THE EFFECT OF GRADE RETENTION ON IMMIGRANT-NATIVE READING GAPS

Comparative research on the Netherlands, Germany, and Austria

Master thesis Sociology by Jildou Tromp
Student: 586758, Prof: Sanneke de la Rie. Words: 9904

Abstract

Because social mobility strongly depends on educational performance, policymakers have expressed concerns about the difference in educational performance between native and immigrant children (I/N gap). One way some countries try to help students to catch up is to make them repeat a grade. Since grade retention is associated with negative outcomes on school performance, this research investigates whether grade retention can account for the I/N reading gap. This paper draws on data from PISA 2018 on 15-year-old children in the Netherlands, Germany, and Austria to examine cross-country differences in the role of grade retention in explaining the I/N reading gap. Because the data is nested, multi-level regression analyses were performed. In all countries, first-generation immigrants are most likely to repeat a grade, followed by second-generation immigrants and, in the last place, natives. In the Netherlands, the odds for first-generation immigrants to repeat a grade are highest among the three countries. The results remained significant even after controlling for socioeconomic conditions. Second-generation students' likelihood of repeating a grade is explained by the control variables, gender, language, track & parental education. Natives have the highest reading scores, followed by second-generation immigrants, and in the last place, first-generation immigrants; this is true for all three countries. In the Netherlands, grade retention can account for (part of) the I/N reading gap for second-generation immigrants; it widens the gap.

In contrast, the gap gets smaller in Germany for first-generation immigrants. For the other groups, the difference in percentages of the I/N was not substantial, and therefore the conclusion is that grade retention does not account for the gap between natives and first-generation immigrants in the Netherlands and both generations in Austria. For second-generation students in Germany, grade retention is not related to the I/N gap.

Keywords: Grade retention, Immigrant/native gap, PISA, Reading performance

Index

1. Introduction

From 2000 until 2016, the immigrant population increased from 78 million to 120 million in the OECD countries (Organisation for Economic Co-operation and Development) (d'Aiglepierre et al., 2020). Children of immigrants often have problems with the new education systems due to cultural differences and the new language. They often come from low socio-economic (SES) families, and their parents are generally lower educated (Herzog-Punzenburger, 2003; Riphahn, 2003; Eldering & Kloprogge, 1989). Education is vital to create more life chances. Better school performance mainly results in higher earnings and better health, and people tend to be happier (McMahon & Oketch, 2013). At the societal level, it benefits political stability, lower crime rates, and lower healthcare costs. In the case of immigrants, education is crucial for successful integration, which lowers social exclusion and discrimination (Dronkers & Korthals, 2016). Because the family's socio-economic status plays a prominent role in the intergenerational transmission of social inequality, it is often hard for children from low-SES families to attain higher education. There is a need for social mobility to break this pattern and for children to be able to climb the social ladder. Because social mobility strongly depends on educational performance, policymakers have expressed concerns about the difference in educational performance between native and immigrant children (Park & Sandefur, 2010).

School performance gap

The differences in educational performance between native and immigrant children are referred to in the literature as the immigrant-native gap. The school performance gaps between immigrant and native students (I/N gaps) vary across countries. In European countries, immigrants are often lower educated than natives (Rindermann & Thompson, 2016). The performance gap is problematic because education is crucial to have chances in the labor market; social mobility is difficult to attain without education. Therefore, it is interesting to investigate and explain the position of immigrant students in education in Europe.

Comparative research on multiple OECD countries, using PISA (Program for International Student Assessment) data of 2018, shows lower overall scores for immigrant students, both first and second generation, on mathematics and science than for their native peers (Rodríguez et al., 2020) (Schnell & Azzolini, 2015). Information from 2018 on the reading performance shows that native students have the highest scores, followed by second-generation immigrant students. First-generation immigrants showed the lowest reading scores (Peña-López,2019). Overall second-generation immigrants perform better than first-generation immigrants of the same origin (Levels, Dronkers & Kraaykamp, 2008).

Possible explanations for the differences in the I/N performance gap can be found in the different educational systems, immigration policies, and cultural differences (Park & Sandefur, 2010). Also, the poorer socioeconomic status of immigrants, compared to natives, is related to the

underachievement in the school performance of immigrant children (Park & Sandefur, 2010). Schnell and Azzolini (2015) showed that parental education plays a larger role in explaining the I/N gap than families' financial circumstances; this is especially the case in traditional immigration countries like the Netherlands, Germany, and Austria.

Grade retention

Different educational systems try to fix the school performance gaps in different ways; One way is to apply grade retention, which is implemented in some countries to give children the opportunity to develop themselves to the required level and close the I/N gap. Although the intention is to decrease the gap between high- and low-performing students, studies examining the effects of grade retention on school performance show detrimental effects. Only short-term beneficial effects show because students tend to get higher marks if they must repeat the same tests as the year before. However, retention shows no effect in the mid-term (Klapproth et al. l., 2016) and even harms school achievement in the long run as it increases dropout and decreases the chances of enrollment in post-secondary education (Hauser, 1999, Rumberger & Larson, 1998).

Between countries, there is much variation in the practice of grade retention. In Norway, there is no retention at all, while in the Netherlands, grade retention in primary education is relatively common, with 21% against 5% in lower secondary education. It is the other way around in Germany; more grade retention happens in lower secondary education (13%) instead of 7% in primary education. In Austria, the differences in grade retention are much lower; for both primary and secondary education, it is around 5% (Goos et al., 2013). Spain has the highest-grade retention, with almost 30 percent in secondary education (Goos et al., 2013).

Despite the rules and regulations of looking at academic performance in the decision to make a student repeat a grade, there is growing evidence that the chances of retention are related to gender, ethnicity, and SES (Frey, 2005). Park & Sandefur (2010) find that in countries where grade retention is practiced, immigrant students are more often in lower grades than their native peers of the same age. They also focused on the influence of grade retention on reading achievement for native and immigrant students and found that the I/N reading gap was wider in countries where immigrant students were more likely to repeat a grade (Park & Sandefur, 2010).

Besides socioeconomic factors on school performance (Park & Sandefur, 2010), the type of educational system in a country influences the I/N gap (Burger, 2016). The fact that immigrants are more likely to repeat a grade and the potential adverse effect of repeating a grade on school performance might explain why grade retention widens the reading performance gap between native and immigrant students. However, the results of Park & Sandefurs' (2010) research might be partly confounded by varying educational systems since they compared countries varying widely in educational systems; in terms of grade retention and tracking by different levels. The European Commission (2018) found that traditional immigration countries (the United Kingdom, the

Netherlands, and Germany) show larger educational performance gaps than other European countries. Also, educational systems with early tracking by different levels, mostly vocational versus academic education, influence the I/N gap negatively. Comparative research on same-model countries that apply grade retention will reduce the influence of systematic and large variations in educational systems. Analyzing only one educational model can better isolate grade retention's role in accounting for cross-national differences in I-N. Therefore, this research will focus on the variation of I/N reading gaps in traditional immigration countries that practice the same educational model of early tracking, including the Netherlands, Germany, and Austria (European Commission, 2018).

Research by Park & Sandefur (2010) did not distinguish between first- and second-generation immigrant students. This distinction will be made in this research because, according to the literature on I/N gaps, there are differences in performance between first- and second-generation immigrants (Rodríguez et al., 2020; Schnell & Azzolini, 2015; Levels, Dronkers & Kraaykamp, 2008). We already know that immigrant students are more likely to repeat a grade. However, whether this is different for first- and second-generation immigrant students is unknown.

Prior research shows conflicting results on the effect of timing of grade retention on educational performance; Fine and Davis (2003) found that enrollment in post-secondary education decreased when the grade a student was retained increased. Other literature showed that early-retained children had better results than later-retained students further in their academic careers (Silberglitt et al., 2006). Meisels and Liaw (2003) found contrary evidence; retention timing is not related to school performance, but retention itself harms school performance. Despite having the same education model, the countries being studied differ in their timing of grade retention; the Netherlands mainly practices retention in primary school, while Germany mostly in lower-secondary school. In Austria, there is not much difference (European Commission, 2018). By studying possible differences in the timing of grade retention between countries, there will be more clarity if it makes any difference in the I/N reading gap if grade retention is mainly practiced in primary or lower secondary education.

Instead of using the PISA 2000 like Park & Sandefur (2010) did, the more recent PISA 2018 dataset will be used. As the immigrant population increased over that time (d'Aiglepierre et al., 2020), the differences in educational performance between I/N students might have shifted. Therefore, it is essential to use more recent data. Since the questionnaire of PISA 2018 was focused on reading, we are only looking into reading performance. Data on the other subjects (mathematics and science) do not give a complete image since only a few questions were included on those subjects.

It is socially relevant to examine how grade retention affects the I/N gap relative to other factors such as gender and SES (Park & Sandefur, 2010). Policymakers need to have this information to make plans to improve the position of immigrants in education and close the I/N reading gap. When grade retention plays a prominent role, it is essential to look at its implementation, considering the effect of timing. Also, grade retention's positive or negative impact on performance might differ between the

different generations of immigrant students. Therefore, it is also essential to make this distinction in research so policies can be based on these differences.

This research will focus on countries with the same educational model to isolate the role of grade retention. Because traditional immigration countries, and countries with educational models of early tracking, increase the I/N gap, it is interesting to focus on countries with these characteristics as well. The countries that meet those characteristics and thus are the focus of this research are the Netherlands, Germany, and Austria (European Commission, 2018).

Research questions

The aim of this research is to discover how grade retention can account for the reading gap between native and immigrant students. To find out if the role differs across countries, the Netherlands, Germany, and Austria will be compared. Therefore, the following research questions are formulated:

"To what extend can grade retention account for the immigrant-native reading gap and how does this vary across the Netherlands, Germany, and Austria?"

A. How do the Netherlands, Germany, and Austria vary in the degree of grade retention of native, second- and first-generation immigrants?
B. How do reading gaps between native, second- and first-generation immigrant students vary across the Netherlands, Germany, and Austria?
C. To what extent is grade retention related to the overall I/N reading gap, does the role of grade retention differ when examining the reading gap between first generation immigrant students and native students, versus the reading gap between second generation immigrant students and native students? To what extent does this vary across the three countries?
D. Can the differences in the timing of grade retention (in primary versus secondary education) account for part of the I/N reading gap? To what extent does this vary across the three countries?

