	1=owns outright or with mortgage, 2=social rental, 3= private rental
income	Log of equivalised household income (TV)	Numerical	
Level 2 - Individuals			
female	Gender	Categorical	1: Female, 0: Male
ethnicity	Ethnic background	Categorical	1: White, 2: Asian, 3: Black, 4: Other

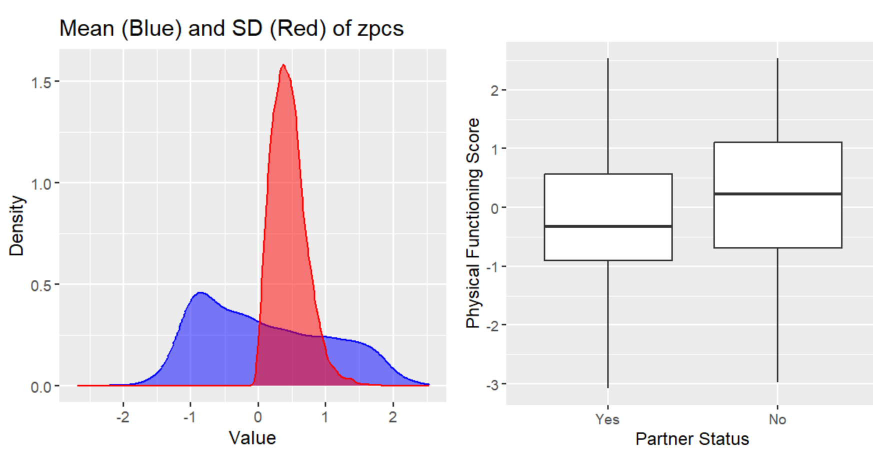
To assess the extent of between- and within-individual variation in physical functioning scores, we use the intraclass correlation coefficient (ICC) for our variance-component baseline model, which is the proportion of total variation that is attributable to differences between individuals, calculated as $\sigma_{u0}^2/(\sigma_{u0}^2 + \sigma_e^2)$. For random effects, variance for intercept is 0.78 and residual variance is 0.28. The ICC is 0.732, showing that nearly three-quarters of the variation in physical functioning scores is between individuals, while the remaining 27% is within individuals over time. This high ICC suggests that individuals generally maintain their relative position in the distribution of physical functioning scores.

$$y_{ij} = \beta_0 + u_{0j} + e_{ij}, \quad u_{0j} \sim N(0, \sigma_{u0}^2), \quad e_{ij} \sim N(0, \sigma_e^2)$$

The plot of physical functioning scores against age suggests a gradual decline with increasing age, particularly after age 70. The trajectory appears to be nonlinear, with the rate of decline accelerating at older ages. When examining the relationship between physical functioning and education level, individuals with a degree or higher tend to have better physical functioning scores than those with lower educational attainment. Similarly, Homeowners exhibiting higher scores on average compared to those in social or private rental housing. The plot of physical functioning scores against log income reveals no clear pattern, with high income levels corresponding to better physical functioning but very high-income earners' health may be distorted by a few outliers. Males tend to have slightly higher scores than females, suggesting better physical functioning on average. Among ethnic groups, the White and Other categories have higher median scores than the Asian and Black ethnicities.



Individuals who have a partner generally have better scores than those without a partner. The density plot of the mean and standard deviation of physical functioning scores across individuals provides insights into the variation in these measures. The distribution of individual means is centered around zero, with a slight negative skew. The distribution of individual standard deviations is positively skewed, with a peak near 0.4, suggesting that most individuals have a moderate amount of within-person variation over time.



The correlation matrix of physical functioning scores (Appendix 2.1) across waves shows decreasing correlations as the time gap between measurements increases. The correlations are generally high, ranging from 0.52 to 0.80. Highest correlation is found between wave 10 and wave 11. This pattern of correlations suggests that an individual's physical functioning at one point in time is strongly related to their functioning at nearby time points, but the strength of the association decreases as the time gap increases.

