

Research Article

ARE MIGRANT STUDENTS CLOSING THE GAP? READING PROGRESSION IN THE FIRST YEARS OF MAINSTREAM EDUCATION

Nicole Marx* 

Institute of German Language and Literature II, University of Cologne

Christian Gill

*School of Languages, Literature, Cultures and Linguistics, Monash University
Melbourne*

Tim Brosowski 

Institute of Psychology and Cognition Research, University of Bremen

Abstract

Since 2015, increased numbers of newly immigrated schoolchildren in Europe have resulted in divergent, often ad hoc measures to provide for their education. Because the basis of classroom learning is information found in written texts, the development of grade-level reading skills is of central importance. However, little is known about immigrant students' reading skills at and following transition, and no data is available for Germany, where the study was conducted. We report the results of a longitudinal study in which migrant students' ($N = 136$) reading subskills after transition into mainstream were investigated at three points over the course of 2 years and compared to cohort performance ($N = 517$) in grades 7 through 9. Results showed that immigrant students performed significantly below mainstream students on all measures for all data points, with little evidence that they are beginning to close the gap even after several years in mainstream.

INTRODUCTION

Since 2015, increased numbers of newly migrated students in Europe have resulted in far-reaching challenges for educational systems. Although exact statistics are difficult to calculate, it is estimated that between 2014 and 2017, Germany took in more than 493,000 refugees under the age of 18, or 38% of all such refugees in Europe (Eurostat, 2019).

* Correspondence concerning this article should be addressed to Nicole Marx, Philosophische Fakultät, Institut für Deutsche Sprache und Literatur II, Albertus-Magnus-Platz 1, 50923 Köln, Germany. E-mail: n.marx@uni-koeln.de

Presently, in most states in Germany, these make up approximately 75% of all students who have immigrated during the course of their school education. Almost all the remaining 25% of immigrant students arrive through processes of European Union domestic migration. Depending on the state, immigrant students comprise up to 10% of the total student population, as is the case in Bremen (Die Senatorin für Kinder und Bildung, Bremen, 2019).

Despite a long history of migration to the country (Reich, 1980, 2017), the German school system was not well prepared for integrating larger numbers of new students (Decker-Ernst, 2017). Ad hoc approaches to educating immigrant students were developed, and Germany's decentralized education system resulted in highly diverse measures to provide for students' education (Massumi et al., 2015).

Within this system, the city-states of Bremen and Hamburg, both in northern Germany, have similar structures regarding the education of new immigrants in the first phase of secondary school, which encompasses grades 5 through 10, with attendees being between about 11 and 16 years old. After allocation to a city district, the children are assigned to a school—often, but not always, lower-streamed secondary programs.¹ There, they attend preparatory DaZ classes (Deutsch als Zweitsprache/German as a Second Language) for newly immigrated students for 20–25 hours a week, or approximately 700–800 classroom hours over the course of a year. Students remain in these classes, provided capacity is available at the school, for a full calendar year. Because students may enter school at any time during the school year, the courses show a high participant fluctuation. At the start of the present study, Bremen had more than 200 DaZ preparatory classes at all levels of schools (Die Senatorin für Kinder und Bildung, Bremen, 2019) and Hamburg had 225 such classes in grades 1–10 (Behörde für Schule und Berufsbildung Hamburg, 2016).

Regardless of educational model, the primary goal of education for newly arrived immigrants is the provision of sufficient language and academic support so that students can integrate successfully into mainstream education—that is, reach, as quickly as possible, academic levels achieved by nonimmigrants. As a result, many schools and states set specific language goals for DaZ students. Such guidelines are, however, arbitrary, in that they lack empirical evidence that this level is sufficient for school success. Further, they are not even compulsory. Bremen, for example, recommends that students achieve a CEFR (Common European Framework of Reference; Council of Europe, 2001) level of A2 in productive skills (speaking and writing) and B1 in receptive skills (listening and reading) before leaving preparatory classes. This is equivalent to being able to carry out everyday activities and understand simple texts related to personally relevant topics²—hardly sufficient for academic participation in the mainstream classroom. Moreover, even this basic proficiency level seems not to be met by students leaving the preparatory system. Recent data from Bremen suggests that very few students—about 12%—are attested with this minimal level of language proficiency (Reichert et al., 2020). Regardless of attainment, students pass into mainstream education at the completion of 1 year of preparatory classes. What happens next is unclear—but given the low level of language skills with which immigrant students begin mainstream education, it is not surprising that they often have difficulties in the regular school system.

In the following, research on linguistic and academic progression of immigrant students in Germany and abroad is discussed and the reasons for the dearth of research in Germany are explained. This is followed by a consideration of the importance of

reading in secondary schools before the aims, research questions, methods, and results of the study are presented. The article concludes with research implications.

The findings of the present study make three important contributions to the literature. First and foremost, information on reading trajectories of new immigrant students will help set a baseline for expectations of their progression in mainstream. Second, results can indicate whether certain variables are associated with higher or lower success in reading development. And finally, the study highlights the need for more research on these students and their success in the school system. Because the study is highly explorative in nature, its results should be interpreted with caution.

LINGUISTIC AND ACADEMIC PROGRESSION OF IMMIGRANT STUDENTS

A major assumption governing not only much of the media discourse but also educational policies and classroom pedagogy is that within a specific period, immigrant students will “catch up” to mainstream students, that is converge with monolingual norms. Consequently, second language education is often of limited duration and encourages full integration as quickly as possible. A central question regarding the education of immigrant students is thus: Do immigrant students close the gap to mainstream students and, if so, at what rate?

This crucial information is difficult to come by in Germany for various reasons. First, and in alignment with such students across Europe and North America, the population is highly diverse regarding such aspects as previous formal education, languages already acquired, and German language knowledge (Gill et al., 2019; Rick & Gill, 2015). This makes general statements about their academic progression difficult. Relevant for their success may be factors such as age at time of immigration (Clark-Gareca et al., 2019), literacy experience in the family language (L1) (Bialystok, 2007; Knapp, 1999), amount of time spent in the new school system (Soto-Corominas et al., 2020), participation in L1 classes after immigration (González, 1986), type of integration model (Collier & Thomas, 2017; Umansky & Reardon, 2014), relatedness of the L1 to the new L2 (Schepens et al., 2020), use of known L2s in learning the new language (Marx, 2005), and possibly refugee status (Woods, 2009), amongst others.

Second, as mentioned in the preceding text, educational models within Germany are decentralized and thus difficult to compare. Even within smaller states such as Bremen and Hamburg, length and content of DaZ education varies. According to official state guidelines, students are schooled in a partial integration model for a full year, meaning that they attend both segregated DaZ classes as well as certain mainstream courses such as music, physical education, or English. However, this is often not the case in reality (for a discussion of this problem pertaining to research, see Collier & Thomas, 2017). Students may first be fully submersed in mainstream classes and only afterward (or never) enter DaZ courses, or they may begin in DaZ classes but be moved into mainstream at short notice due to DaZ classes being overfilled. Individual schools integrate many students into mainstream based on institutional constraints rather than on length of time the student has spent in preparatory classes or students’ language and academic proficiency. Amongst the immigrant students in the present sample, for example, 21% attended preparatory classes for 6 months or less; 25% for more than 7 months and less than one school year; 44% for a full year; and 10% for more than one year—thus, only 54% of students completed the officially required time in DaZ classes.

Third, teachers of preparatory classes often have little or no training because DaZ is not a university subject in Germany. Quite often, young teachers are allocated to teach DaZ classes for which they receive little or no training, for which no textbooks or supplementary materials are available, and for which no supporting curricular guidelines exist (Decker-Ernst, 2017; Gamper et al., 2020). Such uncertainty associated with the education of new immigrants in German schools makes research especially difficult because schools and teachers are—understandably—not excited about granting researchers access to their classes and students.

Fourth, until recently, the number of new immigrants arriving in German schools each year was small enough not to warrant media attention—and thus little research interest since the last wave of migration in the late 1970s and early 1980s. Even then, the focus was primarily on acquisition of morphology and syntax but not on general educational success (e.g., Clahsen et al., 1983).

Fifth, participant attrition is unusually high because immigrant students may change schools at short notice due to factors such as forced relocation, reunion with family in a different state, or simply state- or city-initiated reorganization of school attendees. Finally, laws regarding data protection and privacy are quite stringent in Germany and involve numerous applications at various levels. For example, in the present project, almost a full year was necessary to get permission at all required levels, and for many pupils, parents still did not approve the processing of questionnaire data—despite concerted efforts that included communication with parents, provision of materials in the family languages, and the assurance of anonymity of all data. As well, factors such as socioeconomic status were not considered due to privacy constraints posed by the school authorities.