2. Theoretical framework

In some countries, grade retention occurs when a student does not meet the educational requirements at the end of a school year; the student must repeat the same year. As discussed before, this has primarily negative consequences in the long run; Children's educational achievement does not go up, the chances of early dropout are higher, and the chance of students enrolling in post-secondary education decreases (Rumberger & Larson, 1998). This might be because of the lack of confidence children experience when they must repeat a grade. Also, gender, immigrant status, and low SES are associated with increased chances of repeating a grade; male students, students from lower SES, immigrant students, and students from lower education levels are more likely to repeat a grade than female native students from schools with higher SES and higher education levels (Hauser, Pager, & Simmons, 2001). These results remained after controlling for achievements, which means that even with similar results, some students have higher chances of being retained than others (Klapproth et al. l., 2016). Park & Sandefur (2010) found that immigrant students are more likely to be in lower grades than native students in countries that practice grade retention. (Germany, Austria, Belgium, Luxembourg, Switzerland, and France). In the other five countries (UK, US, Denmark, Norway, and Sweden) that do not practice grade retention, the difference between immigrant and native students is much lower. The differences in grade retention were calculated by looking at the difference in percentage from the average grade (10th grade) for students aged 15. In Austria, 48.8% of native students are in lower grades than the average, while 71.2% of immigrant students are in lower grades. First-generation immigrant students must get used to the new culture and language, which might be of considerable influence on not being able to get along right away and needing some more time, resulting in more grade retention than second-generation immigrant students (Klapproth et al. l., 2016).

To discover if immigrant students are retained more often than natives and if it is different for the two generations, the following hypothesis is created based on sub-question A:

Hypothesis 1: *First generation immigrant students are most likely to be retained, followed by second generation immigrant students, and native students are least likely to be retained in the Netherlands, Germany, and Austria*

Research by Park & Sandefur (2010) shows the difference in reading performance between native and immigrant students per country; the numbers represent the points that immigrant students score lower than natives. In the United Kingdom, the difference is the smallest, with -31 points. In Austria (-79), Belgium (-100), Germany (-80), Switzerland (-81), and Luxembourg (-88), the differences are more considerable. When looking into reading performance in 2018 in the OECD countries, we see the same tendency in all three countries; native students score best, followed by second-generation immigrants, and in the last place first-generation students (Schleichler, 2006) (Rodríguez et al., 2020). The following hypothesis belongs to sub-question B on how the countries vary in the I/N reading gap:

Hypothesis 2: *Native students will show the highest reading performance, followed by second-generation students, and first-generation students will have the lowest scores in the Netherlands, Germany, and Austria.*

The possible negative effect of grade retention on reading performance and the higher likelihood of repeating a grade for immigrant students implies that grade retention might be essential in explaining the reading performance gap between immigrant and native students. Park and Sandefur (2010) tested this hypothesis in 10 countries and proved it true. Their results showed that when the chance of being retained for immigrants increases by 1%, the reading gap increases by 1.37%. In Belgium, the difference between immigrant and native students in the chance of being retained was the highest at 36%, while in Norway, it was the lowest at 4% (Park & Sandefur, 2010). If we take the difference between the countries (36-4=32) and multiply it with the 1.37% effect in the I/N reading gap, we see a difference in reading performance of 44 points. In other words: in Belgium, where grade retention is practiced, the I/N reading gap is 44 points higher than in Norway, where grade retention is not practiced. Therefore, the expectation is that grade retention will have a negative impact on the I/N reading gap. Because of these meaningful and large differences, the importance of investigating the influence of grade retention on the differences in performance and closing the I/N gap gets very clear. This research adds to previous research that a distinction will be made between first- and second-generation immigrants. Since the results of other research showed differences in the generations in performance and retention, it is very likely that these differences also exist when looking at the effect of retention on the I/N reading gap. As we expect first-generation immigrants to be retained more often, we also expect the effect of retention on reading performance to be the largest for this specific group. According to the European Commission (2018), Austria practices less grade retention than the Netherlands and Germany and is therefore expected to show the smallest retention effect on reading performance.

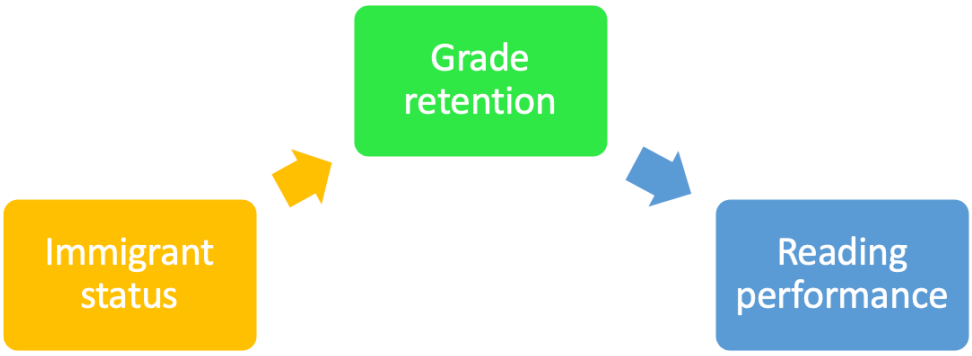
Therefore, the hypothesis accompanying sub-question C on the effect of grade retention on the I/N reading gap for the three countries is:

> **Hypothesis 3:** *The (negative) influence of grade retention on the I/N reading gap is larger for first-generation immigrants than for second-generation immigrants. The effect will be larger for the Netherlands and Germany than for Austria.*

Even though little research has been done on the timing of grade retention, Silberglitt et al. (2006) used longitudinal data on reading performance and found that early-retained students showed better reading performance than later-retained students. Most evidence proved retention to influence performance negatively; only mathematics can have a positive effect in the short term (Klapproth et al. l., 2016). Since this research focuses on reading, a negative relationship is expected.
Therefore, the last hypothesis based on sub-question D on the effect of timing of grade retention is:

> ***Hypothesis 4****: Repeating a grade in lower secondary education will explain more of the I/N gap in reading performance than repeating a grade in primary education*

Figure 2.1. *Conceptual model*



3. Methods and data-analyses

This study will use quantitative data from the Program for International Student Assessment (PISA). PISA measures students' reading, mathematics, and science abilities. Data from 15-year-old students has been collected in over 80 countries since 2000. With the results, policymakers and researchers can compare the quality and equity of various educational systems and react to this. The focus of the questionnaire in 2018 was on the reading skills of 15-year-old students, regardless of the grade.

In addition to information about school performance, PISA also included standardized questions about individual and family characteristics, such as socioeconomic background, parental education and occupation, country of origin, the language spoken at home, and family composition. In 2018 around 600.000 students responded to the questionnaire. At level 1, the students' sample sizes are as follows: the Netherlands (N= 4765), Germany (N= 5451), and Austria (N= 6802). At level 2, the schools, the sample sizes are the Netherlands N=160, Germany N=225, and Austria N=301. The complete dataset is available on the OECD website (OECD, 2020).

PISA used a random sampling method whereby first, the schools were selected randomly, and after that, the students from the schools were randomly picked (OECD, 2020). The selected students could decline when they did not want to participate; the participation was voluntary. All data was anonymized to protect the privacy of the respondents.

Dependent variable: reading performance

All assignments together contain 15.5 hours of testing time. Because it was too much for each student to complete, the tests were grouped into clusters per domain (reading, mathematics, and science). Every student got four clusters which were randomly selected. Because the main domain in 2018 was reading, the students made the most assignments in this domain and only a few in mathematics and science. An estimated proficiency score with IRT analyses (Item Response Theory) could be determined based on the completed assignments and background characteristics (Mislevy, Beaton, Kaplan & Sheehan, 1991). Even if students did not complete all assignments, an estimated score could still be calculated. With the estimated scores, ten plausible scores were calculated per domain. The mean of the ten reading performance scores is used to create the variable: reading performance. The unstandardized coefficients are used.

Immigrant status

This research defines three groups of students: natives, first-generation, and second-generation immigrants. With the questions of the PISA data on the country of birth of father, mother, and child, a variable was created by PISA to define the three groups, coded as native (0), second-generation (1), and first generation (2). With native students, at least one of the parents and the child are born in the test country. Second-generation immigrants are students born in the test country, but both parents were born in another country. Finally, first-generation immigrant students are born in another country than the test country, and at least one of the parents is also born in another country than the test country. Dummies were created for each variable. Native students serve as the reference category in the regression models.

Grade retention

This research examines whether cross-national variation in the I/N reading gap can be accounted for by grade retention. The PISA data asked students if they ever repeated a grade. Now 0 represents the students who never repeated a grade, and 1 represents students that did repeat a grade. Also, a more specific question determines whether students were retained in primary education (ISCED 1) or lower secondary education (ISCED 2; (UNESCO Institute for Statistics, 2012). A new variable was created to combine the information on timing of retention. Hereby 0 represents students who did not repeat, 1 is for the students who repeated a grade in primary education, 2 is for repeating lower secondary education and 3 is for repeating both primary and secondary education. Dummies were made to be able to put this variable in the analyses whereby never repeating a grade is the reference category. To improve the readability, in the text primary education will be referred to as PE and lower secondary education as SE.

Control variables

Performance does not only differ between immigrant and native students but also between boys and girls. Research shows that girls, on average, have better reading performance than boys (Mullis, Martin, Gonzalez, & Kennedy, 2003); a reason mentioned for this gap is gender-specific study cultures whereby girls are more study-focused than boys (Houtte, 2004). These results are similar when looking at children of immigrants (Dronkers & Kornder, 2015). Literature also shows that shows that boys are more likely to be retained (Mullis, Martin, Gonzalez, & Kennedy, 2003); therefore, gender will be a control variable. 0 represents all girls, and 1 represents boys.

The language students speak at home can be found in the available PISA data and is used to determine if children speak the language of the test country or another language at home. Suppose the language spoken at home differs from the language of instruction in schools. In that case, this might

negatively influence reading ability, as this is measured in the native language (Park & Sandefur, 2010). The variable is coded 0 for the language of the test country, and 1 otherwise.

As discussed before, all three countries apply tracking at different levels at an early age. Since the traditional immigration countries show the largest I/N reading gaps, taking the track as a control variable is interesting. Two tracks are distinguished: vocational (0) and academic (1).

Because in early-tracked education systems and traditional immigration countries, parents' educational level matters in children's school performance, parents' highest education will be taken as a control variable. Parental education is divided into three categories based on the ISCED levels: lower secondary education or less (1), upper secondary education completed (2), and tertiary education completed (3), in this research it will be referred to as: lower, middle, and higher education. To perform regression analyses, dummy variables were created for the different groups where low education serves as the reference category.