Model Selection

Random Intercept Models

The model selection process started with a simple variance components model and gradually increasing complexity by adding fixed and random effects for time, testing for nonlinearity, and incorporating time-invariant and finally time-varying covariates. At each step, model fit was compared using likelihood ratio tests and information criteria to determine whether the additional parameters significantly improved the model. To investigate the nature of change in physical functioning with age, we added a linear effect of age (centred at 60 years) as a fixed effect, yielding a random intercept model:

$$y_{ij} = \beta_0 + \beta_1 age60_{ij} + u_{0j} + e_{ij}, \quad u_{0j} \sim N(0, \sigma_{u0}^2), \quad e_{ij} \sim N(0, \sigma_e^2)$$

The likelihood ratio test comparing this model to the variance components model was highly significant ($p < 0.001$), indicating that physical functioning scores varied with age. The ICC decreased slightly to 0.719, suggesting that some of the between-individual variation was explained by age differences. The level 1 and level 2 variances decreased to $\sigma_e^2 = 0.274$ and $\sigma_{u0}^2 = 0.700$, respectively, confirming that age accounted for variance at both levels.

Random Slope in Growth Curve Analysis

To test whether the effect of age varied across individuals, we fitted a random slope model by adding a random effect for the slope of age:

$$y_{ij} = \beta_0 + \beta_1 age60_{ij} + u_{0j} + u_{1j} age60_{ij} + e_{ij}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u01} & \sigma_{u1}^2 \end{bmatrix} \right), \quad e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$$

The random slope variance σ_{u1}^2 was significant (likelihood ratio test, $p < 0.001$), indicating that individuals varied in their rates of change in physical functioning with age. The adjusted ICC increased to 0.757, suggesting a stronger clustering of scores within individuals after accounting for heterogeneous age trajectories. The variance of the random slope for age, denoted as $Var(u_{1j}) = \sigma_{u1}^2$, is estimated as 0.002. The relatively small random slope variance quantifies the amount of between-individual variation in the rate of change in physical functioning per one-year increase in age.

The correlation between the random intercept and random slope, denoted as $Cor(u_{0j}, u_{1j}) = \rho_{u01}$

, is estimated as -0.53. The negative correlation indicates that individuals with higher initial levels of physical functioning tend to have slower rates of decline, while those with lower initial levels tend to have faster rates of decline.

Non-Linearity

To investigate potential nonlinearity in the age effect, we added a quadratic term for age:

$$y_{ij} = \beta_0 + \beta_1 age60_{ij} + \beta_2 age60_{ij}^2 + u_{0j} + u_{1j} age60_{ij} + e_{ij}$$

The quadratic age effect was significant ($p < 0.001$), indicating an accelerating decline in physical functioning scores with advancing age. The level 1 residual variance decreased further to $\sigma_e^2 = 0.250$

, while the level 2 variances remained similar to the linear model.

Time-Invariant Variables

Time-invariant factors were first added to the quadratic age model, followed by time-varying factors. The selection of time-invariant factors involved comparing models with gender, ethnicity, and both gender and ethnicity to the quadratic age model using likelihood ratio tests. The results showed that adding gender alone significantly improved the model fit ($p < 0.001$), with an AIC reduction from 56840 to 56814. Adding ethnicity alone also significantly improved the fit ($p < 0.001$), with an AIC of 56795. When both gender and ethnicity were included, the model fit further improved ($p < 0.001$) with an AIC of 56769, indicating both time-invariant factors were important predictors of physical functioning. A direct comparison confirmed that ethnicity significantly improved the model fit beyond gender ($p < 0.001$). Therefore, the model with both gender and ethnicity was selected as the best model for the time-invariant factors.

$$y_{ij} = \beta_0 + \beta_1 age60_{ij} + \beta_2 age60_{ij}^2 + \beta_3 female_j + \beta_4 ethnicity_j + u_{0j} + u_{1j} age60_{ij} + e_{ij}$$

Time-Varying Variables

Next, the process of adding time-varying covariates involved comparing models with different combinations of education (highed), partner status (partner), housing tenure (tenurehh), and income. Models were fitted with one, two, three, and all four time-varying covariates, and the best model at each level of complexity was determined using likelihood ratio tests and information criteria. For models with a single time-varying covariate, adding education (AIC = 56621), housing tenure (AIC = 56528), or income (AIC = 56758) significantly improved the fit compared to the model with only time-invariant covariates (AIC = 56769), while adding partner status (AIC = 56756) did not. Among the single time-varying covariate models, model with housing tenure had the lowest AIC and was thus selected as the best model with one time-varying covariate.

