VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Finnish Journal of Social Research
Advance access (2026), pp. 1–21
https://doi.org/10.51815/fjsr.144756

# Does the language spoken at home matter for the education and sense of belonging of the children of immigrants?

Elina Kilpi-Jakonen and Jenni Alisaari

## Abstract

Educational disadvantages of children of immigrants have been linked to speaking a language other than that of school instruction at home. Yet both theory and previous research suggest that the continued use of the language of (parental) origin can be beneficial for children of immigrants. We used data from a subset of countries in the 2018 Programme of International Student Assessment (PISA) and analysed reading test scores and sense of belonging at school, paying attention to appropriate measurement of language use. Our results suggest that, when analysed appropriately, there are a number of countries in which language spoken at home is not associated with either learning outcomes or sense of belonging. However, in a few countries, switching the home language is associated with an increased sense of belonging and higher reading scores than maintaining the origin language. We discuss these results with reference to ethnic boundary making and responding to the needs of linguistically diverse students.

**Keywords**: language spoken at home; children of immigrants; learning outcomes; belonging

Kilpi-Jakonen (University of Turku) and Alisaari (University of Eastern Finland). Corresponding author's e-mail: elina.kilpi-jakonen@utu.fi  ISSN 2736-9749 (print), 2814-5038 (online)

## Introduction

In classrooms around the world, there are more students with migrant origins from varied linguistic backgrounds. There is also growing concern regarding their learning outcomes, wellbeing, and sense of belonging. Studies have shown that migrant-origin students have a lower sense of belonging than majority students (Borgonovi, 2018). In several studies, the disadvantages that children of immigrants face have been specifically linked to speaking a language other than the language of instruction at home (e.g., Dustmann et al., 2012; Schnepf, 2007), with the implication that migrant families should switch the language spoken at home. This also comes through in OECD reports (e.g., 2019c, 2023). We argue that these conclusions are partly due to false comparison groups in statistical models, and that to draw this conclusion, it is not sufficient to merely control for whether the student mainly speaks a language that is different from that of instruction.

In this article, we examine whether language switching versus maintaining among migrant families is associated with differential outcomes and whether this varies across countries. We aim to demonstrate how different ways of specifying language spoken at home — and being more precise about the groups that are being compared — change the associations under consideration. By doing this across a range of countries, we aim to show that the language situation in terms of national languages and linguistic minorities, including regional dialects, as well as the composition of the migrant population also influence how the different specifications matter.

We use data from the 2018 Programme of International Student Assessment (PISA), which is a test and survey of 15-year-old students (OECD, 2019a). We analyse the association between language spoken at home (with different specifications) for two dependent variables: reading test scores and sense of belonging at school, with additional robustness checks using two other competence measures (maths and science). Academic achievement is often seen as primary in the educational sphere and which also has long-term implications for socio-economic outcomes. However, sense of belonging is a basic human need (Lambert et al., 2013) and it also has important implications in educational contexts (Allen et al., 2018; Osterman, 2000). Adolescence is a period when sense of belonging plays a particularly important role (Allen et al., 2018). Sense of belonging can include feelings of attachment, as well as positive social relations comprising mutual acceptance and respect (St-Amand et al., 2017). "Sense of belonging at school" or "school belonging" are specific applications of this concept within the educational setting (see e.g., DeNicolo et al., 2017). They pertain to students' feelings of being respected, supported, and included in their school environment (Allen et al., 2018; DeNicolo et al., 2017; Osterman, 2000). A strong feeling of being part of a school or a specific group may also positively impact learning outcomes (Celeste et al., 2019). However, minoritized groups, particularly students with migrant backgrounds or differing home languages, tend to report diminished belonging (Borgonovi, 2018).

In the next section we explain our argument on why different ways of operationalising language spoken at home or the different comparisons that can be made may matter for the results. We then review the theoretical expectations and previous evidence on language use at home and children's outcomes, particularly related to our two dimensions of interest (learning and sense of belonging). After this, we discuss potential reasons for country differences. Our study aims to contribute to research by highlighting the problems of a simplistic use of a variable measuring the language spoken at home when examining the educational outcomes and sense of belonging of children of immigrants vis-à-vis their majority peers, particularly in cross-national research. Although our empirical models cannot establish causality, we also aim to contribute to policy discussions on the ways in which students with migration backgrounds should be supported.

**Different groups of language use at home from a cross-national perspective**

There is major disagreement on whether continuing to speak one's language of origin is beneficial or detrimental for other integration outcomes. The question has been studied across various disciplines and thus from a variety of theoretical approaches. Before reviewing the theoretical expectations and findings from the previous research, we clarify the different distinctions that can be made based on the intersection of language spoken at home and migration background.

In many countries, there are either linguistic minorities (of non-migrant origin) or migrant groups arriving from countries that share a common language with the destination country — in some cases both. This is conceptualised in Figure 1. In the upper part of the figure, we show that among all children, those who do not speak the language of the school at home include both speakers of national or minority languages as well as those who speak a foreign language (i.e. children of immigrants speaking the language of their parent's country of origin). The first group includes both children who cannot go to school in a national minority language (many indigenous languages and official dialects) as well as children who, in multilingual education systems, attend school in the other national language despite schools also being available in their own language (such as French-speaking Canadians attending English-speaking schools). In the lower part of the figure, we show that among migrant-origin children, those who speak the language of the school at home often also include two groups: those with origins in countries that share a common language, i.e. where families may already have been speaking the language of the country of destination, and those with origins in countries that do not share a common language but where families have switched to speaking the language of the country of destination.

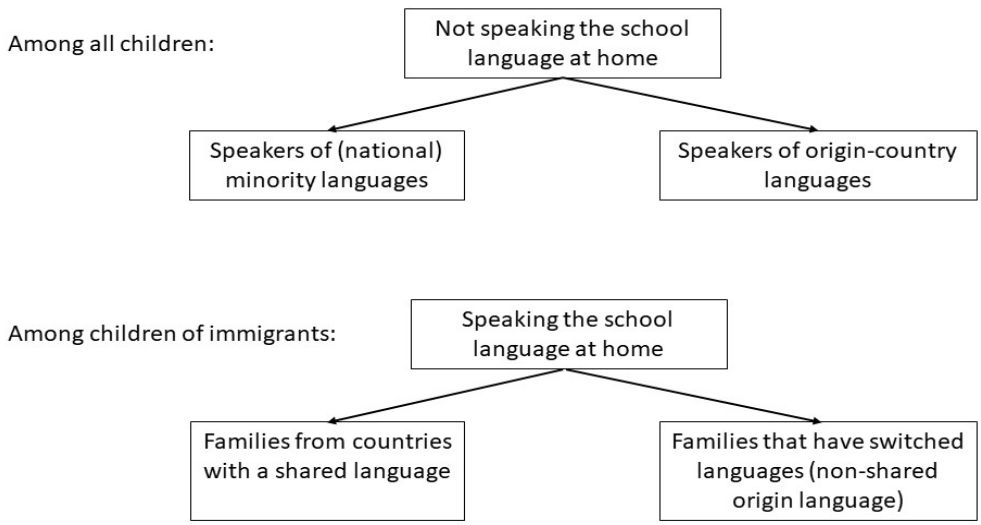


*Figure 1. Conceptualisation of differences in groups based on language spoken at home*

Except for research on specific origin groups, either one or both of these distinctions is often conflated when examining the association between language spoken at home and educational outcomes and how this explains educational disadvantages. Our primary aim is to estimate the association between language spoken at home and school-related outcomes ***among those children of immigrants for whom language switching is a relevant question***. Our secondary aim is to demonstrate that analysing this question without considering the distinctions made above can lead to false conclusions. This risk also applies to the question of whether language spoken at home can explain the disadvantages of migrant-origin students relative to

the majority.