Analytical methods

Multilevel regression analyses will be used to test the hypotheses because there are two levels: students and schools. Students (level 1) are in schools (level 2), so the data is nested. Therefore, there is a dependency between the students; students in the same schools are probably more alike compared to students in different schools. Multilevel analyses take the hierarchical nature of the data into account. When the data hierarchy is not considered, it can result in biased estimates (Field, 2009). All analyses are performed separately for the three countries, as the number of countries (N=3) is too small to be considered as a level in the multilevel analyses (Hox, 2010).

Data-analyses

Because of the nested data in this research, a random intercept model is used. With a random intercept model, the intercepts of the regression line can vary over schools. With that, the scores on reading (dependent variable) for each observation are predicted by the intercept that can vary across schools (Austin, Goel & Van Walraven, 2001). The slopes are held fixed, meaning that associations between the predictors and reading scores are not allowed to vary across schools. This was done as the current study does not aim to test hypotheses regarding varying associations across schools nor to test cross-level interactions. All variables used in the analyses are at the student level (level 1), and there are no variables on the school level (level 2).

The analyses are performed per country, and the models are based on the sub-questions and hypotheses. Hypothesis 1 has a different dependent variable (grade retention) than the other hypotheses, where reading performance is the dependent variable. Because there are two

dependent variables, there are two intercept-only models with the dependent variables included, which decomposes the variance in reading scores into a student-level variance and a school-level variance. These models are used to check whether there is significant variation in reading scores at the school level and if multi-level data analyses are required. To quantify the variance component in reading scores at the school level, the Intraclass Correlation Coefficient (ICC) is calculated (Hox, Moerbeek, & Van de Schoot, 2017).
Since the dependent variable of H1 (Grade retention) is dichotomous, a multi-level logistic regression will be used. The beta coefficients and odds ratio are presented in a table per country. Model 0 contains only the dependent variable grade retention. In model 1, the independent variable immigration status is added. In model 2, the control variables are added.

New tables will be presented for the remaining hypotheses (H2, H3, and H4). Model 0 is the intercept-only model with the dependent variable Reading performance. In model 1, the independent variable Immigrant status was added. In model 2, the control variables are added. With this hypothesis 1, the effect of immigrant status on reading performance is tested. To examine whether grade retention accounts for the I/N reading gap (H3), grade retention is added in model 3. Because second- and first-generation immigrants are added as dummy variables, the differences in the effect of retention for the generations can be discovered. To see if Grade retention changes the I/N reading gap, the percentages of change in reading scores between model 2 (without grade retention) and the models with Grade retention (3) are calculated for second- and first-generation immigrants.
The last hypothesis on the Timing of retention is formulated globally since there are no specific expectations for the different generations, but it will be tested in a way that distinguishes the different immigration statuses. In model 4, grade retention will be removed, but in its place, the dummy variables on Timing of retention will be added (repeat PE, repeat SE, and repeat PE & SE, with no retention as the reference). This model will test whether the Timing of retention influences the reading scores. In model 5, only retention in PE will be added to see how the I/N reading gap chances. In model 6, only retention in SE will be added to test if this influences the I/N gap. The dummy variable of repeating both (PE and SE) is not analyzed since there is no hypothesis on that in this research. For Timing grade retention, the percentages in the change of the I/N reading gap will be calculated to say something about the effect of Timing retention on the gap. The *Statistical Program for Social Science (SPSS)* will be used to analyze the data.

4. Results

Descriptives

This chapter contains the results of the analyses. It is important to note that PISA is based on a 'complete case' at the student level. Students with missing values on one of the variables are not considered (Field, 2009). At first, attention was paid to univariate outliers in the data to check if there were extreme outliers for the variables. According to the Tukey method, there are no extreme outliers that had to be removed (Blaine, 2018). Also, there is checked for multivariate outliers by combining variables (Great mean= GM, reading performance, and immigration status). A few minor outliers are not removed because they are not extreme.

The descriptive statistics for the continuous variable reading performance per country are presented in table 4.1. The average reading score of all students in all OECD countries in 2018 was 456. We see that in all three countries, native students have the highest scores, followed by second-generation immigrants. First-generation immigrants have the lowest reading ability.

The categorical variables per country are presented in table 4.2. An overview of the categorical variables divided by Immigrant status can be found in Appendix A. When looking at grade retention, we see first-generation immigrants repeat a grade most often, followed by second-generation immigrants. Natives are retained least often. In the Netherlands, all groups are retained most often in PE, while in Germany, all groups are retained more often in SE; this is in line with the theory about the timing of grade retention (European Commission, 2018). In Austria, natives are more often retained in PE than in SE, while second- and first-generation immigrants are retained more in SE than in PE. The numbers of the global retainment variable and the timing retention variable differ in the category of students who never repeated a grade because there are more missing values of the Timing retention variable. Also, we are only looking at Timing retention in ISCED 1 and 2, while the global retention variable might also contain students who repeated a grade in ISCED 3.

Table 4.1. Descriptives continuous variable reading performance

		Netherlands	Germany	Austria
Native	**Min**	197.0	204.8	194.4
	Max	768.5	816.4	739.0
	Mean	4936	521.2	501.5
	SD	99.4	95.6	91.8
Immigrants	**Min**	169.8	198.4	202.2
	Max	681.9	761.2	688.9
	Mean	413.0	546.9	436.3
	SD	92.7	103.6	91.9
Second-generation	**Min**	224.4	209.3	221.8
	Max	681.9	761.2	688.9
	Mean	426.5	477.7	446.5
	SD	88.1	102.6	87.0
First-generation	**Min**	169.8	215.9	202.2
	Max	667.3	684.3	666.1
	Mean	388.1	458.4	419.3
	SD	98.2	94.1	94.7

Table 4.2. Descriptives categorical variables Netherlands, Germany, and Austria

	Netherlands		Germany		Austria	
	Count	%	Count	%	Count	%
Grade retention						
Did not repeat	3152	82%	3784	81.0%	5771	86.7%
Did repeat	690	18%	890	19.0%	889	13.3%
Timing retention						
Never	3135	65.8%	3750	68.8%	5878	86.4%
Repeat PE	351	7.4%	218	4.0%	181	2.7%
Repeat SE	188	3.9%	289	5.3%	162	2.4%
Both	21	0.4%	30	0.6%	57	0.8%
Gender						
Girl	2330	48.9%	2525	46.3%	3321	48.8%
Boy	2435	51.1%	2926	53.7%	3481	51.2%
Track						
Vocational	3776	79.2%	5407	99.2%	4951	72.8%
Academic	989	20.8%	44	0.8%	1851	27.2%
Language at home						
Test country	4171	89.5%	3957	82.2%	5460	80.8%
Other	491	10.5%	855	17.8%	1297	19.2%
Parental education						
Lower	206	4.5%	811	18.6%	240	3.7%
Middle	1289	28.2%	1101	25.3%	2669	40.6%
Higher	3083	67.3%	2437	56.0%	3665	55.7%

Correlation

Before conducting the multilevel regression, we will look at how the different variables of this research correlate. A correlation test using Pearson was performed separately for each country. The tables can be found in appendix A. To improve the readability of the results, the country abbreviations are used; Netherlands= NL, Germany= DE, and Austria= AT.

In all countries, there is a negative, weak to very weak, and significant correlation between immigrant status and reading performance *(*NL: *r*=-.264*, p*<.001*,* DE: *r*= -.305*, 0*<.001*,* AT*: r*=-.275*, p*<.001*).* On average, students with an immigration background have lower reading scores.

The other main variable of interest, grade retention, has a weak to very weak, positive, and significant correlation with immigrant status; immigrants are more often retained than natives *(NL: r=.114, p<.001, DE: r= .120, 0<.001, AT: r=.171, p<.001)*. When looking at grade retention and reading performance, a significant weak to very weak, negative correlation can be found; when students are retained, they have lower reading scores *(NL: r=-.254, p<.001, DE: r= -.316, 0<.001, AUT: r=-.238, p<.001)*.

There is a significant correlation between gender and reading performance; girls have slightly higher scores than boys. Also, the correlation between grade retention and gender is significant but very weak: boys have a slightly higher chance of being retained. When looking at the language spoken at home, we see a weak, negative but significant correlation; children who speak a foreign language have lower reading scores. The relation between language and retention is very weak but also significant; children who speak a foreign language are a little more likely to be retained. The correlations for track differ among the three countries; however, they are all positive and significant (p<.001), for Germany, the correlation is very weak (r=.117), and for Austria, it is weak (r=.393). For the Netherlands, it is moderate (r=.563). Parents' education correlates significantly and positively but very weakly with reading performance.

Multilevel analyses

First the tables on H1 are presented per country. Second, the analyses on H2, H3 and H4 are presented in one table per country.

Multilevel logistic regression

H1: First generation immigrant students are most likely to be retained, followed by second generation immigrant students, and native students are least likely to be retained in the Netherlands, Germany, and Austria

The odds of repeating a grade are presented for second- and first-generation immigrants compared to natives. For the control variables, the odds ratios of boys, students who speak another language than the test country, being in an academic track, and having parents with middle or higher education are presented.

Before performing the multilevel regression, the ICC is calculated (ICC= school level variance/ (school level variance + student-level variance); the variance in grade retention at the school level is 34.63% for the Netherlands, 35.67% for Germany, and 29.37% for Austria (Tables 4.3, 4.4, and 4.5, Models 0). We also see that this variation in reading performance at the school level is significant in all countries. In conclusion, we know that students within schools are not independent, and this hierarchical structure needs to be considered when analyzing the data.

When adding Immigrant status in model 1, we see that it is associated with grade retention; both immigrant groups are more likely to be retained than natives. As expected, first-generation immigrants are most likely to be retained; their chances of having to repeat a grade are 3.550 (NL), 2.128 (DE), and 2.617 (AT) times higher than natives. Second-generation immigrants come in second place; they are 1.570 (NL), 1.557(DE), and 1.632 (AT) times more likely to be retained than their native peers.

In all countries, Second-generation immigrant status is not associated with repeating a grade more often than natives, after controlling for gender, language, track, and parental education (models 2); the significant control variables can explain the fact that they are retained more often. For first-generation immigrants, the odds of repeating a grade are 2.869 (NL), 1.654 (DE), and 1.935 (AT) times higher than for natives, after factoring for the control variables. Adding the control variables improves the model fit significantly for Germany and Austria (models 2, rows 13).