For models with two time-varying covariates, the combination of education and housing tenure (AIC = 56423) had the lowest AIC and significantly improved the fit compared to model with housing tenure (p

< 0.001). Similarly, for models with three time-varying covariates, the combination of education, housing tenure, and income (AIC = 56421) had the lowest AIC and significantly improved the fit compared to model with highed and tenurehh (p = 0.031). Finally, the full model with all four time-varying covariates (AIC = 56420) had a slightly lower AIC than model with highed, tenurehh and income, and the likelihood ratio test comparing these two models was not significant (p = 0.115), indicating that adding partner status did not significantly improve the model fit beyond education, housing tenure, and income.

Therefore, model with three time-variant covariates was selected as the final model:

$$y_{ij} = \beta_0 + \beta_1 age60_{ij} + \beta_2 age60_{ij}^2 + \beta_3 female_j + \beta_4 ethnicity_j + \beta_5 highed_{ij} + \beta_6 tenurehh_{ij} + \beta_7 income_{ij} + u_{0j} + u_{1j} age60_{ij} + e_{ij}$$

Interpretation of Final Model

The final model selected to analyse the trajectories and predictors of physical functioning in adults aged 60 and over in the UK includes both fixed and random effects for age, as well as several time-invariant and time-varying covariates. The intercept (-0.325 in Table 2.3) represents the predicted standardized physical functioning score for a 60-year-old white male with a degree or higher education who owns their home and has an average log income, with all covariates centered at their means. The linear age term (0.170) indicates that for every decade increase in age beyond 60, the predicted physical functioning score increases by 0.170 standard deviations on average, reflecting poorer physical health. The quadratic age term (0.080) shows that the rate of decline accelerates with advancing age. Compared to females, males have 0.114 lower predicted scores, indicating better physical functioning on average. Relative to the White ethnic group, Asians have 0.377 higher and Blacks have 0.213 higher predicted scores, reflecting poorer physical functioning, while other ethnicities do not differ significantly. Those with educational attainment lower than a degree have 0.268 higher predicted scores than degree-holders, suggesting education is protective against declines in physical functioning. Compared to homeowners, social renters have 0.393 higher and private renters have 0.235 higher scores, indicating worse physical health. A one-unit increase in log income is associated with a 0.013 decrease in predicted scores, but this effect is small. Income may also be correlated with other predictors in the model, such as ethnicity (0.027 with black), which could partially explain its effect on physical functioning and lead to a less significant p-value when these covariates are included in the same model.

Table 2.2: Final Model Summary

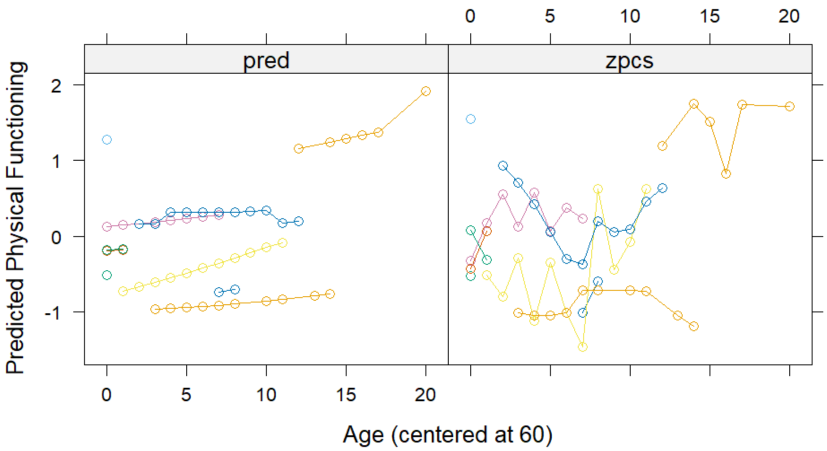
Formula				
zpcs ~ age60by10 + I(age60by10^2) + female + ethnicity + highed + tenurehh + income + (1 + age60by10 | ind)				
Intraclass Correlation (ICC)				
		Adjusted	Unadjusted	
ICC		0.736	0.645	
Random Effects - Number of obs: 28468, groups: ind, 5602				
		Variance	Std. Dev.	
Between-Individual Variance	Intercept	0.8132	0.9018	
	Age60by10	0.2108	0.4591	
Residual		0.2513	0.5013	