We do this from a cross-national perspective. The extent to which these distinctions are relevant differs by country for two reasons. Firstly, countries differ in the number of national languages and the extent to which there are speakers of minority languages. Secondly, some countries share a national language with other countries — and receive substantial numbers of migrants with these shared language backgrounds — whereas others do not. There are also likely to be country differences in the association that we are really interested in: between children whose families have switched to speaking mainly/only the school language at home and those who continue to speak the origin language. We discuss the reasons for these potential differences after reviewing the reasons why language spoken at home may be related to our outcomes of interest.

**The relationship between language used at home and learning outcomes**

How is language use at home associated with learning? One perspective on this is the way in which different languages influence each other. On the one hand, first[1] languages can be vital resources for learning other languages and subjects (e.g., Agirdag & Vanlaar, 2018; Cummins, 2001) in which case languages can support each other as well as further learning. One reason for this is cross-language transfer: skills that children learn through their first language can be applicable for learning further languages (Durgunoğlu, 2017). On the other hand, languages may compete with each other so that learning (or using) a first language that is not the school language takes resources away from learning the second language (Norbahira & Radzuwan, 2018). Learning languages requires input and thus, almost by necessity, greater input of one language is away from input in the other language (Scheele et al., 2010). Learning in a second language may also be a more cognitively challenging task than learning in a first language (Takano & Noda, 1993).

Another perspective on the matter is through the influence of language maintenance on social capital. This is the reasoning in the segmented assimilation framework (e.g., Portes & Rumbaut, 2006). In contrast with straight-line (or neo) assimilation theory, which expects language differences to dissipate over time and generations (Alba & Nee, 1997), segmented assimilation theory expects bi-culturalism and bilingualism to be a major advantage for children of immigrants. This is particularly when parents have low levels of human capital (who may not learn the new language quickly themselves) but who are surrounded by a relatively large co-ethnic group. The resources of this group can be tapped more effectively when the origin language is spoken well. Maintaining this language thus supports social capital within families and their wider ethnic groups (Portes & Hao, 2002; Soehl, 2016), which promotes students' school success (Feliciano, 2001; though see Strobel, 2016, for contradictory evidence).

Linked to social capital, a further mechanism through which language use may influence other outcomes is communication between parents and children. Using the language of parental origin may be beneficial for academic achievement when parents are not proficient in the language of instruction, as maintaining the first language allows for within-family communication and parental support of children's education (Mouw & Xie, 1999). Reductions in the extent to which the language of origin is used with parents can be associated with greater mental health challenges when this reduction happens in an unsupportive family environments (Kilpi-Jakonen & Kwon, 2023) and these may also have repercussions on academic performance.

Previous research on language spoken at home and learning outcomes has produced mixed findings. In the United States, students who maintain their first languages have been found to have better learning outcomes than children of immigrants who mainly use English (e.g., Feliciano, 2001; Lutz & Crist, 2009). However, cross-national research has produced contrasting results. For example, Schnepf (2007) analysed a range of international learning assessments and found that not speaking (only) the school language at home was associated with lower learning outcomes in all countries, with the difference being up to about half a standard deviation in some countries. Using PISA 2006 data, Dustmann et al. (2012) reported that controlling for language spoken at home (school language versus not) explained either all or much of the

gap in reading and maths scores in many countries. As argued in the introduction, these cross-national examinations do not pay attention to the possible different reasons that children (of migrant origin or not) speak another language than that of the school at home — or do not. From the point of view of both theory and societal relevance, it is important to make these distinctions.

**The relationship between language used at home and sense of belonging**

How is language use at home associated with a sense of belonging at school? Whereas the focus for academic achievement is on the interplay between different languages and what happens at home and in the ethnic community (we will discuss educational policies in the next section), school belonging emphasizes the role of peer relationships, teacher support, and school climate (DeNicolo et al., 2017; St-Amand et al., 2017).

The role of teachers is one of the most crucial ones for students' sense of belonging overall (Allen et al., 2018). For children who do not speak the school language at home, teachers' beliefs in the impact of different languages on learning can have a substantial impact on their sense of belonging. Teachers across different countries have been found to uphold strong norms of monolingualism (Alisaari et al., 2019; Bosch et al., 2024; Pulinx et al., 2017), which can undermine their belief that multilingual students can achieve highly. This may have repercussions on students' sense of belonging because teachers' beliefs influence their treatment of students (Osterman, 2000). The beliefs of teachers can also be reflected in school diversity policies; policies explicitly promoting multiculturalism (including multilingualism) tend to reduce the minority–majority gap in sense of belonging (Celeste et al., 2019).

Teachers also shape students' perceptions of others, thus influencing peer relationships (Osterman, 2000). Teachers' monolingual ideals can shape the sense of belonging of multilingual students through the influence they have on other students. Peer relationships and sense of belonging may also be directly influenced by speaking another language at home if this is an impediment to fluency in the school language (Morrison et al., 2003). This is partly because language difficulties, including speaking with a foreign accent, may make forming friendships more challenging (Borgonovi & Ferrara, 2020). Peer groups also have their own views on acculturation ideals and not fitting into these norms may lead to peer rejection (Celeste et al., 2016). The situation is particularly challenging for youth from large ethnic groups with strong norms of cultural maintenance, who navigate between these co-ethnic peer group norms of maintenance and cross-ethnic peer group norms of cultural adoption.

Previous research on the association between language spoken at home and sense of belonging at school is relatively scarce. A study on Asian Americans found that not speaking (only) English was related to lower feelings of belonging and greater difficulties making friends (Yu et al., 2002). This was attributed to lower acculturation but measures such as length of residence or generation were not included. In a cross-national examination across 41 countries, speaking a foreign language at home was associated with a lower sense of belonging at school on average (Chiu et al., 2016). However, there was substantial cross-country variation: the association was significant and negative in 39% of countries but in no country significant and positive.

**Country differences in the association between home language and children's outcomes**

The ways in which the education system supports migrant-origin children is likely to play a strong role for both their learning outcomes as well as how the language spoken at home is associated with these outcomes. In education systems that are better able to improve the language proficiency of children coming from different language backgrounds, language spoken at home should play less of a role, especially for learning outcomes but potentially also for sense of belonging. One measure of the supportiveness of education systems is the Migrant Immigration Policy Index (MIPEX) (Solano & Huddleston, 2020). Taking MIPEX 2014 values, which reflect the most recent situation before the data collection of 2018 in

the countries we analyse, countries with favourable education policies were Canada and Finland, those with slightly favourable ones Australia, Belgium and New Zealand, those with half-way favourable ones Austria, Germany, Luxembourg and Switzerland, with Denmark and Greece being classified as slightly unfavourable.