Considering these difficulties, information about the academic and linguistic progression of immigrant students mostly derives from studies outside of Germany. Large-scale longitudinal data on immigrant students' academic progression is, for example, available for the United States (Collier & Thomas, 2017). Models comparable to the German system are reviewed in these studies and are deemed lacking; these include pullout ESL (English as a Second Language) classes, early exit transitional language classes (the model propagated in Bremen and Hamburg), and even late exit transitional language classes. Regarding such models, the authors concluded: "Apparently, one year of support followed by placement in the English mainstream has about the same effect as no support at all" (Collier & Thomas, 2009, p. 78).

Directly relevant to school policy is the question of how long students need to attain grade-level norms, usually determined as being the 50th percentile reached by nonimmigrant students in diverse school subjects. Various time frames have been proposed. Most widely known and reported is Cummins's conclusion, originally based on data from California in the 1970s and since supported by various studies in different learning contexts, that a minimum of 5 to 7 years are necessary for immigrant students to converge with norms of their mainstream peers in school subjects (Cummins, 2000; Hakuta et al., 2000; Malakoff & Hakuta, 1991; Saunders & O'Brien, 2006).

However, even this may be a gross underestimation of expected learning trajectories, as recent studies suggest. These involve such different aspects as academic language skills (Levin & Shohamy, 2008; Umansky & Reardon, 2014), vocabulary (Cameron, 2002; Soto-Corominas et al., 2020), reading (Farnia & Geva, 2013), morphology (Paradis & Jia,

2017; Soto-Corominas et al., 2020), or attainment in content courses (Collier & Thomas, 2017; Levin & Shohamy, 2008). Depending on various factors, students may in fact need more than 10 years to catch up—or they may never do so, instead plateauing at a level significantly lower than the nonimmigrant average. This group of students has thus also been termed “long-term learners” (Clark-Gareca et al., 2019; Menken & Kleyn, 2010) in the North American context. This result is especially worrisome in light of the fact that many students will prospectively exit the school system before they have been able to complete 10 (or even 5) years of education.

Because a large quantity of school policy relies on (often erroneous) assumptions regarding the length of time that immigrant students need to reach grade-level norms, such results from English-speaking contexts are troubling.

The present study took on the question posed at the beginning of this chapter. It investigates the trajectories of language skills necessary for succeeding in an academic setting, focusing on reading skills, in the first years after leaving the preparatory system.

READING SKILLS IN SECONDARY SCHOOL

Reading skills were chosen as the dependent variable for the study. In contrast to oral language skills, which have received much attention in second language teaching since the advent of Communicative Language Teaching (Hymes, 1972), written skills are often overlooked by teachers. Especially following transition into mainstream education, when teachers are more focused on students’ ability to understand classroom discourse, such skills are less likely to be noticed (Knapp, 1999). However, this is a misconception; written texts are the primary source of information in almost all subjects starting at about grade 3. They are central for students’ knowledge expansion and language development—and thus also school success, which explains the high interest in reading and, especially, reading testing in the past few decades. In the context of L2 (second or additional language) learning, reading is the strongest supporter of vocabulary development throughout the learning process (Jeon & Yamashita, 2014). Although reading comprehension is related to oral skills in an L2, it also draws on a partially different subset of skills, as evidenced by the fact that they are not as closely associated in L2 reading as in L1 reading and that word-level reading skills are not well predicted by oral skills (Geva, 2006).

Although reading *literacy* has been a focus of testing since the advent of the OECD’s Programme for International Student Assessment (PISA) with its first wave of testing in 2000, reading *comprehension* has long been of central interest in the school classroom context and in research. In cognitive views of reading for understanding, reading comprehension is the ability to draw meaning from a printed or digital text and interpret this information appropriately. Comprehension is constructive in nature and relies on a number of interacting lower- and higher-level cognitive, linguistic, and nonlinguistic skills and subskills (Lenhard, 2013; Nassaji, 2014; Richter & Christmann, 2009).

The primary goal of reading a text is to develop and continually improve a situation model of the text (van Dijk & Kintsch, 1983). This is a coherent and nonlinguistic mental representation of the “state of affairs” described in a text and involves, for the most part, similar processes in first and other languages. To develop a situation model, higher-level processes (activation of prior knowledge, self-regulation such as use of reading strategies

and monitoring, construction of global coherence and inferencing) combine with lower-level processes, including visual word recognition (recoding a graphemic representation into a phonological representation of a word according to relevant grapheme-phoneme correspondences and decoding the meaning of a word through access to the mental lexicon) and local coherency building (which relies on semantic and syntactic processing and involves the development of propositions and identification of cohesion). These higher- and lower-level processes are mediated by reading fluency (sometimes understood to be part of lower-level processes; see Lenhard, 2013). Higher- and lower-level processes interact with and are dependent on general cognitive preconditions (basic cognitive skills, pace of access to the mental lexicon, phonological awareness, domain-specific prior knowledge, reading motivation, etc.) and conditions specific to reading in an L2 (such as L2 language skills, reading proficiency in L1) to allow a learner to interpret a text with more or less success (see also Bauerlein, 2014; Grabe, 2009; Lenhard, 2013; Nassaji, 2014; Wolff, 1987).

For L2 learners, including immigrant learners of a new language of schooling, any one of these skill subsets might cause difficulties in comprehending written texts (Jeon & Yamashita, 2014). For example, regarding higher-level processes, students may lack prior knowledge necessary to understand both the topic and specific elements in a text, such as intertextual references to German fairy tales in narrative texts or specific animals in biology texts commonly known to German schoolchildren (sparrows, Labrador retrievers, guinea pigs). At the same time, less experience with reading in general or lack of training in reading strategies may compound problems involving higher-level processes when reading strategies are unsuccessful or poorly implemented, or memory capacity does not allow for the construction of global coherence. Finally, lower-level processes are similarly jeopardized when individual lexical items cannot be quickly deciphered and identified, inhibiting the construction of local coherence.

Successful reading comprehension relies on the continuous interaction of a number of subskills that may play different roles at any given time, depending on the reader, the topic, the complexity of the text, and various other factors involved. For this reason, determining reading comprehension in schoolchildren is no easy matter. Even neglecting related skills (vocabulary breadth and depth, oral fluency, etc.), it should involve at least a measure of both lower-level reading processes and grade-level reading comprehension involving different text genres (Schneider, 2008). As well, for L2 learners, it is advisable to include a measure of L2 reading comprehension to incorporate age-appropriate content at lower levels of language complexity. If possible, these measures should not only be standardized but also normed on larger groups to establish comparable baselines.

The present study attempted to address these concerns while considering long-term learning trajectories in a student population difficult to study.

AIM OF THE STUDY

Given the aforementioned dearth of research on language and academic progression of newly immigrated students in German schools in general and the central importance of reading skills for academic success, we chose to investigate whether immigrant students are closing the gap to mainstream students with regard to reading skills within their first years of mainstream education. To achieve this, it was necessary to address a population

that (a) had already reached an age at which reading skills in the baseline population were developed, (b) were fully integrated in mainstream education, (c) could be followed for at least a 2-year time span (i.e., were young enough at the start of the study so that they would not exit the school system before the completion of the study, but were also old enough so that they did not change schools within the time frame of the study, both of which would result in high participant attrition), (d) had attended comparable schooling before mainstream (i.e., DaZ preparatory classes), and (e) lived in different school districts (to increase external validity). Moreover, (f) oversampling was a goal because immigrant students are exceptionally difficult to study in Germany and make up a small proportion of students. This led to specific considerations and constraints when attempting to gather participant data.

RESEARCH QUESTIONS

Four research questions are addressed to ask how immigrant students are progressing over a 2-year period, and which individual factors may be associated with better results over time.

(1) Have students achieved the recommended CEFR proficiency level of B1 in reading directly after exiting preparatory classes?

(2) With regard to both (a) lower-level reading processes and (b) reading comprehension, are immigrant students closing the gap to mainstream students in their first 2 years of mainstream education?

(3) Is there an association between time spent in mainstream education and reading achievement?

(4) Are specific individual variables associated with higher or lower achievement?