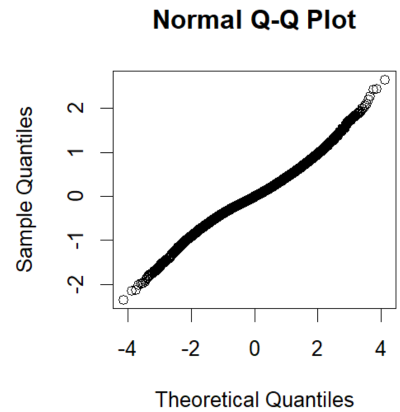
Table 2.3: P Values for Model Summary

	Estimate	Std. Error	p_values
(Intercept)	-0.3254	0.0683	2e-06
(age -60)/10	0.1697	0.0259	0
Squared of (age -60)/10	0.0805	0.0105	0
Male	-0.1143	0.0235	1e-06
Ethnicity: Asian	0.3768	0.0595	0
Ethnicity: Black	0.2126	0.0758	0.005004
Ethnicity: Other	-0.0372	0.1094	0.733757
Education: No Degree	0.2684	0.0266	0
Rent Social Housing	0.3933	0.0278	0
Rent Private	0.2349	0.0308	0
Log Income	-0.0134	0.0062	0.03065

The random intercept variance ($\sigma_{u0}^2 = 0.813$) quantifies the variability in individuals' predicted physical functioning scores at age 60, while the random slope variance ($\sigma_{u1}^2 = 0.211$) captures the between-person variability in the rate of change per decade of age. The negative correlation ($\rho_{u01} = -0.53$) suggests individuals with higher initial functioning tend to decline at a slower rate. The residual variance, denoted as σ_e^2 = 0.5013, represents the remaining within-individual variation in physical functioning that is not accounted for by the age trajectory and the other covariates in the model.

The level-1 residual variance ($\sigma_e^2 = 0.251$) represents the within-person variability in physical functioning scores over time, after accounting for age and the covariates. The adjusted intraclass correlation coefficient of 0.736 indicates 73.6% of the total residual variation in functioning is between individuals. The final model has the lowest AIC (56420.8) of all models compared, indicating the best fit to the data. The likelihood ratio tests confirm the inclusion of the quadratic age term, time-invariant, and time-varying covariates each significantly improve upon simpler models. The residual diagnostics (normal Q-Q plot) show the level-1 residuals are approximately normally distributed, meeting a key assumption of the model. The predicted trajectories show a clear trend of accelerating decline in physical functioning with age with a small non-linear relationship, with significant between-individual variation in both initial status and rate of change.





$\sqrt{\sigma_{u0}^2}$: In this case, the interval is [0.8739, 0.9306], indicating that we are 95% confident that the true standard deviation of the random intercept lies between these values.

ρ_{u01}: The interval [-0.5667, -0.4835] indicates that we are 95% confident that the true correlation between an individual's initial physical functioning and their rate of change over time lies between these values. The negative interval confirms the presence of a significant negative association, with individuals with higher initial functioning tending to have a slower rate of decline, and those with lower initial functioning tending to have a faster rate of decline.

$\sqrt{\sigma_{u1}^2}$: The interval [0.4305, 0.4878] suggests that we are 95% confident that the true standard deviation of the random slope lies between these values. This interval quantifies the plausible range of individual variation in the rate of change in physical functioning per decade increase in age.

$\sqrt{\sigma_e^2}$: The interval [0.4964, 0.5063] indicates that we are 95% confident that the within-individual variation in physical functioning that is not accounted for by the age trajectory and the other covariates in the model lies between the values.

Table 2.4: Confidence Intervals

	2.50%	97.50%
.sig01	0.873914	0.930559
.sig02	-0.56666	-0.48346
.sig03	0.430466	0.487768
.sigma	0.496383	0.506315

Limitations

The model assumes that missing data are missing at random, which may not be entirely accurate. Few individuals participated in all 11 waves, and some dropped out after only 1 wave. For example, healthier individuals might be more likely to drop out of the study at older ages, leading to an overestimation of the decline in physical functioning. Additionally, although the model includes several important predictors, other relevant factors, such as pre-existing health conditions and lifestyle factors, are not accounted for. The analysis focuses on a specific age range (60-90 years) and may not generalize to younger or older populations. The use of self-reported measures of physical functioning is also subject to potential biases. Residuals are assumed to be normal, but given income and physical functioning scores not perfectly normal in their distribution, model assumptions may not be satisfied.