Language use may constitute a social boundary between ethnic majorities and minorities, in addition to others based on citizenship, religion, or race, with the nature of these boundaries varying across countries (Alba, 2005). In some contexts, ethnic boundaries are 'bright' and clearly indicate where people belong. In other cases, the boundaries are more ambiguous and 'blurry', and the assimilation of different groups is relatively easy. Bright boundaries can be crossed only when individuals discard their group's characteristics. Continuing to speak one's first language may constitute a brighter boundary in some countries, which may be particularly relevant to one's sense of belonging, since in these situations language can be used to exclude or include people. Officially multi- or bi-lingual countries (in our data Belgium, Canada, Finland, Luxembourg and Switzerland) may be ones where the national language situation makes language a bright boundary also for migrants. Comparative research on teachers' monolingual norms is relatively scarce but there are some indications that in Greece teachers view multilingualism more favourably than in the Netherlands or Germany, with teachers in Flanders (Belgium) having less positive views (Bosch et al., 2024). Finnish teachers have also been found to have relatively negative views on multilingualism among students (Alisaari et al., 2019).

## Data and methods

The research was carried out using data from the 2018 PISA study (OECD, 2019b) from the following 11 countries: Australia (maximum N in the analyses = 12,734), Austria (max N = 6,650), Belgium (max N = 8,110), Canada (max N = 21,399), Denmark (max N = 7,408), Finland (max N = 5,546), Germany (max N = 4,637), Greece (max N = 6,355), Luxembourg (max N = 5,062), New Zealand (max N = 6,003), and Switzerland (max N = 5,679). These countries had sufficiently detailed information about the students' and parents' countries of birth and the languages spoken at home to support detailed analysis of language use. They also included a sufficient number of students in the relevant groups.

### Independent variables

We examine the associations of interest (between language spoken at home and children's outcomes) with different specifications of the variable measuring language spoken at home. These different specifications exclude certain languages and groups in a stepwise manner to demonstrate that the way in which this variable is often measured (and included in statistical analyses of cross-national data in particular) does not necessarily give an accurate picture of how language switching within a family may be associated with school-related outcomes. These steps are described below and illustrated in Figure 2.

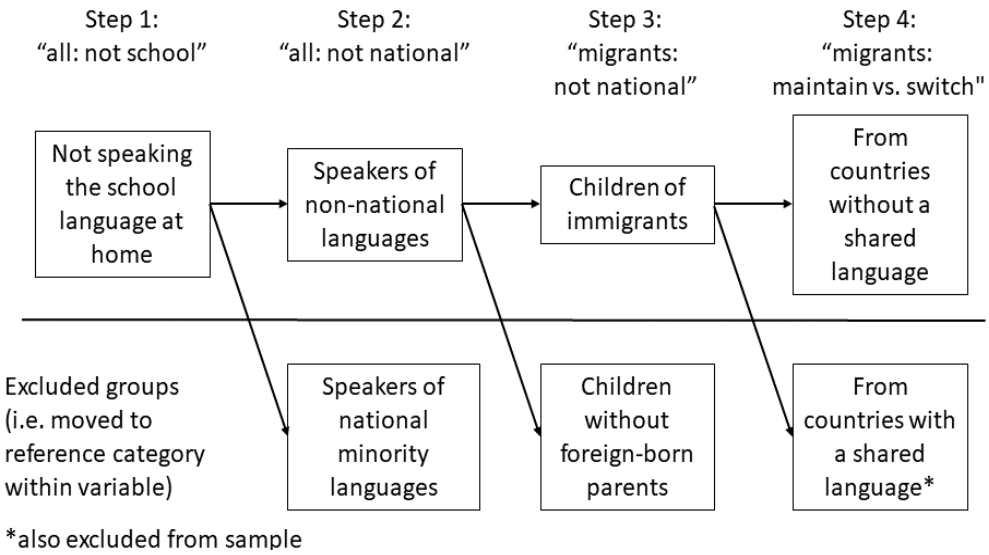


*Figure 2. Steps taken in defining the main independent variable of language spoken at home (and group within which it is measured)*

The main measure related to language spoken at home from previous research using PISA data is the question of whether a language other than the test language is spoken at home most of the time (variable ST022Q01TA in the data). We began our analyses using this same variable (Step 1 in Figure 2). However, this includes majority students who attend schools where the language is different to the one they speak at home, which can include official languages (e.g., French-speaking Canadians attending English-language schools) and regional dialects (e.g., Walloon or German dialects in Belgium). In some countries, speaking another language may also be due to the test (school) language being a foreign language (e.g., in Luxembourg, 262 students tested in English, less than half of whom spoke English at home).

To address this, in the second step we redefined this variable to speaking a *language other than a national language* at home and classified all official languages and regional dialects as national languages (the exact language is in the LANGN variable in the original data). In the third step, we kept this definition but limited the language variable to only consider children of immigrants (based on the COBN_M, COBN_F and COBN_S variables, which include information on the countries of birth of the mother, father and student respectively). Although the second step already excludes many speakers of minority languages, in many countries there are still minority groups whose languages are not national languages or regional dialects, some of whom are groups with longer migration backgrounds (e.g., grandparents) who have continued to speak their origin language.

In the fourth and final step, we limited the language variable as well the sample on the side of children of immigrants to those whose parents were born in countries where the official languages of the countries of destination are not spoken; for example, students of Indian origin in Canada were omitted because English is an official language in both Canada and India. This allowed for more precision about language switching (or in contrast language maintenance) taking place within immigrant families. The full list of countries of origin that were included for each country of destination can be seen in Table A1 (Appendix). PISA does not give students the option of choosing countries of birth from a full list of countries; rather, the categories reflect the main countries of origin in each country and include country groupings. Therefore, our countries of origin are in some cases a rather conservative list since the country groupings may also include countries that share a common language. We also excluded all children from mixed marriages, i.e.

with one foreign-born and one native-born parent, because different processes take place when it comes to the decision of which language to mostly speak at home in comparison to families with two foreign-born parents.

Table 1 includes the descriptives for these four variables by country. By far the largest proportion of students not speaking mainly the school language at home can be found in Luxembourg (83 %). However, this falls by half (to 40 %) when national languages and dialects are removed from this variable. In Switzerland approximately a quarter and in Austria and Canada approximately a fifth of students do not mainly speak the school language at home. In addition to Luxembourg, the countries where there is a relatively substantial difference between the first two ways of defining language spoken at home are Belgium, Canada, Finland and Switzerland – all officially multilingual countries. In Austria, Germany and Greece there is only one national language and no regional dialects (in PISA data), thus there is no difference between the first two specifications. The countries also display wide variation in the proportion of children of immigrants not speaking mainly the language of the school at home. Among all children of immigrants, this ranges from approximately 40 % in Australia, Greece and New Zealand to over 70 % in Austria. Among the groups that we can identify as coming from countries that do not share a common language, the proportion of students maintaining an origin language tends to be higher, with the highest percentages found in Luxembourg (86 %), Austria (77 %) and Switzerland (75 %) and the lowest in Greece (34 %) and Denmark (49 %).