Table 4.3. Multi-level logistic regression grade retention Netherlands

	Model 0		Model 1		Model 2	
	b (SE)	Odds ratio	b (SE)	Odds ratio	b (SE)	Odds ratio
		Fixed Part				
Intercept	-1.558 (1.321)	.211	-1.633 (1.309)	0.195	-1.202 (.922)	.301
Second generation (native=0)			.454*** (.127)	1.570	.247 (.1696)	1.280
First generation (native=0)			1.267*** (.273)	3.550	1.054*** (.303)	2.869
Gender (Girl = 0)					.329*** (.090	1.390
Language at home (0= test country)					.226 (.194)	1.254
Track (0=vocational)					-1.292*** (.140)	.275
Middle education (lower=0)					-.329 (.212)	.720
Higher education (lower=0)					-.386 (.204)	.680
		Random Part				
School level	1.743		1.711		.808	
ICC	0.3463		0.3421		.1972	
Deviance	16008.284		15763.267		18054.331	
Reference model			0		1	
Fit-improvement			χ^2= 245.02 DF= 2 P<.001		χ^2= - 2291.06 DF= 5 P<.001	

Note. ***p<.001, **p<.01, *p<.05; SE = standard error. n. s=non-significant

Table 4.4. Multi-level logistic regression - Grade Retention Germany

	Model 0		Model 1		Model 2	
	b (SE)	Odds ratio	b (SE)	Odds ratio	b (SE)	Odds ratio
				Fixed Part		
Intercept	-1.506 (1.351)	.222	-1.628 (1.349)	0.196	-1.525 (1.287)	0.218
Second generation (native=0)			.443*** (.093)	1.557	.058 (.121)	1.060
First generation (native=0)			.755*** (.142)	2.128	.503* (.197)	1.654
Gender (Girl = 0)					.312*** (.071)	1.366
Language at home (0=test country)					.522*** (.132)	1.685
Track (0=vocational)					-.712* (.359)	.491
Middle education (lower=0)					-.306** (.108)	.736
Higher education (lower=0)					-.435*** (.095)	.647
		Random Part				
School level	1.824		1.792		1.647	
ICC	0.3567		0.3526		0.3336	
Deviance	19615.449		19436.857		17636.365	
Reference model			0		1	
Fit-improvement			χ^2 = 178.59 DF= 2 P<.001		χ^2= 1800.49 DF= 5 P<.001	

Note. ***p<.001, **p<.01, *p<.05; SE = standard error. N. s=non-significant

Table 4.5. Multi-level logistic regression grade retention Austria

	Model 0		Model 1		Model 2	
	b (SE)	Odds ratio	b (SE)	Odds ratio	b (SE)	Odds ratio
			Fixed Part			
Intercept	-1.783 (1.170)	.168	-1.925 (1.147)	.146	-1.698 (1.141)	.183
Second generation (native=0)			.490*** (.083)	1.632	.207 (.113)	1.230
First generation (native=0)			.962*** (.113)	2.617	.660*** (.140)	1.935
Gender (Girl = 0)					.108 (.058)	1.114
Language at home (0=test country)					.282** (.107)	1.326
Track (0=vocational)					-.242*** (.067)	.785
Middle education (lower=0)					-.269 (.165)	.764
Higher education (lower=0)					-.204 (.164)	.815
			Random Part			
School level	1.368		1.314		1.274	
ICC	.2937		.2854		.2791	
Deviance	26811.257		26491.454		25749.396	
Reference model			0		1	
Fit-improvement			χ^2=3 19.80 DF= 2 P<.001		χ^2= 742.06 DF= 5 P<.001	

Note. ***p<.001, **p<.01, *p<.05; SE = standard error. n. s=non-significant

Multilevel regression

Below, the results of H2, H3, and H4 will be presented as these hypotheses share a common dependent variable: Reading performance. The results for the Netherlands can be found in Table 4.6, for Germany in Table 4.7, and Austria in Table 4.8. The estimates of second- and first-generation immigrants represent how much higher the scores of natives are compared to each immigrant group to determine the I/N gap. The estimates of grade retention represent how much higher the reading scores are for students who did not repeat a grade. The same goes for the estimates of the Timing of grade retention (PE, SE, and both). Furthermore, the estimates of girls, students that speak the language of the test country, students in vocational tracks, and students whose parents are lower educated are given in the tables. The ICC shows the percentage of variation in reading performance at the school level (level 2). For the Netherlands, this is 57.31%, for Germany, 47.36%, and for Austria, 50.23 % ($p<.001$). In conclusion, we know that students within schools are not independent, and this hierarchical structure needs to be considered when analyzing the data.

> *H2: Native students will show the highest reading performance, followed by second-generation students, and first-generation students will have the lowest scores in the Netherlands, Germany, and Austria.*

In model 1, the independent variable immigrant status is added. In all three countries, we see that native students have the highest scores, followed by the second generation, and in the last place first-generation immigrants. Differences between the three categories are statistically significant (See Table 4.6 for NL, Table 4.7 for GE, and Table 4.8 for AT). The difference in reading performance between native and second-generation students seems to be the smallest in Germany, slightly larger in the Netherlands, and largest in Austria (Tables 4.6, 4.7, and 4.8, row 2). The gap between first-generation immigrants and natives seems to be smallest in the Netherlands, followed by Austria, and widest in Germany (rows 3).

After adding the control variables (models 2), the influence of immigrant status on reading performance remains significant in the Netherlands and Austria for both generations; Immigrant status is negatively and significantly associated with reading performance. Adding the control variables decreases the differences in reading scores between second-generation immigrants and natives in the Netherlands and Austria since the significant control variables also account (partly) for reading scores (Tables 4.6 and 4.8 rows 3, model 2). Being a second-generation immigrant is no longer associated with lower reading scores compared to being a native student in Germany. For first-generation

students, the results remain significant in all countries after adding the control variables. Being a first-generation immigrant has a significant and negative influence on reading performance. Hypothesis 2 is accepted for the Netherlands and Austria; natives have the highest performance, followed by the second-generation and first-generation immigrant students. For Germany, the hypothesis can be partly accepted; first-generation immigrants have lower reading scores than natives, while second-generation immigrants do not statistically differ from natives.

H3: The (negative) influence of grade retention on the I/N reading gap is larger for first-generation immigrants than for second-generation immigrants. The effect will be larger for the Netherlands and Germany than for Austria.

In models 3 (see Tables 4.6, 4.7, and 4.8), grade retention is added to models 2 as the independent variable. There is a significant relation between grade retention and reading performance in all three countries; on average, students who have not been retained show higher reading performances (see rows 4). Adding grade retention improves the model fit significantly for all countries compared to model 2 (tables 4.6, 4.7, and 4.8 rows 18, models 3). To see whether Grade retention can account for (part of) the I/N gaps in reading performance in the three countries, the differences in the coefficients for Immigrant status between models including and excluding the variable Grade retention are calculated (see Table 4.9). The dummies of second- and first-generation, representing Immigrant status, remain significant after adding Grade retention for both generations in the Netherlands and Austria, and for first-generation immigrants in Germany; Grade retention is associated with the I/N reading performance gap.

For the Netherlands, an interesting thing happens; factoring in Grade retention substantially increases the gap in reading between natives and second-generation immigrants by 7.96 percent (See Table 4.9), but it closes the gap between natives and first-generation immigrants slightly. For Austria, the gap decreases for both generations after adding grade retention, but the changes are minor (Table 4.9, row 2). Second-generation immigrants in Germany are not associated with having lower reading scores than natives when they are retained. The gap in Germany decreases for first-generation students after adding grade retention by 10.11% (see Table 4.9, row 2).

There is no unambiguous answer to the hypothesis that applies to all countries, but the results are different from the expectations. The expectation was that the influence of grade retention on the I/N reading gap would be negative; retention would widen the gap. The negative direction is only true for second-generation immigrants in the Netherlands. The cut-off to say Grade retention makes a substantial difference in the I/N reading gap is made at 5 percent because all changes below are minimal (Omorou et al., 2020). The result for second-generation immigrants in the Netherlands is substantial since it is above 5 %; the gap gets wider (Omorou et al., 2020). The only other substantial change is for first-generation immigrants in Germany, but instead of a negative direction, it is positive; after accounting for grade retention, the I/N reading gap for first-generation immigrants gets smaller. As expected, the influence of Grade retention on the I/N reading gap is smallest in Austria, but since

the I/N reading gap barely changes, the results do not support a larger role for Grade retention in accounting for the I/N gap in reading in Austria. Also, for second-generation students in Germany and first-generation students in the Netherlands, Grade retention does not account for the I/N gap in reading since the differences are too small.

> *H4: Repeating a grade in lower secondary education will explain more of the I/N gap in reading performance than repeating a grade in primary education.*

In models 4, the dummies for Timing grade retention are added to see if it matters if a student is retained in PE, SE, or both, in predicting reading performance. For the Netherlands, being retained in PE and PE/SE matters in predicting the scores; students who have not been retained in these phases have higher reading scores (17.723 for PE and 47.898 for both). Repeating in SE does not matter in predicting reading performance in the Netherlands. For Germany and Austria, all three categories (PE, SE, and both) matter in predicting reading scores; not being retained in one of the phases makes the reading scores higher (see Table 4.7 (DE) and Table 4.8 (AT) models 4, rows 5,6 & 7).
To say something about the influence of Timing of grade retention on the I/N reading gap, the dummy variables are added separately in different models. Models 5 are for repeating PE and models 6 for repeating SE. The changes in the I/N gap per country are presented in Table 4.9. Almost all percentages are negative, meaning the I/N gap closes slightly. Since all percentages are below 5 percent, the Timing of retention in PE and SE barely makes a difference in the reading gap between both immigrant groups and natives. Therefore, the hypothesis can be rejected.