Conclusion and Non-Technical Summary

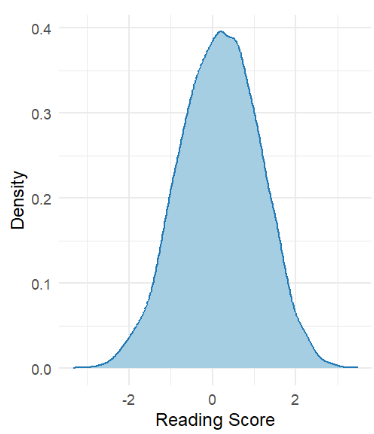
Firstly, we observed a clear trend of declining physical functioning with age, with the rate of decline accelerating as individuals get older. Non-linear age effects were also observed.

Secondly, there is substantial individual variation in both initial levels of physical functioning and the rate of decline. Most of the variation is inside an individual. Factors beyond age contribute to individual differences in physical functioning trajectories.

Thirdly, Females, on average, reported worse physical functioning than males. Individuals of Asian and Black ethnicity also tended to report worse physical functioning compared to the White reference group. Lower educational attainment and living in rented accommodation (both social and private) were associated with poorer physical functioning. Higher income, on the other hand, was linked to better physical functioning as expected.

Appendix

Appendix 1.1 Density of Normalised Reading Score



Appendix 1.2 Confidence Intervals for Final Model

Variable	2.50%	97.50%
SD of school random intercepts	0.294	0.378
SD of school random slopes for lmins	-0.16	0.655
SD of school random slopes for female	-1	-0.162
Corr of intercepts and lmins slopes	0.035	0.138
Corr of intercepts and female slopes	-1	0.301
Corr of lmins slopes and female slopes	0.024	0.207
SD of student-level residuals	0.827	0.856
Intercept	0.555	0.885
Cultural possessions	0.119	0.162
Female	0.102	0.189
Highest parental education	0.031	0.074
Wealth	-0.074	-0.03
Second-generation immigrant	-0.178	-0.001
First-generation immigrant	-0.254	-0.08
Learning minutes	-0.24	-0.033
School size	0.028	0.106
Private government-dependent school	-0.802	-0.44
Public school	-0.874	-0.536
Cultural possessions × Parental education	0.029	0.072
Cultural possessions × Wealth	-0.078	-0.04
Learning minutes × Private gov-dep school	-0.017	0.215
Learning minutes × Public school	0.042	0.26

Appendix 2.1 Correlation Matrix

	zpcs.1	zpcs.2	zpcs.3	zpcs.4	zpcs.5	zpcs.6	zpcs.7	zpcs.8	zpcs.9	zpcs.10	zpcs.11
zpcs.1	1	0.76	0.72	0.68	0.67	0.65	0.62	0.59	0.59	0.57	0.52
zpcs.2	0.76	1	0.78	0.74	0.72	0.71	0.66	0.64	0.64	0.64	0.63
zpcs.3	0.72	0.78	1	0.77	0.74	0.72	0.68	0.66	0.65	0.65	0.63
zpcs.4	0.68	0.74	0.77	1	0.78	0.74	0.73	0.7	0.66	0.65	0.63
zpcs.5	0.67	0.72	0.74	0.78	1	0.76	0.73	0.71	0.68	0.68	0.65
zpcs.6	0.65	0.71	0.72	0.74	0.76	1	0.78	0.73	0.72	0.69	0.68
zpcs.7	0.62	0.66	0.68	0.73	0.73	0.78	1	0.78	0.74	0.72	0.7
zpcs.8	0.59	0.64	0.66	0.7	0.71	0.73	0.78	1	0.77	0.75	0.71
zpcs.9	0.59	0.64	0.65	0.66	0.68	0.72	0.74	0.77	1	0.79	0.75
zpcs.10	0.57	0.64	0.65	0.65	0.68	0.69	0.72	0.75	0.79	1	0.8
zpcs.11	0.52	0.63	0.63	0.63	0.65	0.68	0.7	0.71	0.75	0.8	1