*Table 1. Distribution of different language spoken at home variables within each country (% and N)*

| | Full sample | | Only (relevant) migrant origin | |
|---|---|---|---|---|
| | % | N | % | N |
| Australia | | | | |
| Not test language | 13.0 % | 12718 | | |
| Non-national language | 12.6 % | 12718 | | |
| Non-national language, only migrant origin | 11.1 % | 12734 | 39% | 3416 |
| Non-national language, only specific origins | 4.4 % | 7664 | 59% | 533 |
| Austria | | | | |
| Not test language | 20.3 % | 6644 | | |
| Non-national language | 20.3 % | 6644 | | |
| Non-national language, only migrant origin | 16.7 % | 6650 | 73% | 1398 |
| Non-national language, only specific origins | 16.1 % | 5593 | 77% | 1065 |
| Belgium | | | | |
| Not test language | 17.3 % | 8106 | | |
| Non-national language | 12.4 % | 8106 | | |
| Non-national language, only migrant origin | 9.5 % | 8110 | 52% | 1460 |
| Non-national language, only specific origins | 4.4 % | 5900 | 70% | 365 |
| Canada | | | | |
| Not test language | 20.9 % | 21378 | | |
| Non-national language | 18.2 % | 21378 | | |
| Non-national language, only migrant origin | 17.3 % | 21399 | 49% | 5835 |
| Non-national language, only specific origins | 6.5 % | 14615 | 61% | 1010 |

| | Full sample | | Only (relevant) migrant origin | |
|---|---|---|---|---|
| Denmark | | | | |
| Not test language | 7.8 % | 7404 | | |
| Non-national language | 7.7 % | 7404 | | |
| Non-national language, only migrant origin | 5.3 % | 7408 | 48% | 1555 |
| Non-national language, only specific origins | 5.9 % | 6532 | 49% | 1537 |
| Finland | | | | |
| Not test language | 7.3 % | 5543 | | |
| Non-national language | 5.6 % | 5543 | | |
| Non-national language, only migrant origin | 4.1 % | 5546 | 64% | 347 |
| Non-national language, only specific origins | 2.7 % | 5052 | 61% | 220 |
| Germany | | | | |
| Not test language | 17.0 % | 4636 | | |
| Non-national language | 17.0 % | 4636 | | |
| Non-national language, only migrant origin | 13.5 % | 4637 | 62% | 1018 |
| Non-national language, only specific origins | 6.9 % | 3353 | 68% | 343 |
| Greece | | | | |
| Not test language | 6.2 % | 6345 | | |
| Non-national language | 6.2 % | 6345 | | |
| Non-national language, only migrant origin | 4.8 % | 6355 | 38% | 769 |
| Non-national language, only specific origins | 2.2 % | 5163 | 34% | 309 |
| Luxembourg | | | | |
| Not test language | 82.8 % | 5059 | | |
| Non-national language | 39.5 % | 5059 | | |
| Non-national language, only migrant origin | 35.3 % | 5062 | 64% | 2783 |
| Non-national language, only specific origins | 42.3 % | 2795 | 86% | 1367 |
| New Zealand | | | | |
| Not test language | 14.2 % | 5996 | | |
| Non-national language | 13.2 % | 5996 | | |
| Non-national language, only migrant origin | 12.1 % | 6003 | 44% | 1643 |
| Non-national language, only specific origins | 5.2 % | 3579 | 71% | 265 |
| Switzerland | | | | |
| Not test language | 26.8 % | 5673 | | |
| Non-national language | 23.1 % | 5673 | | |
| Non-national language, only migrant origin | 20.3 % | 5679 | 59% | 1963 |
| Non-national language, only specific origins | 20.4 % | 3504 | 75% | 948 |

Weighted percentages (calculated using repest)
Based on samples used for the models with reading scores
Source: Own calculations based on PISA 2018.

Table S1 (supplementary online materials) provides information on missingness in the variables used. In most countries less than 2 % of students have missing information on the language mainly spoken at home, but this is more substantial in Australia (9.8 %) and Germany (11.7 %). These students also tend to have missing information on the socio-demographic variables used as control variables as well as the second dependent variable. The sample for the third specification is slightly larger than for the first two because all majority students are included there even if they had missing information on language.

**Dependent variables**

In PISA 2018, reading literacy is defined as "understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential and to participate in society" (OECD, 2019a, p. 28). The reading test used a multistage adaptive design to improve measurement accuracy and efficiency (OECD, 2021). As in previous PISA cycles, this resulted in a set of plausible values for each student. The mean reading scores in the countries in our sample ranged between 460–520, with a standard deviation of around 100. For the analyses, we z-standardised each plausible value (mean zero and standard deviation of one).

Feeling of belonging is a "feeling of being accepted, respected, included and socially supported in the school environment" (OECD, 2019a, p. 273). This was measured with six items on a four-point Likert scale, including "I make friends easily at school." and "I feel awkward and out of place in my school." (OECD, 2021). This scale was produced by the OECD and is z-standardised (the BELONG variable). The mean level of belonging in the countries in our sample ranged between -0.20–0.40, with a standard deviation around 1.

**Control variables**

In all our models we control for gender (ST004D01T) and student's grade in comparison with the modal grade for 15-year-olds in the respective countries (GRADE). While girls tend to have higher reading scores than boys, they have often been found to have a lower sense of belonging at school, particularly as teenagers. The grade in which students study has a strong relationship with their learning outcomes but may also be associated with their sense of belonging through developmental processes. Moreover, students who have been held behind, may have greater problems in integrating with their classmates.

We also control for the migration generation of students with the categories consisting of majority students with only native-born parents, 2.5 generation students with one foreign-born and one native-born parent, 2nd generation students who were born in the country of destination but have only foreign-born parents, and 1st generation students who were born abroad and have foreign-born parents. It should be noted that when the language specification changes to only measure language within the group of children of immigrants, the interpretation of the generational differences switches slightly to refer mainly to generational differences between the majority and children who speak the test language at home (though with the simplification that we assume the language difference to be the same within both the 1st and 2nd generation). For 1st generation students, we control for age at arrival (ST021Q01TA) — all other students are coded as 0. This means that the categorical variable of generation essentially compares 1st generation students who arrived at the age of 0 to the other generation groups. Finally, we control for parental country of birth; as mentioned above the countries of birth available depend on the survey country. We used the mother's country of birth unless this was missing or the mother was native born and the father was not, in these cases we used the father's.

We present models with and without controlling for parental socioeconomic (SES) status. As our measure of parental SES, we use the OECD's index of economic, social, and cultural status (ESCS). This is a composite index including parents' education, occupation, and home possessions (OECD, 2021). ESCS is scaled to have a mean of zero and a standard deviation of one across OECD countries. We know from

previous research that parental SES is very influential for educational outcomes, and children of parents with low levels of education are more likely to use their first languages at home (Soehl, 2016), although children of highly educated parents are more likely to be fluently bilingual (Portes & Hao, 2002, Soehl, 2016). Therefore, parental SES may confound the relationship between language spoken at home and the outcomes we consider.

### Methods and supplementary analyses

Linear regression models were run separately for each country. We took into account the stratification of the sample by using the recommended weighting procedure from the OECD. For analyses of the reading scores, we ran the models with the ten plausible values and replicate weights. We used the user-written Stata repest command (Avvisati & Keslair, 2014) for all our analyses, including the descriptive results in Table 1. For each language variable we present the results both with and without the control for parental SES. We applied listwise deletion for missing values, and the analyses of the two dependent variables used different samples based on missingness in the respective variables.

The main results are shown as figures in the main text. The full models are in the supplementary materials (Tables S2a and S3a). We also ran the models with the second specification (non-national language spoken at home) including an interaction between language and generation. These results are in the supplementary materials (Tables S2b and S3b). The supplementary materials also include the results using the two other competence measures: mathematics and science (Tables S4 and S5).

In order to further assess country differences, we tested the equality of coefficients between the models based on the separate country samples using seemingly unrelated equations. We did this only for the final model, where language used at home is assessed for the groups that do not come from countries sharing a common language and including the control for parental SES. Because these tests cannot be run after models estimated with repest, we re-ran these models using Stata's survey estimation command, which allows for the use of the overall weights, balanced repeated replicate (BRR) weights and Fay's adjustment. As the dependent variable for reading scores, we used the average of the plausible values. Overall, the results change rather minimally in comparison to the repest-results, the main difference in many countries being slightly smaller standard errors. The results of these models and the test for the equality of coefficients relative to Canada, Denmark, New Zealand and Switzerland are the supplementary materials (Tables S2c and S3c).