Table 4.6. Regression reading performance Netherlands

Reading performance Netherlands							
	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	b(SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)
Fixed Part							
Intercept	479.527***	414.380***	468.537***	447.371***	399.378***	453.104***	464.785***
	(6.266)	(9.157)	(10.453)	(12.537)	(19.529)	(10.929)	(11.468)
Second generation (=0)		28.380***	23.658***	25.541***	23.435***	23.591***	23.615***
		(3.461)	(3.942)	(4.657)	(3.929)	(3.933)	(3.942)
First generation (=0)		43.011***	30.127***	29.727***	28.521***	28.948***	30.049***
		(5.938)	(6.637)	(8.973)	(4.308)	(6.626)	(6.638)
Grade retention (0=did repeat)				20.306***			
				(2.854)			

Repeat PE (0=yes)					17.723 *** (3.638)	17.127 *** (3.632)	
Repeat SE (0=yes)					5.793 (4.965)		3.956 (4.970)
Repeat PE & SE					47.898** (15.311)		
Gender (Boy = 0)			18.34 *** (1.896)	17.411 *** (2.095)	18.092 *** (1.892)	18.282 *** (1.892)	18.271 *** (1.898)
Language at home (0= other language)			12.594 ** (4.160)	14.107 ** (5.165)	12.389 ** (4.147)	12.553 ** (4.151)	12.604 ** (4.161)
Track (0= Academic)			-63.342 *** (3.000)	-61.335 *** (3.031)	-62.370 (4.868)	-62.782 *** (2.996)	-63.154 *** (3.009)
Middle education (=0)			-8.553 (4.882)	-2.041 (5.960)	-7.817 (4.868)	-8.179 (4.871)	-8.600 (4.882)
Higher education (=0)			-1.652 (4.747)	3.368 (5.824)	-.786 (4.735)	-1.126 (4.738)	-1.715 (4.748)
		Random Part					
Student level	4444.43** (92.58)	4296.24 *** (91.04)	3843.86 *** (82.77)	3889.72 *** (92.40)	3817.78 *** (82.23)	3825.83 *** (82.39)	3843.89 *** (82.78)
School level	5967.69** (694.678	5501.33** (642.801	3551.93 *** (427.50	3093.75 *** (388.48	3533.73** (425.380	3531.61** (425.086	3559.80** (428.512
ICC	0.5731	.05615	.4803	.4430	.4807	.500	.4808
Deviance	54117.30	52217.36	50183.8	41590.7	50133.89	50157.25	50178.18
Reference model		0	1	2	2	2	2
Fit-improvement		χ^2= 1899.95 DF= 2 P<.001	χ^2= 2033.50 DF= 5 P<.001	χ^2= 8593.15 DF= 1 P<.001	χ^2= 49.97 DF= 3 P<.001	χ^2= 26.61 DF= 1 P<.001	χ^2= 5.56 DF= 1 P<.001

Note. ***p<.001, **p<.01, *p<.05; SE = standard error. n. s=non-significant

Table 4.7. Regression reading performance Germany

	Reading performance Germany						
	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)
	Fixed Part						
Intercept	491.033 ** (5.209)	419.060 *** (7.374)	503.467 *** (15.607)	480.473 *** (15.515)	390.730 *** (22.165)	478.687 *** (16.538)	486.383 *** (16.144)
Second generation (=0)		18.230 *** (3.109)	-2.453 (3.793)	-2.626 (3.768)	-2.397 (3.769)	-2.273 (3.786)	-2.520 (3.786)

First generation (=0)		65.230 *** (4.625)	30.040 *** (6.078)	27.003 *** (6.091)	28.918 *** (6.043)	30.028 *** (6.068)	29.468 *** (6.069)
Grade retention (0=did repeat)				36.368 *** (3.025)			
Repeat PE (0=yes)					27.028 *** (5.623)	25.099 *** (5.639)	
Repeat SE (0=yes)					21.099 *** (4.699)		19.010 *** (4.710)
Repeat PE & SE (0=yes)					69.084 *** (14.259)		
Gender (Boy = 0)			13.824 *** (2.198)	11.793 *** (2.182)	13.193 *** (2.189)	13.974 *** (2.195)	13.274 *** (2.200)
Language at home (0= other language)			42.634 *** (4.118)	41.450 *** (4.120)	41.364 *** (4.100)	42.902 *** (4.111)	42.213 *** (4.112)
Track (0= Academic)			-56.066 *** (12.950)	-52.065 *** (12.772)	-55.033 *** (12.869)	-55.901 *** (12.927)	-55.234 *** (12.930)
Middle education (=0)			-15.792 *** (3.375)	-15.554 *** (3.354)	-15.936 *** (3.356)	-15.550 *** (3.370)	-16.070 *** (3.370)
Higher education (=0)			-13.277 *** (3.045)	-12.789 *** (3.026)	-13.114 *** (3.028)	-12.818 *** (3.041)	-13.621 *** (4.710)
		Random Part					
Student level	5221.580** (102.16)	5048.095*** (106.42)	4830.567*** (107.00)	4703.413*** (104.61)	4771.006*** (105.74)	4814.615*** (106.67)	4814.111*** (106.65)
School level	5803.64** (578.17)	5137.27*** (525.19)	4409.74*** (466.44)	3512.67*** (377.17)	4247.59*** (451.74)	4299.51*** (456.48)	4382.67*** (464.01)
ICC	0.4736	.5044	.4772	.4274	.4710	.4717	.4765
Deviance	62860.754	54375.555	49440.486	48724.174	49361.567	49415.460	49419.289
Reference model		0	1	2	2	2	2
Fit-improvement		χ^2= 8485.20 DF= 2 P<.001	χ^2= 4935.07 DF= 5 P<.001	χ^2= 716.31 DF= 1 P<.001	χ^2= 78.93 DF= 3 P<.001	χ^2= 25.03 DF= 1 P<.001	χ^2= 21.20 DF= 1 P<.001

Note. ***p<.001, **p<.01, *p<.05; SE = standard-error. n. s=non-significant

Table 4.8. Regression reading performance Austria

	Reading performance Austria						
	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)
	Fixed Part						
Intercept	470.044 ** (4.269)	383.274 *** (5.859)	456.807 *** (9.095)	429.730 *** (9.357)	372.80 *** (15.816)	427.220 *** (10.409)	444.163 *** (10.748)
Second generation (=0)		41.656 *** (2.630)	16.997 *** (3.326)	16.957 *** (3.327)	17.089 *** (3.315)	16.906* ** (3.319)	17.020 *** (3.326)
First generation (=0)		57.462 *** (3.542)	31.841 *** (4.147)	30.496 *** (4.162)	30.322 *** (4.143)	30.855 *** (4.141)	31.298 *** (4.153)
Grade retention (0=did repeat)				27.520 *** (2.686)			
Repeat PE (0=yes)					31.442 *** (5.270)	30.307 *** (5.270)	
Repeat SE (0=yes)					15.127** (5.790)		12.732* (5.800)
Repeat PE & SE (0=yes)					38.579 *** (9.939)		
Gender (Boy = 0)			15.140 *** (1.960)	14.355 *** (1.954)	14.718 *** (1.955)	15.174 *** (1.954)	15.105 *** (1.958)
Language (0= other)			34.910 *** (3.110)	34.431 *** (3.111)	34.432 *** (3.101)	34.962 *** (3.103)	34.666 *** (3.120)
Track (0= Academic)			-74.045 *** (5.782)	-64.113 *** (5.677)	-70.841 *** (5.783)	-73.241 *** (5.754)	-72.317 *** (5.831)
Middle education (=0)			-8.602 (4.751)	-6.104 (4.814)	-7.661 (4.739)	-7.796 (4.742)	-8.900 (4.751)
Higher education (=0)			-1.879 (4.727)	.159 (4.791)	-1.209 (4.715)	-1.184 (4.718)	-2.231 (4.728)
	Random Part						
Student level	4819.144* * (84.46)	4500.99 *** (79.46)	4303.70 *** (77.15)	4240.77 *** (76.41)	4274.20 *** (76.65)	4284.52 *** (76.81)	4301.60 *** (77.12)
School level	4862.65 ** (439.68)	4244.76 *** (386.10)	2593.89 *** (248.41)	2297.20 *** (224.28)	2512.83 *** (241.92)	2552.95 *** (244.97)	2586.09 *** (248.09)

ICC	0.5023	.4854	.3761	.3514	.3702	.3734	.3755
Deviance	77840.638	76309.368	73749.584	72791.769	73678.590	73711.442	73739.413
Reference model		0	1	2	2	2	2
Fit improvement		χ^2= 1531.27 DF= 2 P<.001	χ^2= 2559.78 DF= 5 P<.001	χ^2= 957.82 DF= 1 P<.001	χ^2= 70.99 DF= 3 P<.001	χ^2= 38.14 DF= 1 P<.001	χ^2= 10.17 DF= 1 P<.001

Note. ***p<.001, **p<.01, *p<.05; SE = standard-error. n. s=non-significant

Table 4.9. Changes in coefficients for Immigrant status when predicting Reading performance after adding Grade retention

		Gap second generation	**Gap First generation**
Grade retention	Netherlands	+ 7.96%	-1.33%
	Germany	n. s	-10.11%
	Austria	-0.24%	-4.22%
PE	Netherlands	-0.28%	-3.91%
	Germany	n.s	-0.04%
	Austria	-0.54%	-1.71%
SE	Netherlands	-0.18%	-0.39%
	Germany	n. s	-1.90%
	Austria	+0.14%	-1.71%

5. Discussion

Conclusion

This research aimed to discover whether and to what extent grade retention can account for the immigrant-native reading gap in the Netherlands, Germany, and Austria. An essential addition to the literature compared to previous research is that the distinction between first- and second-generation immigrants was made to see whether the effect is different for the two groups.

When looking at grade retention, the outcomes are as expected; first-generation immigrants are most likely to repeat a grade, followed by second-generation immigrants and, in the last place, natives. The odds for first-generation immigrants to repeat a grade are higher than natives. As mentioned in the theoretical chapter, an explanation might be that first-generation immigrants still have to learn the language and get used to the cultural differences (Klapproth et al. l., 2016). The countries differ in grade retention; in the Netherlands, the degree to which first-generation immigrants are more likely to repeat a grade than natives is the largest of the three countries.

For second-generation immigrants in all countries, the control variables account for the likelihood of grade retention instead of immigration status. The fact that they are retained more often than natives can be explained by the control variables. For example, second-generation immigrants are more often in a vocational track, which explains a higher likelihood of repeating a grade.

The results show that in all three countries, the order of reading performance is the same: natives have the highest reading performance, followed by second-generation immigrants, and in the last place first-generation immigrants. A difference is that in Germany, Second-generation immigrants do not statistically differ from native students in reading performance, while in the Netherlands and Austria, both immigrant groups have lower reading scores than natives.

For second-generation Dutch and first-generation German immigrants, grade retention is related to the I/N reading gap; in the Netherlands, the gap gets wider, while retention closes the gap in Germany. Although grade retention is related to reading performance, for first-generation Dutch and both generations of immigrants in Austria, grade retention does not account for the I/N reading gap since the percentage differences in the gap are too small (below 5%). For second-generation German immigrants, grade retention is not related to the I/N gap.