METHOD

SAMPLE

Students beginning grades 7 and 8 (average age at the start of the study in September 2016, Grade 7: 13.3, $SD = 0.8$; Grade 8: 14.2, $SD = 0.8$) were recruited for the study. This allowed for data collection over a 2-year time span while reducing the chance of participant attrition due to students exiting the school system (nonacademic secondary school ends after grade 9 or 10 in Germany). The focus was chosen for both grades 7 and 8 to maximize potential participant numbers and allow for higher external validity through the comparison of data over three different grades. These 2 years represent a more stable period of secondary education, following the transitional period of grades 5/6 (beginning of secondary school), but before the grades oriented at attaining a Middle School Leaving Certificate (German: *Mittlerer Schulabschluss*) begin. Recruitment was carried out at school level. More than 80 schools in the city-states of Bremen and Hamburg that offered DaZ preparatory classes were approached and asked if they would participate in the research project. Fifteen schools reacted positively and were included in the sample. Of these, 13 were lower-streamed secondary schools, reflecting the typical assignment of immigrant students to the nonacademic stream. Most schools allowed entire classrooms to participate, so that immigrant students' results could be nested within their cohort data.

Bremen and Hamburg were chosen as focus states, as they have similar integration practices and move students from full segregation in DaZ classes to partial integration and finally to full integration in the mainstream after 1 year. As well, they lie geographically close to one another and could thus be visited regularly by the researchers and research assistants, who were working in Bremen at the time of data collection. All regulations regarding student and school data protection and privacy were fulfilled in both states. This involved a four-step process of gaining permission from the state data-protection agency, the school and teachers involved, the individual parents of the students taking part, and the students.

For comparison purposes, participants were categorized into one of three groups. The reference group (mainstream students, or MS) was composed of students who had entered the German school system at the start of grade 1 and had never attended school in another country. Immigrant students in the group SE-1 (“Seiteneinsteiger,” a common noun denoting immigrant students) transitioned from preparatory classes into mainstream education at the beginning of the study. SE-2 immigrant students had already spent, on average, 17.0 months in mainstream education (*SD*: 11.9; range: 3–61 months). Although the average time in mainstream of the SE-2 students was highly divergent, their test results were quite homogeneous, supporting the decision to analyze them as one group. Only students for whom full data sets over 2 years are available were included in the final study. The distribution of the three groups studied is presented in [Table 1](#).

Of particular importance in the present study is that the reference group of mainstream students was composed solely of students who did not immigrate during their school career. However, 298, or 58%, of these students spoke diverse heritage languages (HL) at home. This is especially important in lieu of the fact that in Germany, mainstream students with HLs consistently score lower than their cohorts without HLs. For example, students with “migrant backgrounds” (i.e., second- or third-generation learners), who often speak heritage languages at home, score significantly below their cohorts in the PISA studies even after controlling for gender and schools’ and students’ socioeconomic profiles (Secretary-General of the OECD, 2019).³

This was supported in the present data, in which nonimmigrant HL speakers scored between a half and a full standard deviation below their non-HL speaking counterparts. However, because the study was primarily interested in comparisons of nonimmigrant students with immigrant students, the comparison group deliberately represented the actual composition of classrooms in Germany to ensure ecological validity.

The 136 immigrant students in the sample hailed from 33 different countries, although the three most common countries of origin were Syria (30 students), Bulgaria

TABLE 1. Distribution of students in the sample

Starting grade in September 2016	Grade 7	Grade 8	Total
SE-1: Immigrant students beginning mainstream education, 9% in Gymnasium	30	24	54
SE-2: Immigrant students already in mainstream education, 7% in Gymnasium	32	50	82
MS: Mainstream students (reference group in same classes), 15% in Gymnasium	287 (169 HL)	230 (129 HL)	517 (298 HL)

(14 students), and Poland (10 students); all other countries were named by seven or fewer students (for nine students, questionnaire data is missing). This reflects the present character of immigrant students in general in Germany.

MATERIALS

The full study involved both quantitative and qualitative components with differing instruments. For the purposes of this article, three reading test instruments⁴ were used. At each testing time, a different version of each was administered to avoid inflated results due to retesting. The first instrument was the Reading Comprehension subset of the DSD I (Deutsches Sprachdiplom/German Language Certificate, Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany [KMK], n.d.), a national exam originally developed to assess language skills of students studying at German-language schools outside of Germany and thus intended for L2 learners. Unfortunately, the KMK refuses to release reliability measures, and internal testing showed little consistency; for the SE-1 students, a split-half reliability test resulted in a low, albeit significant, correlation ($r = .30$, $p = .034$). However, due to both the importance of the DSD I in the school system—outside of Germany, it is a high-stakes exam, as it may be used to determine entrance into university, and it is recommended by both Bremen and Hamburg for students exiting preparatory classes—and to the dearth of similar tests, it was included in the present study. Of primary interest is whether students have achieved the recommended language level of B1. It involves two subtasks typical in L2 contexts: a Cloze task and a reading text with multiple choice comprehension questions with a maximum of 11 points. This test was used to address RQ1.

The second instrument involved different subtests of the normed and standardized reading battery LESEN 6–7 (for grades 6 and 7; Bäuerlein et al., 2012a) and LESEN 8–9 (for grades 8 and 9, Bäuerlein et al., 2012b). Lower-level reading processes (involving subskills of recoding, decoding, and reading fluency; German: *basale Lesefertigkeit*) were investigated with Subtest 1. This comprised a sentence judgment task consisting of a total of 100 sentences that the test taker is required to judge as correct or incorrect as quickly as possible. A total of 3 minutes is allotted for the test. Reading comprehension (German: *Leseverständnis*) of expository (biology) texts was investigated with Subtest 2.1, of narrative texts with Subtest 2.2. These tests were developed to mirror common academic language demands both within and outside of language arts classes and thus show high external validity. Both tests contain a one-page written text (approximately 45 lines) followed by 17 (LESEN 6–7) or 19 (LESEN 8–9) multiple-choice comprehension questions with five possible solutions each, of which only one is correct. Because only two subtest versions were available, but students were tested at three points, the version at the second testing for both narrative and expository texts was developed from grade-level textbooks and modeled on the tests in LESEN. A full description of the test construction, norming, item analyses, reliability measures, and model conformity can be found in Bäuerlein (2014). For the present group of participants, test-retest reliability was measured for the first two data points and was close enough to 0.7 to be considered acceptable for the purposes of this study ($r = .78$ for lower-level reading processes, $r = .65$ for expository texts, and $r = .65$ for narrative texts; all $p < .001$). These tests were implemented to address RQ2, 3, and 4.

Immigrant students also completed a comprehensive questionnaire on language use, school career, and further relevant information. Mainstream students completed a more compact version focusing on language use and reading preferences. The questionnaire data was used to provide more information on the population tested, control for possible confounding variables, and answer RQ4.

PROCEDURE

All students completed the reading exercises in their regular classrooms during German class. To expedite data collection and analysis, and reduce the strain on the schools, classrooms without immigrant students were not included in the study. Tests were administered by trained research assistants to ensure procedural consistency and that participants understood each task. Students did not receive any further support once the testing phase had begun. Students were informed at each testing point that participation was voluntary and that results would not be shared with their teachers, parents, or schools. All tests began with the development of a student-specific pseudonym, which could be consistently replicated but not traced back to individual participants.

The study involved five data-collection points as well as follow-up testing within 1 week to include students who missed the tests due to absenteeism. Students who missed both testing points of any one task were excluded from the final analyses. At the first data collection point (“time,” i.e. t1), all students completed the short form of the questionnaire before continuing with the reading tasks. The DSD I reading subtask was administered once at the beginning of the study to estimate students’ L2 reading comprehension. Although all students took part, only the results of the SE-1 group were analyzed because the test is specified for students exiting DaZ classes into mainstream. Total time for all aspects was approximately 1.5 hours. The group of SE students completed the longer questionnaire separately later on the same day. To reduce the testing load for students and schools, and to allow for noticeable learning gains, each of the standardized reading instruments was staggered during the first year. Lower-level reading processes were tested in September 2016 (t1) and June 2017 (t4). Reading comprehension of expository texts was tested in September 2016 (t1) and March 2017 (t3), while reading comprehension of narrative texts was tested in December 2016 (t2) and June 2017 (t4). All three tests were administered a third and final time at the end of the second year of testing, in June 2018 (t5).

DATA CODING

Due to the nested character of the data and thus possible confounding effects of variables such as classroom and school—which often reflect further social variables due to location of residence within a particular neighborhood—as well as effects of gender, the data was z-standardized and transformed based on means and SDs of these variables (see the following text). All results were entered in SPSS 25 before data analysis. Measures were coded by trained research assistants.

Questionnaire

Quantitative questionnaire data was double-coded according to guidelines in a codebook developed for the study. In addition, for RQ3, the variable “Time spent in mainstream education at start of study” (used to categorize students into SE-1 and SE-2 groups) was recoded as a continuous variable reflecting the number of months that each student had been in mainstream education at the specific point of testing.

DSD I

The DSD I reading comprehension task was coded using only raw scores. Students were judged to have reached the level B1 if at least 8 (of max. 11) points were achieved.