We acknowledge some limitations when using PISA data. For example, we do not know whether the families we defined as 'switching' ever spoke the languages of their country of origin; we made these assumptions based on their country of birth and its official language(s). Similarly, the estimated difference between the different language use categories cannot be interpreted as causal; we relied on cross-sectional data, language choices may also have been endogenous to the outcomes studied. Furthermore, since the students were only able to choose one language (other than the test language) to describe their home language, the characterization of their linguistic biography is in many cases simplified.

## Results

### Reading scores

The results for the different specifications of language spoken at home and their associations with reading scores are presented in Figure 3 (also Table S2a). The first results showed that across all countries, children who mainly speak a language that is not that of the test (i.e. the school) have lower reading scores than children who mainly speak that language at home. In many cases, this disadvantage was reduced somewhat

when controlling for parental SES, though in Belgium the difference hardly changed and in Finland the difference grew slightly.

The second set of results excluded national languages and regional dialects from the language specification. As described in the data section, the difference between the first specification and the second is relevant for countries that have multiple official languages and/or substantial regional dialects as minority languages, such as Belgium, Finland, Luxembourg and Switzerland. In countries where there is a difference between these two, the estimate for the second tended to be smaller than for the first, although this difference was substantial only in Belgium, Canada and New Zealand. Moreover, in some countries the distinction is relevant but the results hardly changed, such as Finland and Switzerland. Again, parental SES explained some of the disadvantages of speaking a different language at home. In most countries, speaking a language that is not that of the test or a national/regional one was associated with reading scores 0.2–0.4 standard deviations lower than for those speaking a test/national language at home, which in many cases was in the same order of magnitude as the gender difference.

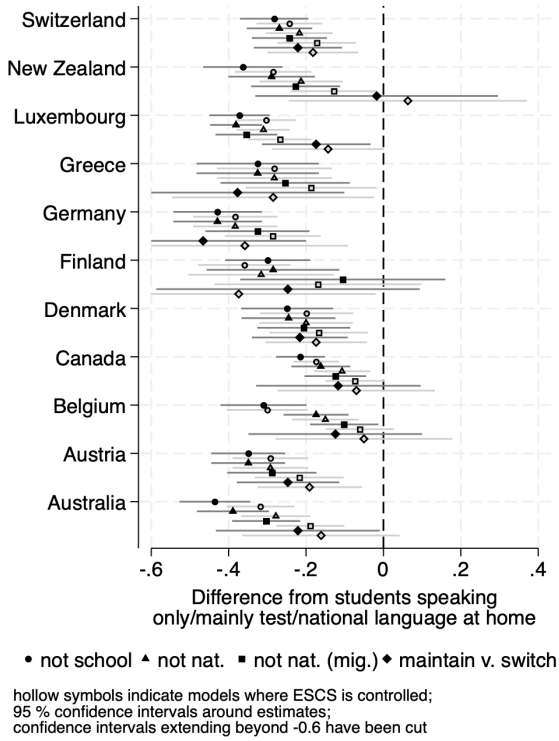


*Figure 3. Estimated association of (not) speaking the test or national language at home and reading scores based on different specifications of the language variable and different samples (full models in Table S2a).*

In the third specification, we restricted the variable to measure language use specifically among children of immigrants (also excluding the 2.5 generation). Across all the countries analysed, this reduced the negative association. This reduction was the most notable in Finland, whereas the smallest change was for Luxembourg and Switzerland. This suggests that speaking a non-national/non-minority language at home is more detrimental for children who have native-born parents. These are likely either third (or further) generations with migrant origins or the 2.5 generation. Controlling for parental SES once again reduced

the size of the estimates in most countries.

In the final specification, the language variable was only measured for the sample of migrant-origin children from countries where the national language of the destination country is not an official (or substantial minority) language. In most countries, this reduced the estimate, but most notably in Finland, Germany and Greece the fourth estimate was larger than the third. In this final specification, the difference between the two language groups is negligible (and not statistically significant) in Belgium, Canada and New Zealand. After controlling for parental resources, the estimate was also not statistically significant in Australia or Luxembourg, though in these cases the estimate was similar in size as in other countries where the estimate was significant: approximately -0.15. Testing the equality of coefficients across the countries (Table S2c) suggests that quite a number of countries differ significantly from New Zealand, namely Austria, Finland, Germany and Switzerland, with Denmark and Luxembourg close to being so. However, only Finland and Germany are close to being significantly different from Canada.

Rather than categorising all language minority students who are not children of immigrants into the reference category of the language variable, as is done in the third specification, another strategy could also be to include an interaction between the language variable used in the second specification and generation. In the supplementary materials (Table S2b) we show these results: across all countries the coefficient for the majority was negative, whereas the interaction coefficient for the first and second generation was positive. The interaction coefficient for the 2.5 generation also tended to be positive but not as large as the ones for the first and second generations. This also suggests that language spoken at home has a differential impact for children of immigrants in comparison with national language minorities.

The supplementary tables also display the results for the other two competence measures (Tables S4 and S5). In many countries, the results for 'maintaining' vs. 'switching' were consistent across all three measures. However, the difference in scores tended to be slightly smaller for the other two measures, in particular math, and in the final specification, the difference was not statistically significant in more countries (for math Denmark, Finland, Greece and Switzerland in addition to the already mentioned Australia, Belgium, Canada and New Zealand; for science Denmark and Finland – though in the case of Finland and Greece the estimates remained relatively large in size).

**Sense of belonging**

Figure 4 shows the main results for the sense of belonging at school (also Table S3a). There were some similarities to the results for reading: in many cases the estimates were progressively less negative with each specification and slightly less negative (or more positive) when controlling for parental resources. However, these differences were relatively small and there was substantial country variation in these patterns.

In Australia, Germany, New Zealand and Switzerland, none of the specifications produced statistically significant associations: language spoken at home did not seem to matter for the sense of belonging in these countries. Other countries showed somewhat differing trends. In Belgium, Denmark, and Greece, the association was consistently negative. In Belgium and particularly Greece, the negative estimate was the largest in the final specification. This was also the case for Finland, though in this case the confidence intervals were so large that the estimates were not statistically significant for the third and fourth specifications. In these four countries, the negative estimates were approximately in the range of 0.2–0.4 standard deviations.

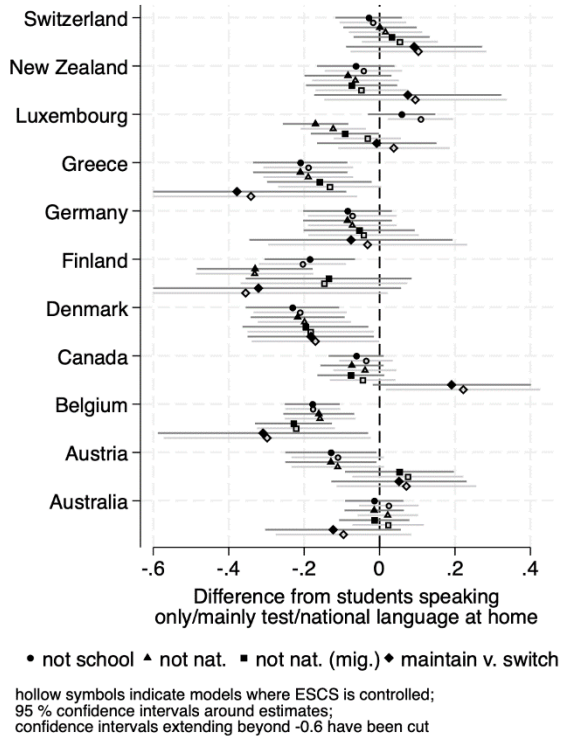


*Figure 4. Estimated association of (not) speaking the test or national language at home and sense of belonging based on different specifications of the language variable and different samples (full models in Table S3a).*

Luxembourg showed an interesting pattern whereby in the first specification speaking a non-test language at home was associated with a higher sense of belonging, which was statistically significant when controlling for parental resources. When redefining this to exclude national/regional languages (or a non-test language in English-speaking schools), the estimate was negative. In the third and fourth specification, for children of immigrants, the estimates were small and not statistically significant.