The timing of grade retention is related to reading performance; in most cases, it closes the gap slightly. However, the percentage differences are again not substantial (below 5%); therefore, the conclusion is that the retention timing does not account for the I/N reading gap.

Overall, grade retention was expected to widen the I/N reading gap. The results showed contradicting evidence since, in most cases, the gap closed even if the percentages were not substantial. As expected, Austria shows the least extreme results; this might be because Austria practices less grade retention than the other two countries (European Commission, 2018) (Table 4.8,

Appendix A). An explanation for the decreasing gap for first-generation immigrants in Germany might be that they must get used to a new culture and language; being retained might give them the time they need to catch up (Klapproth et al. l., 2016). An explanation of why for second-generation immigrants in the Netherlands, the gap widens after grade retention cannot be found in this research.

Limitations and recommendations

In contrast to prior research by Park and Sandefur (2010), there is a distinction between first-and second-generation immigrants in this research. This appeared to be necessary since the results for the two groups differed. Although, immigrants were not selected by ethnicity or country of origin. Previous research investigated students across Europe's reading and mathematics performance and discovered that Turkish immigrant students fall most behind compared to immigrants of other origins (Arikan, Van de Vijver, & Yagmur, 2017). Because this research only includes three European countries, the sample sizes would have been too small for statistical analyses when focusing on one specific group of immigrants. For this research, the decision was made to compare more homogenous countries in terms of school policies to better isolate the role of grade retention in reading performance. It would have been preferred to test whether the differences between countries were significant, which was not done in this research. Future research should focus on getting enough data on specific immigration groups to be able to compare them and test the differences on significance. Also, in some schools, there were only native students or only a few immigrant students; this could have impacted the results. To take this factor into account, multi-level analyses were used to consider the influence of the school level. Still, it is essential for future research that a large amount of data from all OECD countries is available to have a good representation of all groups in the different schools. Also, this research is not based on longitudinal data, which might have given a complete image of how the performance gap changed in the countries over time. With more extensive data in the future, it is interesting to do a longitudinal study on the role of grade retention on the I/N performance gap.

Since some students are more likely to be retained levels (Hauser, Pager, & Simmons, 2001), even with the same results as others, it is essential to pay attention to whether retention is assessment-based. When it is less assessment-based, there is more room for the opinion of teachers, which might increase the influence of parental SES in the decision. In the Netherlands, in 98% of the cases, assessments determine whether a student gets promoted to the next grade or must repeat the same year. With an average in OECD countries of 77%, this is very high. In Germany, this is 95.8%, while in Austria, it is lower than average at only 12% (OECD, 2014). Future research can focus on the role of assessment-based retention on the likelihood of immigrant students repeating a grade in relation to school performance.

A factor that is also not considered in this research is immigration policy. Obligated language courses might be of influence the reading performance of children. Also, other cultural courses might impact first-generation students who will make them get used to the new culture and school system; this might influence the number of immigrants retained. The age at which first-immigrant students came to the test country might be an explanation for being retained in PE or SE; when they are at the age where they start in SE, they will not be retained in PE. These factors should be considered in future research.

The track is of meaningful influence on students' school performance and, therefore, a control variable in this research. It should be considered that respondents of this questionnaire are all around 15 years old. A distinction between vocational and academic education is not applicable in all cases; in PE, there is no difference in tracks. Therefore, there are few categories for academic tracks in the Netherlands, Germany, and Austria data.

To further investigate the negative influence of retention on the I/N reading gap in the Netherlands and the positive influence in Germany, it is interesting for future research to use a more qualitative approach. When there is more information on what factors make retention positively or negatively impact performance, school policies can be adjusted to get the maximum benefits out of it to close the gap.

Biography

Arikan, S., Van de Vijver, F. J., & Yagmur, K. (2017). PISA mathematics and reading performance differences of mainstream European and Turkish immigrant students. Educational Assessment, Evaluation and Accountability, 29(3), 229-246.

Austin, P., Goel, V., & van Walraven, C. (2001). An introduction to multilevel regression models. Canadian Journal of Public Health / Revue Canadienne De Sante'e Publique, 92(2), 150–154.

Blaine, B. E. (2018). Winsorizing. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation, 1817.

Burger, K. (2016). Intergenerational transmission of education in Europe: Do more comprehensive education systems reduce social gradients in student achievement? Research in Social Stratification and Mobility, 44, 54–67. https://doi.org/10.1016/j.rssm.2016.02.002

d'Aiglepierre, R., David, A., Levionnois, C., Spielvogel, G., Tuccio, M., & Vickstrom, E. (2020). A global profile of emigrants to OECD countries: Younger and more skilled migrants from more diverse countries.

Dronkers, J., & Kornder, N. (2015). Can gender differences in educational performance of 15-year-old migrant pupils be explained by societal gender equality in origin and destination countries?. Compare: A Journal of Comparative and International Education, 45(4), 610-634.

Dronkers, J., & Korthals, R. (2016). Tracking, school entrance requirements and the educational performance of migrant students. Education systems and inequalities: International comparisons, 185-205.

Eldering, L., & Kloprogge, J. (Eds.). (1989). Different cultures same school: Ethnic minority children in Europe. Amsterdam: Swets & Zeitlinger B.V.

European Commission. (2005). *Key data on education in Europe 2005*. Luxembourg: Office for Official Publications for the European Communities.

European Commission/EACEA/Eurydice, 2018. The Structure of the European Education Systems 2018/19: Schematic Diagrams. Eurydice Facts and Figures. Luxembourg: Publications Office of the European Union.

Field, A. (2009). Discovering Statistics Using SPSS, 3rd Edition (Introducing Statistical Methods) (3rd editie). SAGE Publications Ltd.

Fine, J. G., & Davis, J. M. (2003). Grade retention and enrollment in post-secondary education. Journal of School Psychology, 41, 401–411.

Frey, N. (2005). Retention, social promotion, and academic redshirting: what do we know and need to know? Remedial and Special Education, 26, 332–346.

Goos, M., Schreier, B. M., Knipprath, H. M. E., De Fraine, B., Van Damme, J., & Trautwein, U. (2013). How can cross-country differences in the practice of grade retention be explained? A closer look at national educational policy factors. Comparative Education Review, 57(1), 54-84.

Hauser, R. M. (1999). Should we end social promotion? Truth and consequences. Working Paper, Center for Demography and Ecology, University of Wisconsin-Madison.

Herzog-Punzenberger, B. (2003). Ethnic segregation in school and labor market – 40 year legacy of austrian guestworker policy. International Migration Review, 37, 1120–1144.

Houtte*, M. V. (2004). Why boys achieve less at school than girls: The difference between boys' and girls' academic culture. Educational studies, 30(2), 159-173.

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). Multilevel analysis: Techniques and applications. Routledge.

Klapproth, F., Schaltz, P., Brunner, M., Keller, U., Fischbach, A., Ugen, S., & Martin, R. (2016). Short-term and medium-term effects of grade retention in secondary school on academic achievement and psychosocial outcome variables. Learning and Individual Differences, 50, 182-194.

Levels, M., Dronkers, J., & Kraaykamp, G. (2008). Immigrant children's educational achievement in western countries: origin, destination, and community effects on mathematical performance. American Sociological Review, 73(5), 835-853.

McMahon, W. W., & Oketch, M. (2013). Education's effects on individual life chances and on development: An overview. British Journal of Educational Studies, 61(1), 79-107.

Meisels, S. J., & Liaw, F.-R. (1993). Failure in grade: do retained students catch up? Journal of Educational Research, 87, 69–77.

Mislevy, R. J., Beaton, A. E., Kaplan, B. & Sheehan, K. M. (1991). Estimating population characteristics from sparse matrix samples of item responses. Journal of Educational Measurement, 29, 133-161.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). PIRLS 2001 international report. Boston, MA: Boston College.

OECD. (2014) EDUCATION POLICY OUTLOOK NETHERLANDS

Omorou, A. Y., Manneville, F., Langlois, J., Legrand, K., Böhme, P., Muller, L., Guillemin, F., Briançon, S., Lecomte, E., Böhme, P., Briançon, S., Lavenne, R. de, Gailliard, C., Langlois, J., Lecomte, E., Legrand, K., Muller, L., Omorou, A. Y., Pourcher, C., … Toussaint, B. (2020). Physical activity rather than sedentary behaviour is socially determined in French adolescents with overweight and obesity. Preventive Medicine, 134, 106043. https://doi.org/10.1016/j.ypmed.2020.106043

Organisation for Economic Co-operation and Development. (2020). Education at a glance 2020 : oecd indicators (Ser. Education at a glance). OECD Publishing. https://doi.org/10.1787/69096873-en

Park, H., & Sandefur, G. (2010). School performance gaps between immigrant and native students in Europe: The role of grade. In Quality and inequality of education (pp. 113-136). Springer, Dordrecht.

Peña-López, I. (2019). PISA 2018 results. Where all students can succeed.

Rindermann, H., & Thompson, J. (2016). The cognitive competences of immigrant and native students across the world: An analysis of gaps, possible causes and impact. Journal of biosocial science, 48(1), 66-93.

Riphahn, R. T. (2003). Cohort effects in the educational performance of second generation immigrants in Germany: An analysis of census data. Journal of Population Economics, 16, 711–737.

Rodríguez, S., Valle, A., Gironelli, L. M., Guerrero, E., Regueiro, B., & Estévez, I. (2020). Performance and well-being of native and immigrant students. Comparative analysis based on PISA 2018. Journal of Adolescence, 85, 96-105.

Rumberger, R. W., & Larson, K. A. (1998). Student mobility and the increased risk of high school dropout. American Journal of Education, 107, 1–35.

Silberglitt, B., Jimerson, S. R., Burns, M. K., & Appleton, J. J. (2006). Does the timing of grade retention make a difference? Examining the effects of early versus later retention. School Psychology Review, 35, 134–141.

Schleicher, A. (2006). Where immigrant students succeed: a comparative review of performance and engagement in PISA 2003: © OECD 2006. Intercultural education, 17(5), 507-516.

Schmid, C. L. (2001). Educational achievement, language-minority students, and the new second generation. Sociology of Education, 74, 71–87.

Schnell, P., & Azzolini, D. (2015). The academic achievements of immigrant youths in new destination countries: Evidence from southern Europe. Migration Studies, 3(2), 217-240.

UNESCO Institute for Statistics. (2012). International standard classification of education: ISCED 2011. Int. Stand. Classif. Educ.(ISCED) 2011.