Lower-Level Reading Processes and Reading Comprehension Tests (Expository and Narrative)

Although these tests are normed for grade levels and school type in Germany, it was deemed necessary to renorm the data based on the populations studied, that is the group of students in the states of Bremen and Hamburg at the time of data collection. First, baseline scores were computed. Because the mainstream student data showed significant group differences regarding grade, gender, and school type, scores for eight subgroups (2 genders \times 2 grades \times 2 school types) were computed to reflect the mean and standard deviation of the mainstream students for each test and test time. Individual scores were then z-standardized according to results of the corresponding reference group for each test time and test. This minimized possible confounding factors due to population differences. The standardized data was normally distributed. To prevent outlier bias, data was winsorized, that is data points outside of the 1st and 99th percentile were replaced with values at the 1st and 99th percentiles, respectively.

DATA ANALYSIS

Questionnaire Data

Data was analyzed according to the specific research question and levels of measurement involved.

DSD I

Because the intention of the DSD I was only to provide a ballpark estimate of whether students had reached the recommended B1 level at the beginning of their integration into mainstream education and the test is not normed (although centrally developed and standardized), only descriptive statistics were calculated for the SE-1 participants on this test. Due to plateau effects, the DSD I results were not included for the other students.

Lower-Level Reading Processes and Reading Comprehension Tests

Generalized Estimating Equations (GEEs) were used to calculate the effects of group (SE-1/SE-2) on the z-standardized test values over time. These allow the use of general linear models on repeated-measure data sets with differently distributed dependent variables and different predictor variables. For the standardized tests used in the present study, normally distributed, metric dependent variables were analyzed using the identity link function. For each of the eight dependent variables, both a main effect of group and an interaction effect between group and time was calculated (due to the z-standardization of results, it was not logical to investigate a main effect of time alone).

To calculate the GEEs, three different models were tested for each data set with robust model estimation that assumed different autocorrelation assumptions (due to the repeated measures design, this step was deemed necessary; see also Liang & Zeger, 1986 on the expansion of the General Linear Model). The best fit for all outcomes was the assumption of independence of data. Thus, the independence model parameter was chosen for data interpretation for all measures.

The association between time spent in mainstream education and reading skills (RQ3) was calculated with Pearson's product-moment correlations using bias corrected and accelerated bootstrapping for the three dependent variables of interest.

To specify the effects of further covariates on the individual test results (RQ4), explorative post hoc residual analyses were carried out. Nonparametric Kruskal–Wallis tests were chosen for residuals with nominal variables. Spearman's rho correlation coefficients were calculated for residuals with ordinal and metric data.

RESULTS

L2 READING COMPREHENSION DIRECTLY AFTER TRANSITION INTO MAINSTREAM (RQ1)

Because the DSD I was administered to investigate students' level directly after entry into mainstream, only the SE-1 data, that is data for those students who had transitioned into mainstream in September 2016, are reported here. Results are available for 49 SE-1 students.

Students achieved an average of 5.6 out of 11 total points on the test ($SD = 2.10$). Only 10 students, that is 20%, reached 8 points, the cutoff for a level of B1. Twenty-two students—almost half—achieved 5 points or less; none achieved the maximum number of points (Figure 1). This indicates that most students had not reached the recommended CEFR proficiency level of B1 in reading, despite having already been reclassified into mainstream education.

READING SKILLS OF IMMIGRANT STUDENTS COMPARED TO MAINSTREAM STUDENTS (RQ2)

RQ2 represented the issue of central importance to the study. The results are presented separately for each of the three measures considered.

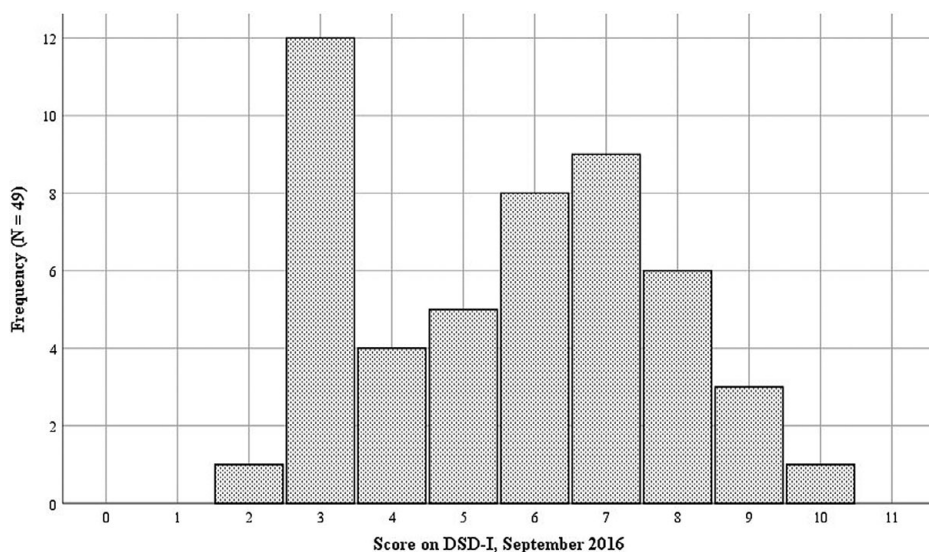


FIGURE 1. Frequency of correct answers on the L2 reading subtest of the DSD I written by SE-1 students beginning mainstream education.

Note: A score of 8 or above represents having reached the recommended level of B1.

Lower-Level Reading Processes

Tests were administered in September 2016 (t1), June 2017 (t4), and June 2018 (t5). Results of the GEEs showed that, at all times, students in both immigrant groups did not achieve a level even approaching that of their mainstream cohorts. Beta values of students in the SE-1 group (reference group at t1: mainstream students; reference group at following two times: within-group at t1) were, at the three test points respectively: $\beta = -2.16, p < .001$; $\beta = 0.74, p < .001$; and $\beta = 0.62, p < .001$. However, compared to September 2016, students did improve significantly by June 2017. Students in the SE-2 group were slightly better compared to SE-1 students, but still far below expected levels for their grade levels ($\beta = -1.47, p < .001$; $\beta = 0.22, p = .076$; $\beta = 0.20, p = .236$); there was no significant change in their scores over time. In general, the immigrant students in this sample thus seemed to improve over their first year in mainstream education, but then to plateau at about -1.5 to -1.2 *SD* compared to nonimmigrant students (exact beta values see Appendix Table A1. N.B. Note that the reference group for the analyses of the interaction between time and group at t4 and t5 is the group, i.e. SE-1 or SE-2, and not MS). The progress over time is presented in Figure 2.

Reading Comprehension of Expository (Biology) Texts

Tests were administered in September 2016 (t1), March 2017 (t3), and June 2018 (t5). Results of the GEEs showed that both immigrant groups performed significantly poorer than mainstream students at all data collection points. Students who had just transitioned into mainstream (SE-1) had beta values of $\beta = -1.16, p < .001$; $\beta = -0.41, p = .010$; and

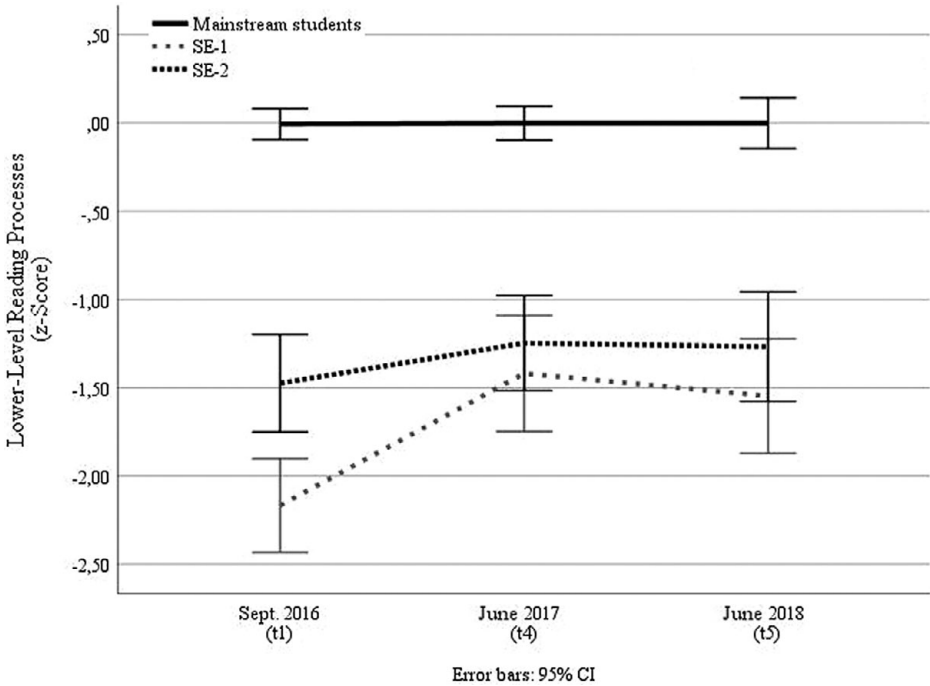


FIGURE 2. Standardized score of lower-level reading processes.