Testing the equality of coefficients for the final model across the different countries (Table S3c), the positive coefficients in Canada, New Zealand and Switzerland were found to be significantly different from the negative ones in Belgium, Denmark, Finland and Greece, with the one for Canada also being significantly different from the one for Australia. On the other hand, the coefficient for Denmark was only significantly different from the three aforementioned ones.

With regard to the supplementary analyses that introduced an interaction between generation and the second specification, the results were much in line with those for reading (Supplementary Table S3b). In most cases, the association for the majority was negative whereas for children of immigrants the interaction was positive (i.e. the association less negative than for the majority). However, most coefficients were not statistically significant.

## Discussion and conclusion

We have analysed how language spoken at home is associated with test scores and sense of belonging at school, particularly among migrant-origin children. Our results suggest that some overly simplistic conclusions about the influence of home language (and language proficiency more broadly) have been drawn previously due to the way that language use has been specified using international assessment data. One part of the problem is that different processes are at play when migrant-origin students do not speak the test/school language at home versus when native-origin students do not speak a national language (including regional dialects). The relevance of this distinction differs from country to country, with it naturally being most accentuated in multilingual countries such as Belgium, Canada, Finland, and Luxembourg. In addition, comparisons need to be done within groups where language switching is relevant. Our first analyses found a negative association between speaking 'another' language at home and reading test scores across all the analysed countries. However, the different specifications used in our analyses show how results change when language use is measured in different ways and for different groups. Our preferred specification is the last one, which focuses on groups where we can ascertain that language switching is a relevant process – in contrast with arriving in the country already speaking the language of the destination country as a home language – and in these models the initial negative associations for language maintenance disappeared in a number of countries.

To summarize our main results from this final specification in more detail, countries could be grouped into approximately three groups when it came to the difference in reading test scores between those maintaining the language of the country of origin at home and those switching to a national language. The countries where there was no difference were Belgium, Canada and New Zealand. The countries with a moderate association (approximately 0.15–0.20 of a standard deviation) were Austria, Denmark and Switzerland. In-between these first two groups were Australia and Luxembourg, where the coefficients were in-line with those of the second group, but the difference was not statistically significant. In the final group were Finland, Germany and Greece, with relatively large differences (approximately 0.3–0.4). To put this into the perspective of other social differences: in many countries gender differences are approximately of the same magnitude as the differences in the second group of countries, whereas a standard deviation change in parental SES is of approximately the same magnitude as the difference in the last group of countries.

In terms of sense of belonging at school, maintaining the language of the country of origin at home was associated with a lower sense of belonging than switching in a smaller number of countries: namely Belgium, Denmark, Finland and Greece, with the difference in Denmark being slightly smaller than in the others (under 0.2 versus 0.3–0.4). In Australia, the estimate was also negative but not statistically significant, whereas in other countries those maintaining the origin language were estimated to have a higher sense of belonging than those switching, though this difference was not statistically significant except for in Canada (0.2). Overall social differences are smaller for sense of belonging than for reading, with gender, grade-level and one standard deviation parental SES differences being between 0 and 0.2.

There is thus considerable variation in how patterns of language maintenance and switching are associated with incorporation measures in different countries. Our results suggest that continuing to speak the origin language is most disadvantageous in Denmark, Finland and Greece, where it is associated with both lower reading scores and a lower sense of belonging. In other countries, there is an association with either lower reading scores (Austria, Switzerland, and Germany) or a lower sense of belonging (Belgium). Of the countries analysed, Australia, Canada, Luxembourg and New Zealand stand out as having no significant disadvantages for those continuing to speak the origin language; in Canada they also have a higher sense of belonging than those switching.

Our results thus suggest that the more classical (Anglophone) countries of immigration provide contexts where speaking different languages does not constitute a major barrier for incorporation (also supported by previous research from the U.S.: Feliciano, 2001; Lutz & Crist, 2009; though not Yu et al., 2002), whereas in the newer (European) countries of immigration, non-national languages constitute barriers in

some form or another, with the possible exception of Luxembourg. In these latter contexts, families who choose to switch to the language of the destination country seem to provide their children with better opportunities. However, we cannot say whether the associations are causal; better incorporation may lead to language switching or vice versa. It is also unknown whether it is parents who switch languages or if children refuse to speak their first language with their parents. It should also be noted that the migrant groups included in these analyses do not cover all the groups for whom language switching would be relevant since not all countries of birth can be identified separately; we also compare different groups across countries. Among specific groups, the results may be slightly different (e.g., Kilpi-Jakonen & Alisaari, 2022).

At the policy level, the findings related to learning outcomes seem to reflect the extent to which immigrant integration is supported in the education system (using the 2014 education index of the MIPEX), with the exception of Finland and partly also Luxembourg. More specifically, Canada, Australia, New Zealand, and Belgium score highly in terms of having a comprehensive approach to integration, and there is no gap in learning outcomes due to language spoken at home. Austria, Germany and Switzerland have relatively middling scores, and Denmark and Greece somewhat low ones, and we find learning gaps in all of these countries. In Luxembourg, MIPEX scores are also middling, but there is no language-related learning gap. Finland, on the other hand, scores highly in the MIPEX, but language differences are still evident, reflecting previous research on how Finland is supportive of multiculturalism at the policy level but not in practice (Saukkonen, 2013). In Denmark and Finland, clear language hierarchies exist in educational policies: the official national languages are at the top of a language ideological hierarchy, followed by national minority languages (Alisaari et al., 2023).

The low MIPEX scores of Denmark and Greece may also be linked to the results for sense of belonging. Moreover, it seems that teachers' relatively positive attitudes towards multilingualism in schools in Greece (Bosch et al., 2024) do not help multilingual students to thrive in the school environment. The results for Belgium and Finland may be related to historically power-related language hierarchies that can still be reflected in school contexts as ideologies that maintain monolingual norms. In these officially bilingual countries, national languages are still mainly kept separate from one another, which may relate to how new languages are perceived and tolerated in schools (Alisaari et al., 2019; Pulinx et al., 2017). This kind of bright boundary (Alba, 2005) seems to show as a difference in the sense of belonging at school.

In some countries there thus seems to be a clear challenge to close the gap in learning outcomes between students who maintain their origin languages and those who switch to using the school language at home. This may be due to different orientations towards first languages and whether they are seen as a resource or a barrier for learning within schools. For students, this may influence the ways in which they incorporate multilingualism into their identities and whether they can draw on it as a resource (Rutgers et al., 2021). Previous research has shown that in Germany, where speaking another language at home is associated with lower learning outcomes, this disadvantage does not exist for students with higher levels of German proficiency (Strobel, 2016). In other words, the focus should be on developing language proficiency regardless of language spoken at home, and some countries appear to be more successful than others in doing this. This focus should also extend beyond children of immigrants.