VandeCandeLaere, M. (2015). De effectiviteit van doorkleuteren. Basis, (9), 23-25.

Appendix A: Descriptives

	Netherlands		Germany		Austria	
	Count	%	Count	%	Count	%
Native	3951	85.7%	3677	77.8%	5303	79%
Grade retention						
Did not repeat	2805	83.6%	3018	83.5%	4697	89.6%
Did repeat	551	13.9%	598	16.5%	546	10.4%
Timing retention						
Never	2973	86.2%	2996	89.2%	4781	95.2%
Repeat PE	285	8.8%	150	4.5%	120	2.4%
Repeat SE	150	4.6%	199	5.9%	83	1.7%
Both	13	0.4%	15	0.4%	36	0.7%
Gender						
Girl	1938	49.1%	1723	46.9%	2593	48.9%
Boy	2013	50.9%	1954	53.1%	2710	51.1%
Track						
Vocational	335	76.8%	3639	99.0%	3803	71.7%
Academic	916	23.2%	38	1.0%	1500	28.3%
Language at home						
Test country	3834	97.5%	3511	95.5%	5061	95.5%
Other	97	2.5%	165	4.5%	236	4.5%
Parental education						
Lower	130	3.3%	609	17.3%	71	1.4%
Middle	1078	27.6%	883	25.1%	2153	41.3%
Higher	2696	69.1%	2022	57.5%	2986	57.3%
Second-generation	518	11.2%	751	15.9%	933	13.9%
Grade retention						
Did not repeat	271	74.2%	551	74.4%	726	79.5%
Did repeat	94	25.8%	190	25.6%	187	20.5%
Timing retention						
Never	267	77.6%	546	83.0%	746	91.1%
Repeat PE	44	12.8%	48	7.3%	30	3.7%
Repeat SE	28	8.1%	55	8.4%	35	4.3%
Both	5	1.0%	9	1.4%	8	1.0%
Gender						
Girl	262	50.6%	351	46.7%	467	50.1%
Boy	256	49.4%	400	53.3%	466	49.9%
Track						
Vocational	470	90.7%	746	99.3%	715	76.6%
Academic	48	9.3%	5	0.7%	218	23.4%
Language at home						
Test country	250	48.4%	357	47.5%	255	27.3%
Other	267	51.6%	394	52.5%	678	72.7%
Parental education						
Lower	55	11.6%	149	25.4%	121	13.7%
Middle	157	33.2%	167	28.4%	348	39.4%
Higher	261	55.2%	271	46.2%	414	46.9%
First-generation	143	3.1%	299	6.3%	474	7.1%
Grade retention						
Did not repeat	42	57.5%	195	67.9%	319	69.8%
Did repeat	31	42.5%	92	32.1%	138	30.2%
Timing retention						

Never	42	61.8%	189	77.8%	321	80.3%
Repeat PE	18	26.5%	17	7.0%	30	7.5%
Repeat SE	6	8.8%	32	13.2%	40	10.0%
Both	2	2.9%	5	2.1%	9	2.3%
Gender						
Girl	64	44.8%	138	46.2%	227	47.9%
Boy	79	55.2%	161	53.8%	247	52.1%
Track						
Vocational	130	90.9%	299	100%	362	76.4%
Academic	13	9.1%	-	-	112	23.6%
Language at home						
Test country	35	25.2%	38	12.8%	111	23.4%
Other	104	74.8%	258	87.2%	363	76.6%
Parental education						
Lower	18	14.2%	44	22.9%	44	10.1%
Middle	27	21.3%	50	22.9%	151	34.7%
Higher	82	64.6%	124	56.9%	240	55.2%

Appendix B: Correlations

Table 1. Correlations Netherlands
Netherlands

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. Reading performance	—						
2. Immigrant status	-.264***	—					
3. Grade retention	-.254***	.114***	—				
4. Gender (Girl = 0)	-.141***	.004 n. s	.063***	—			
5. Language at home (test country = 0)	-.239***	.628***	.092***	.018 n. s	—		
6. Track (Vocational= 0)	.563***	-.113***	-.190***	-.039**	-.102***	—	
7. Education parents	.164***	-.110***	-.071***	.026 n. s	-.094***	.197***	—

Note. ***p<.001, **p<.01, *p<.05, n.s=non-significant; All correlations are calculated with Pearsons r.

Table 2. Correlations Germany
Germany

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. Reading performance	—						
2. Immigrant status	-.305***	—					
3. Grade retention	-.316***	.120***	—				
4. Gender (Girl = 0)	-.124***	.003 n. s	.084***	—			
5. Language at home (test country = 0)	-.339***	.662***	.134***	.000 n. s	—		
6. Track (Vocational= 0)	.117***	-.028 n. s	-.047***	-.011 n. s	-.027 n. s	—	
7. Education parents	,204***	-.056***	-.096***	-.017 n. s	-.064***	.043**	—

Note. ***p<.001, **p<.01, *p<.05, n. s=non-significant; All correlations are calculated with Pearsons r.

Table 3. Correltations Austria

Austria

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. Reading performance	—						
2. Immigrant status	-.275***	—					
3. Grade retention	-.238***	.171***	—				
4. Gender (Girl = 0)	-.157***	.000n. s	.046***	—			
5. Language at home (test country = 0)	-.302***	.157***	.157***	.016 n. s	—		
6. Track (Vocational= 0)	.393***	-.076***	-.076***	-.113***	-.069***	—	
7. Education parents	.129***	-.030*	-.030*	.009 n. s	-.144***	.236***	—

Note. ***p<.001, **p<.01, *p<.05, n. s=non-significant; All correlations are calculated with Pearsons r.

Syntax SPSS

```
* Encoding: UTF-8.
* Encoding: UTF-8.


GM reading performance NL variable

DATASET ACTIVATE DataSet1.
USE ALL.
COMPUTE filter_$=(CNTRYID = 528).
VARIABLE LABELS filter_$ 'CNTRYID = 528 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

COMPUTE NLreadGM=PVReading- 479.822304.
EXECUTE.

GM reading performance DE variable

USE ALL.
COMPUTE filter_$=(CNTRYID = 276).
VARIABLE LABELS filter_$ 'CNTRYID = 276 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

COMPUTE DEreadGM=PVReading- 479.822304.
EXECUTE.

GM reading performance AUT variable

USE ALL.
COMPUTE filter_$=(CNTRYID = 40).
VARIABLE LABELS filter_$ 'CNTRYID = 40 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

COMPUTE AUTreadGM=PVReading- 479.822304.
EXECUTE.

Univariate Outliers checken GM reading *per land filter aangezet
```

```
EXAMINE VARIABLES=NLreadGM
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES EXTREME
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.


EXAMINE VARIABLES=DEreadGM
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES EXTREME
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.


EXAMINE VARIABLES=AUTreadGM
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES EXTREME
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Multi variate outliers GM reading and immigration status (per land filter aangezet

```
EXAMINE VARIABLES=NLreadGM BY immigrationstatusnew
  /PLOT=BOXPLOT
  /STATISTICS=NONE
  /NOTOTAL.

EXAMINE VARIABLES=DEreadGM BY immigrationstatusnew
  /PLOT=BOXPLOT
  /STATISTICS=NONE
  /NOTOTAL.


EXAMINE VARIABLES=AUTreadGM BY immigrationstatusnew
  /PLOT=BOXPLOT
  /STATISTICS=NONE
  /NOTOTAL.
```

Descriptives categorical variables

```
SORT CASES  BY IMMIG.
SPLIT FILE LAYERED BY IMMIG.
```

```
DATASET ACTIVATE DataSet1.
FREQUENCIES VARIABLES=REPEAT Track Language educationparents
  /ORDER=ANALYSIS.
```

Correlaties Nederland, duitsland en Oostenrijk (filter per land aangezet)

```
CORRELATIONS
  /VARIABLES=PVReading IMMIG REPEAT ST004D01T ST022Q01TA Track
educationparents
  /PRINT=TWOTAIL NOSIG FULL
  /MISSING=PAIRWISE.
```

Model 0 reading performance for culculating ICC (filter per country)

```
MIXED PVReading
  /CRITERIA=DFMETHOD(SATTERTHWAITE) CIN(95) MXITER(100) MXSTEP(10)
SCORING(1)
    SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0,
ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  /FIXED=| SSTYPE(3)
  /METHOD=REML
  /PRINT=DESCRIPTIVES G  SOLUTION TESTCOV
  /RANDOM=INTERCEPT | SUBJECT(CNTSCHID) COVTYPE(VC).
```

Dummies immigration maken

```
comp native= (IMMIG=1).
comp second= (IMMIG=2).
comp first= (IMMIG=3).
```

Dummies education level parents

```
comp lower= (educationparents=1).
comp middle= (educationparents=2).
comp upper= (educationparents=3).
```

Model 0 dependent variable grade retention

```
*Generalized Linear Mixed Models.
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=CNTSCHID
  /FIELDS TARGET=REPEAT TRIALS=NONE OFFSET=NONE
  /TARGET_OPTIONS DISTRIBUTION=BINOMIAL LINK=LOGIT
  /FIXED USE_INTERCEPT=TRUE
```

```
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=DESCENDING
    MAX_ITERATIONS=100 CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL
COVB=MODEL PCONVERGE=0.000001(ABSOLUTE)
    SCORING=0 SINGULAR=0.000000000001
  /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Model 1 dependent variable grade retention

```
*Generalized Linear Mixed Models.
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=CNTSCHID
  /FIELDS TARGET=REPEAT TRIALS=NONE OFFSET=NONE
  /TARGET_OPTIONS DISTRIBUTION=BINOMIAL LINK=LOGIT
  /FIXED  EFFECTS=second first USE_INTERCEPT=TRUE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=DESCENDING
    MAX_ITERATIONS=100 CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL
COVB=MODEL PCONVERGE=0.000001(ABSOLUTE)
```

```
   SCORING=0 SINGULAR=0.000000000001
  /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Model 2 dependent variable grade retention

*Generalized Linear Mixed Models.