$\beta = 0.08, p = .612$. SE-2 students performed slightly better, but still far behind MS ($\beta = -0.93, p < .001$; $\beta = -0.40, p = .021$; $\beta = 0.13, p = .421$). Over 2 years, immigrant students did not close the gap on mainstream students regarding reading comprehension of expository texts (Figure 3; see also Appendix Table A2).

Reading Comprehension of Narrative Texts

Tests were administered in December 2016 (t2), June 2017 (t4), and June 2018 (t5). Again, near the start of the study, immigrant students performed significantly below their mainstream counterparts. Although both groups improved within the first year, they appeared to plateau, that is in the second year of the study, they did not make any further appreciable gains on the mainstream students. With reference to MS, SE-1 students showed values at t1 of $\beta = -1.63, p < .001$, but did not make substantial gains on their scores thereafter (at t4: $\beta = 0.54, p = .003$ and at t5: $\beta = 0.586, p = .032$). SE-2 students were only minimally (and not significantly) better than SE-1 students ($\beta = -1.52, p < .001$; at t4: $\beta = 0.63, p < .001$; t5: $\beta = 0.53, p = .008$) (Figure 4; see also Appendix Table A3).

READING ACHIEVEMENT AS A FACTOR OF TIME (RQ3)

RQ 3 asked if students improved with increased time in mainstream education. The independent (metric) variable was time spent in mainstream classes at time of testing. As

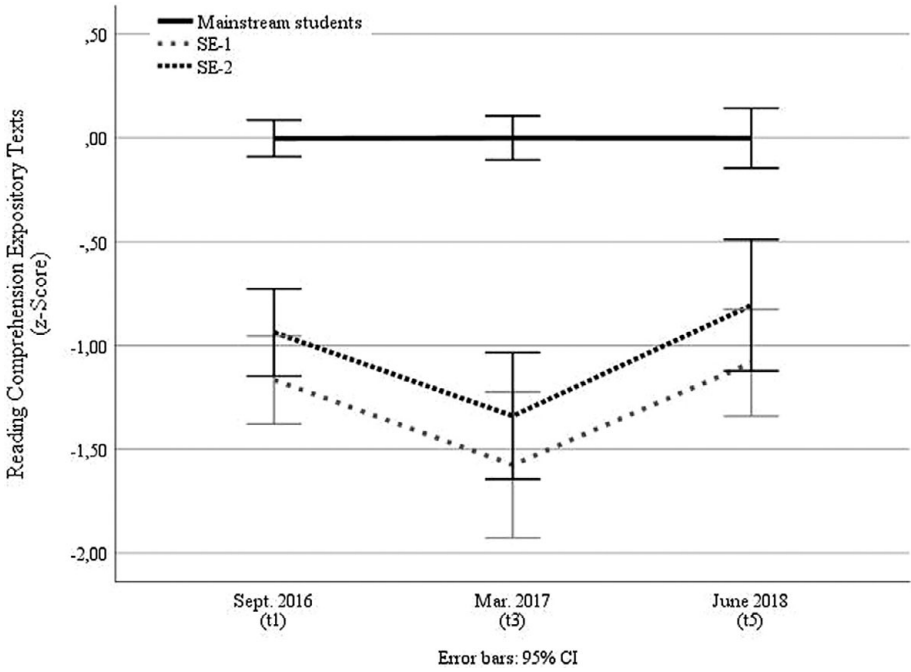


FIGURE 3. Standardized score of reading comprehension of expository texts.

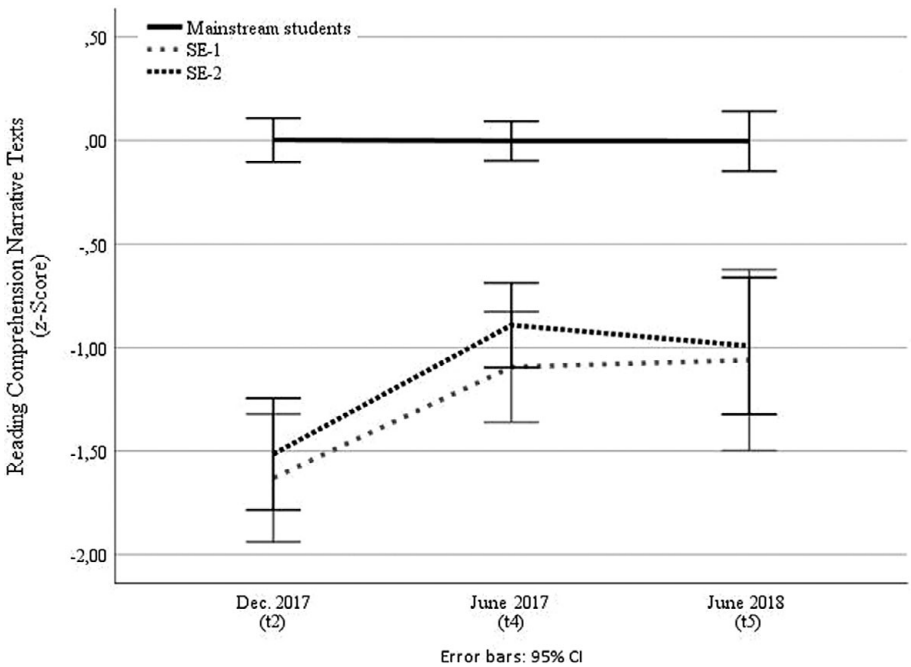


FIGURE 4. Standardized score of reading comprehension of narrative texts.

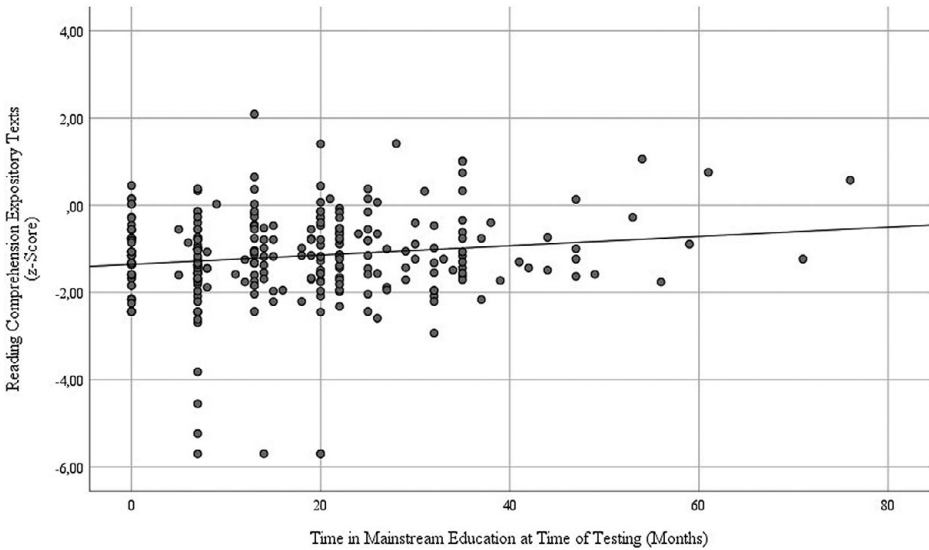


FIGURE 5. Relationship between time spent in mainstream education at time of testing and standardized test score.

for the purposes of RQ2 immigrant students were grouped into two categories without accounting for time as a metric variable, this question required further data analyses (see preceding text for an explanation of the recoding of data). The relationship between time and test was calculated as a Pearson correlation coefficient for each of the three relevant dependent variables. Bias corrected and accelerated bootstrap 95% CIs are reported in the following text in square brackets.

Results showed that all dependent variables were significantly correlated with time spent in mainstream, but the effects are small. Of all, the strongest relationship was in lower-level reading processes, $r = .251$, $p < .001$, $R^2 = .063$ [.145, .356]. Correlations between time and reading comprehension of expository texts ($r = .144$, $p = .015$, $R^2 = .021$ [.042, .244]) and narrative texts ($r = .165$, $p = .006$, $R^2 = .027$ [.052, .273]) were even lower. Thus, the common variance of reading measure and time was only 2–6% and can be neglected as a contributing factor to students' performance. To illustrate this, the scatterplot for expository texts is presented in Figure 5.

ASSOCIATION OF ADDITIONAL VARIABLES WITH HIGHER OR LOWER ACHIEVEMENT (RQ4)

To investigate the influence of additional covariates on the performance of immigrant students, standardized residuals per person and time of measurement were calculated for possible associations with 54 further possible covariates. These included such factors as city of residence, type of migration, reading interests, L1 literacy, language preferences, and length of schooling in the country of origin. Because these analyses were of exploratory nature, the alpha error accumulation was not adjusted. Only effects of at least medium size are reported here, as small effect sizes are generally less relevant to

pedagogical practice. Thus, nominal covariates with an effect size of $\eta^2 \geq 0.06$ and $p < 0.05$, and ordinal and metric covariates (both treated as ordinal variables in the calculations to minimize Type 1 error) with a rho value of at least $\rho = 0.25$ and $p < 0.05$ are included in the following results. Exact values are given in the tables in the Appendix (nominal variables: Table A4; ordinal and metric variables: Table A5).

Of the variables investigated, only five showed relevant effects on students' test results. For both the SE-1 and the SE-2 groups, both *age at the start of the study* and the *length of schooling in the country of origin* were significant negative correlates of results for lower-level reading processes. However, these variables were significantly intercorrelated (older students tend to have spent more time in school in the home country) at $r = .28$.

For the SE-1 group, *age* was also negatively associated with the results on reading comprehension of both expository and narrative texts, and with L2 reading comprehension (DSD I). As well, students hailing from a *crisis region* (i.e., a region for which refugee status may be granted) had significantly worse results in lower-level reading processes than the students who had not emigrated from a crisis region.

For the SE-2 group, *female* students scored significantly better on reading comprehension of narrative texts than male students—an effect also commonly seen amongst nonimmigrant students of similar ages (see, e.g., the most recent PISA results for reading; Secretary-General of the OECD, 2019, p. 16). As well, the *length of schooling in the country of origin* was significantly and negatively associated with both reading comprehension of narrative texts and foreign language reading comprehension (DSD I).

Finally, students in the SE-1 group who had spent *more time in preparatory DaZ classes* scored better on the L2 reading comprehension task (DSD I).

In sum, younger students and students who spend more time in preparatory classes might have an advantage on the reading measures evaluated here, whereas students originating from a crisis region may be at a disadvantage. The effects of age of testing and length of schooling in the country of origin are intercorrelated and thus warrant further investigation.

DISCUSSION

The results of the present study shed light on immigrant students' academic reading progression in the school system in their first years of mainstream education.

First, immigrant students entering mainstream classes after DaZ preparatory classes do not achieve even the minimal recommended foreign language reading level of B1. Of the 49 new students for whom data was available, only 10 seemed to reach B1—a level that is not sufficient for engaging with written language beyond that needed for everyday situations using high-frequency vocabulary. This aligns with unofficial data from the state of Bremen, as reported previously.

Second, immigrant students lag far behind their mainstream peers in lower-level reading processes and reading comprehension of both narrative and subject-specific texts. This applies both to students in their first 2 years of mainstream education, which comprised approximately half the immigrant population in the present study, and to immigrant students who had already attended school for, on average, 1.5 years before the study began. At all testing points, students remained over one standard deviation behind their nonimmigrant peers regarding lower-level reading processes and hovered around

one standard deviation below them regarding reading comprehension of both expository and narrative texts. This corresponds to international data, most notably the longitudinal research by Collier and Thomas (2017) mentioned in the preceding text. These results are especially disquieting given the fact that the majority of educational content at these grades (7–9) is delivered through written texts. Combined with the fact that German students in general do not rank amongst the highest in reading literacy in international studies (cf. the results of the most recent PISA study in which they scored just below the average, at 498; Secretary-General of the OECD, 2019, p. 17), this does not bode well for their future school success.

Third, it appears that students are also not improving further after the first years, as evidenced by the analyses with time spent in mainstream as a metric independent variable. Instead, they maintain a consistent gap behind mainstream students, regardless of whether they have been in mainstream education for only a few months or up to 7 years. Thus, whilst they are improving, they are not doing so at a rate that would indicate that they will at any time be able to close the gap on their nonimmigrant cohort. This result is in line with both broader quantitative studies on students' academic progression on standardized tests (Collier & Thomas, 2017; Cummins, 1979; Genesee, 1987) and with more recent studies on specific populations (Farnia & Geva, 2013; Levin & Shohamy, 2008; Paradis & Jia, 2017; Soto-Corominas et al., 2020). Reasons for this plateau in comparison to nonimmigrant students can only be guessed at; however, it is safe to assume that the dearth of continued German language support both within and outside of mainstream classrooms is not aiding the acquisition of the language.

Fourth, the exploratory residual analyses uncovered certain covariates that are related to better or poorer achievement of immigrant students. While these analyses are tentative, relying on the data available, it does seem that the younger students in the study—that is those starting grade 7 at the beginning of the project—performed better in comparison to their peers than the students just one grade level above them. Note that this does not imply that grade 7 immigrant students' reading was better than that of grade 8 immigrant students', but rather merely that the discrepancy between immigrant and mainstream students was not as high for students in grade 7. This is understandable, as students are originally schooled in the same DaZ classes, but then advance to different grade levels depending on their age. If students progress, on average, similarly within the same preparatory class, but are then integrated at different levels in schools, they might perform more poorly in comparison to their nonimmigrant peers at higher grades than at lower grades.

In addition to age, length of schooling in the country of origin was associated with poorer results on some measures. Due to this being correlated with age (as discussed in the preceding text), this result is difficult to interpret. As well, a number of further factors may play a role, such as differences between school systems attended prior to migration (Levin & Shohamy, 2008), leading to uncertainty about classroom participation and learning goals in the new country. As well, benefits of more education in the L1 may only appear in the medium, but not in the short term, as suggested by previous research (Genesee, 1987; Umansky & Reardon, 2014).

Finally, length of time spent in DaZ courses before entering mainstream may also—positively—affect outcomes. This corresponds to Collier and Thomas's (2009) conclusion quoted in the preceding text that only one year of language instruction has little or no

value in supporting new immigrant students. As with the other covariates, however, this needs to be studied with in-depth analyses, which as yet have not been carried out in the German context.

LIMITATIONS

Although various measures were implemented to increase internal and external validity of the study, and objectivity of data analysis, some limitations do apply.

In addition to population effects due to having only been carried out at 15 schools in two states, five main limitations of the present study can be noted. First, despite concerted efforts to increase the number of students studied, the sample size remains relatively small. Although this is understandable given that the study was dealing with a specific and small cohort with high rates of participant attrition, it does mean that analyses should be interpreted with caution.

Second, with the exception of school form and city, covariates are based solely on student report data. For this reason, questions that students might not be able to answer (such as highest education attained by parents) or desire to answer truthfully (such as those pertaining to socioeconomic status) were excluded from the study. This strengthened external validity and, at the same time, avoided items that might make students uncomfortable (and thus less willing to complete the questionnaire or to answer truthfully). Despite these precautions, self-report data should be interpreted with caution.

Third, no reading test can represent a full view of reading, reading literacy, and reading comprehension. Tests used here were chosen for being expeditious, having either already been validated through measures of reliability or being standardized and normed for the relevant population, and for showing external validity due to their similarity with classroom reading activities for the age group tested.

Fourth, and related to the third limitation, more comprehensive data is necessary to allow for broader conclusions about the language and academic progression of new immigrants in the school system. While large-scale studies might be preferable in this regard, limitations, especially regarding protection of personal data, posed within the German system make this seem unlikely.

Finally, the study was limited to the length of 2 school years, whereby in the second year, only one data-collection point (end of year) was possible. This allows for a certain time frame in which growth in reading skills may be followed, but not enough to truly map growth curves over the full time students need to reach grade-level norms.

CONCLUSIONS AND OUTLOOK

The results reported here shed some light on a previously under-researched area of the German school population. At the same time, they give some rather negative information about the academic progression of new immigrants in secondary schools.

The results, coupled with the lack of continuing language and subject support for immigrant students, raise troubling questions for the academic careers of immigrant pupils in Germany. While not surprising, considering studies from North America reported in the preceding text, they do allow the question to be posed of whether

expectations of learning trajectories reflect reality. A number of conclusions can be drawn from the results reported here.

First, there appears to be a discrepancy between government guidelines and practice. New immigrant students are allotted a specific amount of time in preparatory classes, during which they should learn enough of the language to be able to participate in mainstream schooling. Often, these classes are the only language support that students are certain to receive; depending on funding, teacher and classroom availability, and various other factors, they may or may not attend further classes. In the present study, for example, less than half continued to receive DaZ support after transitioning into mainstream. Of those with continuing DaZ classes, almost none attended more than 3 hours weekly. In fact, they may not even have spent the recommended time deemed necessary for them to transition into mainstream, as evidenced by the fact that only half of the immigrant students in this sample completed the recommended year of preparatory classes.

Second, and possibly in part due to the discrepancy discussed in the preceding text, many immigrant students in this sample entering mainstream classes do not reach even the recommended L2 reading level of B1. Considering the minimal level of reading skills necessary to reach this level, this result is troubling.

Third, immigrant students' performance does not converge with mainstream norms within their first years of education. This is evident on all three measures of reading conducted: lower-level reading processes and reading comprehension of both biology and narrative texts. Immigrant students are attempting to catch up to the moving target of mainstream students—and they are failing. Over the 2-year time span of the study, progression of z-standardized scores was minimal or nonexistent. Even more troubling, immigrant students do not appear to close the gap with increased time spent in mainstream education beyond the first 2 years. Although longer-term learners do (moderately) outperform students who have just recently transitioned into mainstream, the relationship between time spent in the education system and convergence with nonimmigrant norms is negligible. This holds even for students who had been in mainstream education for a full 7 years by the completion of the study.

Fourth, the low scores of immigrant students cannot be simply due to their bilingualism. Fifty-eight percent of the reference group of mainstream students were (nonimmigrant) heritage language speakers, meaning that this possibly confounding variable was accounted for by choosing an ecologically valid comparison population.

Fifth, certain covariates appear to support (younger age and a longer time spent in preparatory classes) or hinder (origin from a crisis region) students' reading development. Although these results are explorative in nature, they are worth noting—especially regarding those aspects that might be directly relevant to policy, such as time spent in preparatory classes.

Finally, further longitudinal research on immigrant students' learning trajectories is imperative if these students are to be academically successful. Very little is known about this highly diverse group of learners, their learning trajectories, and the various aspects that affect their learning progression, and this dearth of information contributes to the general inability to develop proper curricula, materials, and integration plans.

In summary, it appears that the belief that newly immigrated students can begin to close the gap on mainstream students after only 1 year (or less) of preparatory language classes

is overly optimistic. Immigrant students to Germany might not reach grade-level norms even after up to 7 years in the system. In an educational system like that in Germany, where students in nonacademic streams finish school after grade 9 or 10, this means that a large proportion of the immigrant student population may never be granted the opportunity to catch up academically to their nonimmigrant peers before leaving the educational system for good.

NOTES

¹Germany's education system is organized so that, upon leaving primary school after grade 4, students are allocated to different academic streams. These vary from state to state, but only the higher stream, the "Gymnasium," allows students to continue with college or university education. The other streams ("lower-streamed," in this publication the "Oberschule" in Bremen and the "Gesamtschule" in Hamburg) end after grade 9 or 10, after which students may leave the system or continue with vocational training. Although transfer between streams is technically possible, current data shows that, once a student has entered a lower-level secondary school in grade 5, they have very few chances of moving up to a Gymnasium or of attending postsecondary education at the end of their school career.

²The CEFR describes reading at this level as being able to understand "texts that consist mainly of high frequency everyday language" and "the description of events, feelings and wishes in personal letters" (Council of Europe [2001]: 27).

³N.B. In Germany, new immigrants do not take part in the PISA study until they have been in mainstream education for at least 1 year and do not have potential emotional reasons not to take part (Reiss et al., 2019, p. 17).

⁴The DSD I and the internally developed questionnaire as well as procedural and coding manuals can be made available to interested readers upon request to the corresponding author. Due to their length, they are not included in the Appendix. The tests from LESEN 6–7 and LESEN 8–9 are under copyright protection and can be requested through the publisher.

REFERENCES

- Bäuerlein, K. (2014). *Leseverständnisdagnostik in der Sekundarstufe*. Würzburg University Press.
- Bäuerlein, K., Lenhard, W., & Schneider, W. (2012a). *LESEN 6-7. Lesetestbatterie für die Klassenstufen 6-7. Verfahren zur Erfassung der basalen Lesekompetenz und des Textverständnisses*. Hogrefe.
- Bäuerlein, K., Lenhard, W., & Schneider, W. (2012b). *LESEN 8-9. Lesetestbatterie für die Klassenstufen 6-7. Verfahren zur Erfassung der basalen Lesekompetenz und des Textverständnisses*. Hogrefe.
- Behörde für Schule und Berufsbildung Hamburg. (2016). *Basisklassen und Internationale Vorbereitungsklassen an allgemeinbildenden Schulen im Schuljahr 2016/17 (ohne Sekundarstufe II, berufliche Bildungsgänge an Sonderschulen und Erwachsenenbildung)*. <https://www.hamburg.de/contentblob/8163550/1a20387a8e69d053845d156f334afb3/data/2016-17-klassen.pdf>
- Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, 10, 210–223.
- Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research*, 6, 145–173. <https://doi.org/10.1191/1362168802lr103oa>.
- Clahsen, H., Meisel, J., & Pienemann, M. (1983). *Deutsch als Zweitsprache: Der Spracherwerb ausländischer Arbeiter*. Narr.
- Clark-Gareca, B., Short, D., Lukes, M., & Sharp-Ross, M. (2019). Long-term English learners: Current research, policy, and practice. *TESOL Journal*, 49, 1–15. <https://doi.org/10.1002/tesj.452>.
- Collier, V., & Thomas, W. (2009). *Educating English learners for a transformed world*. Fuente Press.
- Collier, V., & Thomas, W. (2017). Validating the power of bilingual schooling: Thirty-two years of large-scale, longitudinal research. *Annual Review of Applied Linguistics (ARAL)*, 37, 203–217.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press. <https://rm.coe.int/1680459f97>.

- Cummins, J. (1979). *Research findings from French immersion programs across Canada: A parent's guide*. Canadian Parents for French.
- Cummins, J. (2000). *Language, power and pedagogy. Bilingual children in the crossfire*. Multilingual Matters.
- Das Deutsche Sprachdiplom I (DSD-I). (n.d.). *Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany*. <https://www.kmk.org/themen/deutsches-sprachdiplom-dsd.html>
- Decker-Ernst, Y. (2017). *Deutsch als Zweitsprache in Vorbereitungsklassen. Eine Bestandsaufnahme in Baden-Württemberg*. Schneider Verlag Hohengehren.
- Die Senatorin für Kinder und Bildung, Bremen. (2019). *Zahlen und Fakten*. https://www.bildung.bremen.de/zahlen_und_fakten-117153
- Eurostat. (2019). *Asylum and first time asylum applicants by citizenship, age and sex. Annual aggregated data (rounded)*, December 3. https://ec.europa.eu/eurostat/en/web/products-datasets/-/MIGR_ASYAPPCTZA
- Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, 27, 389–421. <https://doi.org/10.1111/jrir.12003>.
- Gamper, J., Röttger, E., Steinbock, D., & Falke, U. (2020). "Aber es ist insgesamt zu wenig und es fehlt ein klares Konzept"—Willkommensklassen in Berlin. *Informationen Deutsch als Fremdsprache*, 47, 410–428. <https://doi.org/10.1515/infodaf-2020-0064>
- Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education*. Heinle & Heinle.
- Geva, E. (2006). Second-language oral proficiency and second-language literacy. In D. August & L. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth* (pp. 123–139). Erlbaum and the Center for Applied Linguistics.
- Gill, C., Marx, N., Reichert, M.-C., & Rick, B. (2019). Der Unterricht von Deutsch als Sprache der Bildung in Vorbereitungsklassen. Ein curricularer Vorschlag für die Sekundarstufe I. *Deutsch als Fremdsprache*, 56, 152–161.
- González, L. A. (1986). *The effects of first language education on the second language and academic achievement of Mexican immigrant elementary school children in the United States*. University of Illinois at Urbana-Champaign.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* The University of California Linguistic Minority Research Institute Policy Report 2000–1. <https://web.stanford.edu/~hakuta/Publications>
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates. A meta-analysis. *Language Learning*, 64, 160–212. <https://doi.org/10.1111/lang.12034>.
- Knapp, W. (1999). Verdeckte Sprachschwierigkeiten. *Die Grundschule*, 5, 30–33.
- Lenhard, W. (2013). *Leseverständnis und Lesekompetenz: Grundlagen—Diagnostik—Förderung*. Psychologie 2014. W. Kohlhammer.
- Levin, T., & Shohamy, E. (2008). Achievement of immigrant students in mathematics and academic Hebrew in Israeli school: A large-scale evaluation study. *Studies in Educational Evaluation*, 34, 1–14. <https://doi.org/10.1016/j.stueduc.2008.01.001>.
- Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13–22. <http://www.biostat.jhsph.edu/~fdominic/teaching/bio655/references/extra/liang.bka.1986.pdf>.
- Malakoff, M., & Hakuta, K. (1991). Translation skills and metalinguistic awareness in bilinguals. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 141–157). Cambridge University Press.
- Marx, N. (2005). *Hörverstehensleistungen im Deutschen als Tertiärsprache: Zum Nutzen eines Sensibilisierungsunterrichts in "DaF/E"*. Schneider Verlag Hohengehren.
- Massumi, M., von Dewitz, N., Griebach, J., & Terhart, H. (2015). *Neu zugewanderte Kinder und Jugendliche im deutschen Schulsystem Bestandsaufnahme und Empfehlungen*. Universität zu Köln.
- Menken, K., & Kleyn, T. (2010). The long-term impact of subtractive schooling in the educational experiences of secondary English language learners. *International Journal of Bilingual Education and Bilingualism*, 13, 399–417.
- Nassaji, H. (2014). The role and importance of lower-level processes in second language reading. *Language Teaching*, 47, 1–37.

- Paradis, J., & Jia, R. (2017). Bilingual children's long-term outcomes in English as a second language: Language environment factors shape individual differences in catching up with monolinguals. *Developmental Science*, 20, 1–15. <https://doi.org/10.1111/desc.12433>.
- Reich, H. (1980). Deutschlehrer für Gastarbeiterkinder. Eine Übersicht über Ausbildungsmöglichkeiten in der Bundesrepublik. In Deutscher Akademischer Austauschdienst. Arbeitskreis Deutsch als Fremdsprache (Ed.), *Materialien Deutsch als Fremdsprache. Diskussion eines Rahmenplanes "Deutsch als Fremdsprache an den Hochschulen."* Vorträge und Materialien der Jahrestagung "Deutsch als Fremdsprache" an der Universität Regensburg vom 7–9 Juni 1979 (pp. 187–213). AKDaF.
- Reich, H. (2017). Geschichte der Beschulung von Seiteneinsteigern im deutschen Bildungssystem. In M. Becker-Mrotzek & H.-J. Roth (Eds.), *Sprachliche Bildung: Band 1. Sprachliche Bildung—Grundlagen und Handlungsfelder* (pp. 77–94). Waxmann.
- Reichert, M.-C., Rick, B., Marx, N., & Gill, C. (2020). Sprachliche Integration neu zugewandelter SchülerInnen in den Regelunterricht der Sekundarstufe I am Beispiel des Curriculumsentwurfs für die Bremer Vorkurse. *Informationen Deutsch als Fremdsprache*, 2020, 443–458. <https://doi.org/10.1515/infodaf-2020-0062>.
- Reiss, K., Weis, M., Klieme, E., & Köller, O. (Eds.). (2019). *PISA 2018: Grundbildung im internationalen Vergleich* (1. Auflage). Waxmann.
- Richter, T., & Christmann, U. (2009). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben & B. Hurrelmann (Eds.), *Lesesozialisation und Medien. Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (3rd ed., pp. 25–58). Juventa Verlag.
- Rick, B., & Gill, C. (2015). Seiteneinsteiger(innen) in Bremen. Einblicke in die Sprachförderpraxis. *Die Grundschulzeitschrift*, 30, 21–23.
- Saunders, B., & O'Brien, G. (2006). Oral language. In F. Genesee (Ed.), *Educating English language learners: A synthesis of research evidence* (pp. 14–63). Cambridge University Press. <https://doi.org/10.1017/CBO9780511499913.003>
- Schepens, J., van Hout, R., & Jaeger, T. F. (2020). Big data suggest strong constraints of linguistic similarity on adult language learning. *Cognition*, 194, 104056. <https://doi.org/10.1016/j.cognition.2019.104056>
- Schneider, W. (2008). Entwicklung, Diagnose und Förderung der Lesekompetenz im Kindes- und Jugendalter. In C. Fischer, F. J. Mönks, & U. Westphal (Eds.), *Begabungsforschung: Vol. 6. Individuelle Förderung: Begabungen entfalten - Persönlichkeit entwickeln* (pp. 131–168). LIT-Verl.
- Secretary-General of the OECD. (2019). *PISA 2018 Results. Combined executive summaries. Volume I, II & III*. https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf
- Soto-Corominas, A., Paradis, J., Rusk, B. V., Marinova-Todd, S., & Zhang, X. (2020). Oral language profiles of English second language learners in adolescence: Cognitive and input factors influence how they compare to their monolingual peers. *Studies in Second Language Acquisition*, 42, 697–720. <https://doi.org/10.1017/S0272263119000767>
- Umansky, I., & Reardon, S. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal*, 51, 879–912.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.
- Wolff, D. (1987). Some assumptions about second language text comprehension. *Studies in Second Language Acquisition (SSLA)*, 9, 307–326. <http://dx.doi.org/10.1017/S0272263100006707>.
- Woods, A. (2009). Learning to be literate: Issues of pedagogy for recently arrived refugee youth in Australia. *Critical Inquiry in Language Studies*, 6, 81–101. <https://doi.org/10.1080/15427580802679468>.

APPENDIX

TABLE A1. Model estimates predicting lower-level reading processes

Predictor	Beta	SE	<i>p</i>	Lower 95%-CI	Upper 95%-CI
MS	Reference				
SE-1	-2.160	0.138	≤0.001	-2.431	-1.890
SE-2	-1.468	0.145	≤0.001	-1.752	-1.185
SE-1 * t1	Reference				
SE-1 * t4	0.743	0.138	≤0.001	0.473	1.013
SE-1 * t5	0.615	0.179	0.001	0.265	0.965
SE-2 * t1	Reference				
SE-2 * t4	0.223	0.125	0.076	-0.023	0.469
SE-2 * t5	0.202	0.170	0.236	-0.132	0.535

TABLE A2. Model estimates predicting reading comprehension of expository texts

Predictor	Beta	SE	<i>p</i>	Lower 95%-CI	Upper 95%-CI
MS	Reference				
SE-1	-1.163	0.114	≤0.001	-1.386	-0.940
SE-2	-0.934	0.114	≤0.001	-1.157	-0.712
SE-1 * t1	Reference				
SE-1 * t3	-0.411	0.160	0.010	-0.725	-0.097
SE-1 * t5	0.081	0.160	0.612	-0.233	0.396
SE-2 * t1	Reference				
SE-2 * t3	-0.404	0.161	0.012	-0.721	-0.088
SE-2 * t5	0.130	0.161	0.421	-0.186	0.445

TABLE A3. Model estimates predicting reading comprehension of narrative texts

Predictor	Beta	SE	<i>p</i>	Lower 95%-CI	Upper 95%-CI
MS	Reference				
SE-1	-1.632	0.161	≤0.001	-1.947	-1.317
SE-2	-1.517	0.144	≤0.001	-1.800	-1.235
SE-1 * t2	Reference				
SE-1 * t4	0.540	0.180	0.003	0.187	0.893
SE-1 * t5	0.575	0.268	0.032	0.049	1.100
SE-2 * t2	Reference				
SE-2 * t4	0.628	0.143	≤0.001	0.347	0.908
SE-2 * t5	0.528	0.201	0.008	0.135	0.922

TABLE A4. Relevant influences of nominal variables on residuals (LLRP = lower-level reading processes)

Covariate	Variable	Residual of:	<i>n</i>	Mean (95%-KI)	<i>p</i>	<i>Eta</i> ²
SE-1: From crisis region	no	LLRP	60	0.21 (−0.05; 0.46)	0.008	0.06
	yes	LLRP	47	−0.27 (−0.50; −0.03)	0.008	0.06
SE-2: Gender	female	RC_Narrative	70	0.29 (0.06; 0.52)	0.002	0.08
	male	RC_Narrative	88	−0.23 (−0.40; −0.05)	0.002	0.08

TABLE A5. Relevant Spearman correlations of ordinal and metric variables on residuals (LLRP = lower-level reading processes)

Covariate	Residual of:	<i>n</i>	<i>p</i>	<i>Rho</i>
SE-1: Length of schooling in country of origin	LLRP	102	≤0.001	−0.355
SE-1: Age	LLRP	102	≤0.001	−0.423
	RC_Expository	104	0.001	−0.327
	RC_Narrative	101	0.001	−0.321
	DSD I	81	0.002	−0.335
SE-1: Length of preparatory classes	DSD I	85	0.011	0.276
SE-2: Length of schooling in country of origin	LLRP	144	≤0.001	−0.360
	RC_Narrative	141	0.001	−0.281
	DSD I	119	≤0.001	−0.340
SE-2: Age	LLRP	149	0.002	−0.250