To conclude, some previous studies have claimed that the language used at home defines students' linguistic abilities (Schnepf, 2007; for a review of similar approaches, see Agirdag & Vanlaar, 2018) and suggested that migrant parents should eventually change the language spoken at home (e.g., Dustmann et al., 2012; see also Casey & Dustmann, 2008). This also comes through in OECD analyses (e.g., 2019c, 2023). However, the disadvantages that students face when continuing to speak the language of parental origin are by no means universal. Moreover, when parents with low levels of proficiency use the language of the surrounding society at home, their children's language development is not optimally supported (e.g., Scheffner Hammer et al., 2009). Schools have a high potential to influence how cultural diversity is perceived in an academic setting (Celeste et al., 2019); a supportive school culture values students' identities and considers first languages to be valuable learning resources that should be maintained, not lost (Cummins, 2001). This is also highlighted in the OECD reports that focus on policies to support migrant

children (e.g., OECD, 2015). At both national and local levels, policies and strategies should be further developed to support immigrant parents and their children in linguistic choices

## Endnotes

1 We use the term first language to refer to the languages that children may have learnt first, with a specific assumption that many children of immigrants are likely to have learnt a language other than that of the surrounding society first. As a synonym, we also use origin language to more concretely refer to the languages spoken in the parents' countries of origin. However, we acknowledge that these terms are not completely neutral

## Funding

This research was funded by the Research Council of Finland (Academy Fellowship, 316247, and Research Council of Finland Flagship Programme, 345546).

## Data availability statement

The data that support the findings of this study are openly available from the OECD at https://www.oecd.org/pisa/data/2018database/. The Stata do-files used for data manipulation and analysis can be found at the INVEST GitHub page (https://github.com/INVEST-flagship).

## References

Agirdag, O., & Vanlaar, G. (2018). Does more exposure to the language of instruction lead to higher academic achievement? A cross-national examination. *International Journal of Bilingualism*, *22*(1), 123–137. https://doi.org/10.1177/1367006916658711

Alba, R. (2005). Bright vs. blurred boundaries: Second-generation assimilation and exclusion in France, Germany, and the United States. *Ethnic and Racial Studies*, *28*(1), 20-49. https://doi.org/10.1080/0141987042000280003

Alba, R., & Nee, V. (1997). Rethinking assimilation theory for a new era of immigration. *International Migration Review*, *31*(4), 826–874. https://doi.org/10.2307/2547416

Alisaari, J., Heikkola, L. M., Commins, N., & Acquah, E. O. (2019). Monolingual ideologies confronting multilingual realities. Finnish teachers' beliefs about linguistic diversity. *Teaching and Teacher Education*, *80*, 48–58. https://doi.org/10.1016/j.tate.2019.01.003

Alisaari, J., Møller Daugaard, L., Dewilde, J., Harju-Autti, R., Heikkola, L. M., Iversen, J. Y., Kekki, N., Pesonen, S., Reath Warren, A., Straszer, B., & Yli-Jokipii, M. (2023). Mother tongue education in four Nordic countries - problem, right or resource?. *Apples - Journal of Applied Language Studies*, *17*(2), 52–72. https://doi.org/10.47862/apples.113671

Allen, K., Kern, M. L., Vella-Brodrick, D., Hattie, J., & Waters, L. (2018). What schools need to know about fostering school belonging: A meta-analysis. *Educational Psychology Review*, *30*, 1–34. https://doi.org/10.1007/s10648-016-9389-8

Avvisati, F., & Keslair, F. (2014). *REPEST: Stata module to run estimations with weighted replicate sam-*

*ples and plausible values* [Computer software]. Statistical Software Components, Boston College Department of Economics.

Borgonovi, F. (2018). How do the performance and well-being of students with an immigrant background compare across countries? *PISA in Focus, 82*. OECD Publishing. https://doi.org/10.1787/a9e8c1ab-en

Borgonovi, F., & Ferrara, A. (2020). Academic achievement and sense of belonging among non-native-speaking immigrant students: The role of linguistic distance. *Learning and Individual Differences*, *81*, 101911. https://doi.org/10.1016/j.lindif.2020.101911

Bosch, J. E., Olioumtsevits, K., Santarelli, S. A. A., Faloppa, F., Foppolo, F., & Papadopoulou, D. (2024). How do teachers view multilingualism in education? Evidence from Greece, Italy and The Netherlands. *Language and Education*, *39*(2), 337–356. https://doi.org/10.1080/09500782.2024.2380073

Casey, T., & Dustmann, C. (2008). Intergenerational transmission of language capital and economic outcomes. *Journal of Human Resources*, *43*(3), 660–687. https://doi.org/10.3368/jhr.43.3.660

Celeste, L., Baysu, G., Phalet, K., Meeussen, L., & Kende, J. (2019). Can School Diversity Policies Reduce Belonging and Achievement Gaps Between Minority and Majority Youth? Multiculturalism, Colorblindness, and Assimilationism Assessed. *Personality and Social Psychology Bulletin*, *45*(11), 1603–1618. https://doi.org/10.1177/0146167219838577

Celeste, L., Meeussen, L., Verschueren, K., & Phalet, K. (2016). Minority acculturation and peer rejection: Costs of acculturation misfit with peer-group norms. *British Journal of Social Psychology*, *55*(3), 544–563. https://doi.org/10.1111/bjso.12147

Chiu, M. M., Chow, B. W.-Y., McBride, C., & Mol, S. T. (2016). Students' Sense of Belonging at School in 41 Countries: Cross-Cultural Variability. *Journal of Cross-Cultural Psychology*, *47*(2), 175–196. https://doi.org/10.1177/0022022115617031

Cummins, J. (2001). *Negotiating identities: Education for empowerment in a diverse society* (2nd ed.). California Association for Bilingual Education.

DeNicolo, C. P., Yu, M., Crowley, C. B., & Gabel, S. L. (2017). Reimagining critical care and problematizing sense of school belonging as a response to inequality for immigrants and children of immigrants. *Review of Research in Education*, *41*(1), 500–530. https://doi.org/10.3102/0091732X17690498

Durgunoğlu, A. Y. (2017). An updated review of cross-language transfer and its educational implications. In Segers, E., & van den Broek, P. V. D. (Eds.), *Developmental perspectives in written language and literacy: In honor of Ludo Verhoeven* (pp. 167–182). John Benjamins Publishing Company. https://doi.org/10.1075/z.206.11dur

Dustmann, C., Frattini, T., & Lanzara, G. (2012). Educational achievement of second-generation immigrants: An international comparison. *Economic Policy*, *27*(69), 143–185.

Feliciano, C. (2001). The benefits of biculturalism: Exposure to immigrant culture and dropping out of school among Asian and Latino youths. *Social Science Quarterly*, *82*(4), 865–879. https://doi.org/10.1111/0038-4941.00064

Kilpi-Jakonen, E., & Alisaari, J. (2022). Language choices at home and their relationship with educational outcomes, with a special focus on children with origins in Former Yugoslavia and Turkey in six European countries. *Frontiers in Sociology*, *7*, 841847. https://doi.org/10.3389/fsoc.2022.841847

Kilpi-Jakonen, E., & Kwon, H. W. (2023). The Behavioral and Mental Health Benefits of Speaking the Heritage Language within Immigrant Families: The Moderating Role of Family Relations. *Journal of Youth and Adolescence, 52*(10), 2158–2181. https://doi.org/10.1007/s10964-023-01807-5

Lambert, N. A., Stillman, T. F., Hicks, J. A., Kamble, S., Baumeister, R. F., & Fincham, F. D. (2013). To belong is to matter: Sense of belonging enhances meaning in life. *Personality and Social Psychology Bulletin*, *39*(11), 1418–1427. https://doi.org/10.1177/0146167213499186

Lutz, A., & Crist, S. (2009). Why do bilingual boys get better grades in English-only America? The impacts of gender, language and family interaction on academic achievement of Latino/a children of immigrants. *Ethnic and Racial Studies*, *32*(2), 436–468. https://doi.org/10.1080/01419870801943647

Morrison, G.M., Cosden, M.A., O'Farrell, S.L., & Campos, E. (2003). Changes in Latino Students' Percep-

tions of School Belonging Over Time: Impact of Language Proficiency, Self-Perceptions and Teacher Evaluations. *The California School Psychologist*, *8*, 87–98. https://doi.org/10.1007/BF03340898

Mouw, T., & Xie, Y. (1999). Bilingualism and the academic achievement of first- and second-generation Asian Americans: Accommodation with or without assimilation? *American Sociological Review*, *64*(2), 232–252. https://doi.org/10.2307/2657529

Norbahira, M. N., & Radzuwan, A. R. (2018). A review of theoretical perspectives on language learning and acquisition. *Kasetsart Journal of Social Sciences*, *39*(1), 161–167. https://doi.org/10.1016/j.kjss.2017.12.012.

OECD. (2015). *Closing the Gap for Immigrant Students. Policies, Practice and Performance*. OECD Publishing. https://doi.org/10.1787/9789264075788-en

OECD. (2019a). *PISA 2018 Assessment and Analytical Framework*. OECD Publishing. https://doi.org/10.1787/b25efab8-en.

OECD. (2019b). *PISA 2018 database*. https://www.oecd.org/pisa/data/2018database/

OECD. (2019c). *PISA 2018 Results (Volume II). Where All Students Can Succeed*. OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en

OECD. (2021). *PISA 2018 Technical Report*. OECD Publishing. https://www.oecd.org/pisa/data/pisa2018technicalreport/

OECD. (2023). *PISA 2022 Results (Volume I). The State of Learning and Equity in Education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research, 70*(3), 323–367. https://doi.org/10.3102/00346543070003323

Portes, A., & Hao, L. (2002). The price of uniformity: Language, family and personality adjustment in the immigrant second generation. *Ethnic and Racial Studies*, *25*(6), 889–912. https://doi.org/10.1080/0141987022000009368

Portes, A., & Rumbaut, R. (2006). *Immigrant America: A Portrait*. University of California Press.

Pulinx, R., Van Avermaet, P., & Agirdag, O. (2017). Silencing linguistic diversity: the extent, the determinants and consequences of the monolingual beliefs of Flemish teachers. *International Journal of Bilingual Education and Bilingualism*, *20*(5), 542–556. https://doi.org/10.1080/13670050.2015.1102860

Rutgers, D., Evans, M., Fisher, L., Forbes, K., Gayton, A., & Liu, Y. (2021). Multilingualism, multilingual identity and academic attainment: Evidence from secondary schools in England, *Journal of Language, Identity & Education*, *23*(2), 210–227. https://doi.org/10.1080/15348458.2021.1986397

Saukkonen, P. (2013). Multiculturalism and nationalism: The politics of diversity in Finland. In P. Kivisto & Ö. Wahlbeck (Eds.), *Debating Multiculturalism in the Nordic Welfare States* (pp. 270–294). Palgrave Macmillan. https://doi.org/10.1057/9781137318459_11

Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*(1), 117–140. https://doi.org/10.1017/S0142716409990191

Scheffner Hammer, C., Dunn Davison, M., Lawrence, F. R., & Miccio, A.W. (2009). The effect of maternal language on bilingual children's vocabulary and emergent literacy development during Head Start and kindergarten. *Scientific Studies of Reading*, *13*(2), 99–121. https://doi.org/10.1080/10888430902769541

Schnepf, S. V. (2006). Immigrants' Educational Disadvantage: An Examination Across Ten Countries and Three Surveys. *Journal of Population Economics*, *20*(3), 527–545. https://doi.org/10.1007/s00148-006-0102-y

Soehl, T. (2016). But do they speak it? The intergenerational transmission of home-country language in migrant families in France. *Journal of Ethnic and Migration Studies*, *42*(9), 1513–1535. https://doi.org/10.1080/1369183X.2015.1126171

Solano, G., & Huddleston, T. (2020). *Migrant Integration Policy Index 2020: Education*. MIPEX. https://www.mipex.eu/education

St-Amand, J., Girard, S., & Smith, J. (2017). Sense of belonging at school: Defining attributes, determinants, and sustaining strategies. *IAFOR Journal of Education, 5*(2). https://doi.org/10.22492/ije.5.2.05

Strobel, B. (2016). Does family language matter? The role of foreign language use and family social capital in the educational achievement of immigrant students in Germany. *Ethnic and Racial Studies*, *39*(14), 2641–2663. https://doi.org/10.1080/01419870.2016.1145712

Takano, Y., & Noda, A. (1993). A Temporary Decline of Thinking Ability During Foreign Language Processing. *Journal of Cross-Cultural Psychology*, *24*(4), 445–462. https://doi.org/10.1177/0022022193244005

Yu, S. M., Huang, Z. J., Schwalberg, R. H., Overpeck, M. D., & Kogan, M. D. (2002). Association of language spoken at home with health and school issues among Asian American adolescents. *Journal of School Health*, *72*(5), 192–198. https://doi.org/10.1111/j.1746-1561.2002.tb06545.x

## Author biographies

**Elina Kilpi-Jakonen** is a Professor of Sociology at the INVEST Research Flagship Centre at the University of Turku.

**Jenni Alisaari** is a Professor of Education at the School of Applied Educational Science and Teacher Education at the University of Eastern Finland and a visiting senior researcher at the INVEST Research Flagship Centre at the University of Turku.

## Appendix

*Table A1. Countries of origin in the final language specification (countries that do not share languages)*

| PISA country | Countries of parental birth included |
|---|---|
| Australia | China<br>Vietnam<br>Smaller numbers: Greece, Italy |
| Austria | Balkans<br>Turkey<br>Smaller numbers: Afghanistan, Syria, other non-European countries |
| Belgium | Turkey<br>Eastern Europe |
| Canada | China<br>Smaller numbers: Iran, Korea, Syria, UAE |
| Denmark | Iraq<br>Somalia<br>Turkey<br>Smaller numbers: Afghanistan, Former Yugoslavia, Lebanon, Pakistan, Syria, Nordic countries, other countries |

| | |
|---|---|
| Finland | Somalia<br>Smaller numbers: Iraq, Balkans, China, Vietnam, Turkey, Afghanistan, other countries |
| Germany | Turkey<br>Smaller numbers: Former Yugoslavia, Poland, Greece |
| Greece | Former USSR Republics |
| Luxembourg | Former Yugoslavia<br>Portugal<br>Smaller numbers: Cape Verde, Italy |
| New Zealand | China<br>Korea |
| Switzerland | Former Yugoslavia<br>Portugal<br>Smaller numbers: Spain, Turkey |

Note: The way Balkans is defined differs slightly from country to country, as does the residual category ’other countries’