```
GENLINMIXED
  /DATA_STRUCTURE SUBJECTS=CNTSCHID
  /FIELDS TARGET=REPEAT TRIALS=NONE OFFSET=NONE
  /TARGET_OPTIONS DISTRIBUTION=BINOMIAL LINK=LOGIT
  /FIXED  EFFECTS=second first Gender Language Track middle upper
USE_INTERCEPT=TRUE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /RANDOM USE_INTERCEPT=TRUE
COVARIANCE_TYPE=VARIANCE_COMPONENTS SOLUTION=FALSE
  /BUILD_OPTIONS TARGET_CATEGORY_ORDER=DESCENDING
INPUTS_CATEGORY_ORDER=DESCENDING
    MAX_ITERATIONS=100 CONFIDENCE_LEVEL=95 DF_METHOD=RESIDUAL
COVB=MODEL PCONVERGE=0.000001(ABSOLUTE)
    SCORING=0 SINGULAR=0.000000000001
  /EMMEANS_OPTIONS SCALE=ORIGINAL PADJUST=LSD.
```

Nieuwe variabele timing retention

```
COMPUTE timingretention=$SYSMIS.
  IF ((Repeat1=0) AND (Repeat2= 0)) timingretention = 0.
  IF ((Repeat1=1) AND (Repeat2= 0)) timingretention = 1.
  IF ((Repeat1=0) AND (Repeat2= 1)) timingretention = 2.
  IF ((Repeat1=1) AND (Repeat2= 1)) timingretention = 3.
EXECUTE.
```

Dummies timing retention

```
comp never= (timingretention=0).
comp PE= (timingretention=1).
comp SE= (timingretention=2).
comp both= (timingretention=3).
```

Model 1 dependent variable Reading

```
MIXED PVReading BY second first
  /CRITERIA=DFMETHOD(SATTERTHWAITE) CIN(95) MXITER(100) MXSTEP(10)
SCORING(1)
    SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0,
ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  /FIXED=second first | SSTYPE(3)
  /METHOD=REML
  /PRINT=DESCRIPTIVES G  SOLUTION TESTCOV
  /RANDOM=INTERCEPT | SUBJECT(CNTSCHID) COVTYPE(VC).
```

model 2 dependent variable reading

```
MIXED PVReading BY second first Gender Language Track middle upper
  /CRITERIA=DFMETHOD(SATTERTHWAITE) CIN(95) MXITER(100) MXSTEP(10)
SCORING(1)
    SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0,
ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  /FIXED=second first Gender Language Track middle upper | SSTYPE(3)
  /METHOD=REML
  /PRINT=DESCRIPTIVES G  SOLUTION TESTCOV
  /RANDOM=INTERCEPT | SUBJECT(CNTSCHID) COVTYPE(VC).
```

Model 3 dependent variable reading

```
MIXED PVReading BY second first REPEAT Gender Language Track middle upper
  /CRITERIA=DFMETHOD(SATTERTHWAITE) CIN(95) MXITER(100) MXSTEP(10)
SCORING(1)
    SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0,
ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  /FIXED=second first REPEAT Gender Language Track middle upper | SSTYPE(3)
  /METHOD=REML
  /PRINT=DESCRIPTIVES G  SOLUTION TESTCOV
  /RANDOM=INTERCEPT | SUBJECT(CNTSCHID) COVTYPE(VC).
```

Model 4 dependent variable reading

```
DATASET ACTIVATE DataSet1.
MIXED PVReading BY second first PE SE both Gender Language Track middle upper
```

```
  /CRITERIA=DFMETHOD(SATTERTHWAITE) CIN(95) MXITER(100) MXSTEP(10)
SCORING(1)
    SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0,
ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  /FIXED=second first PE SE both Gender Language Track middle upper | SSTYPE(3)
  /METHOD=REML
  /PRINT=DESCRIPTIVES G  SOLUTION TESTCOV
  /RANDOM=INTERCEPT | SUBJECT(CNTSCHID) COVTYPE(VC).
```

Model 5 repeat PE

```
MIXED PVReading BY second first PE Gender Language Track middle upper
  /CRITERIA=DFMETHOD(SATTERTHWAITE) CIN(95) MXITER(100) MXSTEP(10)
SCORING(1)
    SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0,
ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  /FIXED=second first PE Gender Language Track middle upper | SSTYPE(3)
  /METHOD=REML
  /PRINT=DESCRIPTIVES G  SOLUTION TESTCOV
  /RANDOM=INTERCEPT | SUBJECT(CNTSCHID) COVTYPE(VC).
```

Model 6 repeat SE

```
MIXED PVReading BY second first SE Gender Language Track middle upper
  /CRITERIA=DFMETHOD(SATTERTHWAITE) CIN(95) MXITER(100) MXSTEP(10)
SCORING(1)
    SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0,
ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  /FIXED=second first Gender Language Track middle upper SE | SSTYPE(3)
  /METHOD=REML
  /PRINT=DESCRIPTIVES G  SOLUTION TESTCOV
  /RANDOM=INTERCEPT | SUBJECT(CNTSCHID) COVTYPE(VC).
```

Checklist ethical and privacy aspects of research

CHECKLIST ETHICAL AND PRIVACY ASPECTS OF RESEARCH

INSTRUCTION

This checklist should be completed for every research study that is conducted at the Department of Public Administration and Sociology (DPAS). This checklist should be completed *before* commencing with data collection or approaching participants. Students can complete this checklist with help of their supervisor.

This checklist is a mandatory part of the empirical master's thesis and has to be uploaded along with the research proposal.

The guideline for ethical aspects of research of the Dutch Sociological Association (NSV) can be found on their website (http://www.nsv-sociologie.nl/?page_id=17). If you have doubts about ethical or privacy aspects of your research study, discuss and resolve the matter with your EUR supervisor. If needed and if advised to do so by your supervisor, you can also consult Dr. Jennifer A. Holland, coordinator of the Sociology Master's Thesis program.

PART I: GENERAL INFORMATION

Project title: Master thesis families and inequalities

Name, email of student: Jildou Tromp, 586758jt@student.eur.nl

Name, email of supervisor: Sanneke de la Rie, delarie@essb.eur.nl

Start date and duration: 17-01-2022, 5 months

Is the research study conducted within DPAS YES - ~~NO~~

If 'NO': at or for what institute or organization will the study be conducted?
(e.g. internship organization)

PART II: HUMAN SUBJECTS

1. Does your research involve human participants. YES - ~~NO~~

If 'NO': skip to part V.

If 'YES': does the study involve medical or physical research? ~~YES~~ - NO

Research that falls under the Medical Research Involving Human Subjects Act (WMO) must first be submitted to an accredited medical research ethics committee or the Central Committee on Research Involving Human Subjects (CCMO).

2. Does your research involve field observations without manipulations that will not involve identification of participants. ~~YES~~ - NO

If 'YES': skip to part IV.

3. Research involving completely anonymous data files (secondary data that has been anonymized by someone else). YES - ~~NO~~

If 'YES': skip to part IV.

PART III: PARTICIPANTS

1. Will information about the nature of the study and about what participants can expect during the study be withheld from them? ~~YES~~ - NO
2. Will any of the participants not be asked for verbal or written 'informed consent,' whereby they agree to participate in the study? ~~YES~~ - NO

3. Will information about the possibility to discontinue the participation at any time be withheld from participants? ~~YES~~ - NO

4. Will the study involve actively deceiving the participants? ~~YES~~ - NO

Note: almost all research studies involve some kind of deception of participants. Try to think about what types of deception are ethical or non-ethical (e.g. purpose of the study is not told, coercion is exerted on participants, giving participants the feeling that they harm other people by making certain decisions, etc.).

Does the study involve the risk of causing psychological stress or negative emotions beyond those normally encountered by participants? ~~YES~~ - NO

Will information be collected about special categories of data, as defined by the GDPR (e.g. racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data for the purpose of uniquely identifying a person, data concerning mental or physical health, data concerning a person's sex life or sexual orientation)? YES - ~~NO~~

Will the study involve the participation of minors (<18 years old) or other groups that cannot give consent? YES - ~~NO~~

Is the health and/or safety of participants at risk during the study? ~~YES~~ - NO

Can participants be identified by the study results or can the confidentiality of the participants' identity not be ensured? ~~YES~~ - NO

Are there any other possible ethical issues with regard to this study? ~~YES~~ - NO

If you have answered 'YES' to any of the previous questions, please indicate below why this issue is unavoidable in this study.

Because the research it based on existing data of PISA, I can say that they already took care of this matter when conducting the research

What safeguards are taken to relieve possible adverse consequences of these issues (e.g., informing participants about the study afterwards, extra safety regulations, etc.).

_____
_____
_____
_____

Are there any unintended circumstances in the study that can cause harm or have negative (emotional) consequences to the participants? Indicate what possible circumstances this could be.

_____
_____
_____
_____

Please attach your informed consent form in Appendix I, if applicable.

Continue to part IV.

PART IV: SAMPLE

Where will you collect or obtain your data?

PISA 2018

Note: indicate for separate data sources.

What is the (anticipated) size of your sample?
the Netherlands (N=4765), Germany (N=5451) , and Austria (N=6802) From PISA 2018
Note: indicate for separate data sources.

What is the size of the population from which you will sample?
The whole population of 15 year old students in the Netherlands, Germany and Austria in 2018

Note: indicate for separate data sources.

Continue to part V.

Part V: Data storage and backup

Where and when will you store your data in the short term, after acquisition?

The SPSS data-file will be saved in a secured folder in the cloud

Note: indicate for separate data sources, for instance for paper-and pencil test data, and for digital data files.

Who is responsible for the immediate day-to-day management, storage and backup of the data arising from your research?
OECD was responsible for the data collection. I (Jildou Tromp) am responsible for the daily storage and backup of the results of this study

How (frequently) will you back-up your research data for short-term data security?
Daily to weekly

In case of collecting personal data how will you anonymize the data?
N.A
The PISA data is made anonymous by OECD, I myself also don't know any personal data that can be linked to a person

Note: It is advisable to keep directly identifying personal details separated from the rest of the data. Personal details are then replaced by a key/ code. Only the code is part of the database with data and the list of respondents/research subjects is kept separate.

PART VI: SIGNATURE

Please note that it is your responsibility to follow the ethical guidelines in the conduct of your study. This includes providing information to participants about the study and ensuring confidentiality in storage and use of personal data. Treat participants respectfully, be on time at appointments, call participants when they have signed up for your study and fulfil promises made to participants.

Furthermore, it is your responsibility that data are authentic, of high quality and properly stored. The principle is always that the supervisor (or strictly speaking the Erasmus University Rotterdam) remains owner of the data, and that the student should therefore hand over all data to the supervisor.

Hereby I declare that the study will be conducted in accordance with the ethical guidelines of the Department of Public Administration and Sociology at Erasmus University Rotterdam. I have answered the questions truthfully.

Name student: Jildou Tromp Name (EUR) supervisor: Sanneke de la Rie

Date: Date: