*Article*

# Tracking Biliteracy Skills in Students Attending Gaelic Medium Education: Effects of Learning Experience on Overall Reading Skills

Euan Dickson [1], Laura Manderson [2], Mateo Obregon [3] and Maria Garraffa [1,*]

[1] School of Social Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK; e.dickson@hw.ac.uk
[2] School of Psychological Sciences and Health, University of Strathclyde, Glasgow G1 1QE, UK; laura.manderson.2020@uni.strath.ac
[3] Institute for Language, Cognition and Computation, School of Informatics, The University of Edinburgh, Edinburgh EH8 9AB, UK; mateo.obregon@ed.ac.uk
* Correspondence: m.garraffa@hw.ac.uk

**Abstract:** This study describes the validation of a reading assessment developed for speakers of Scottish Gaelic, an endangered language spoken in Scotland. The test is designed to investigate the areas of reading for understanding, reading errors and reading speed. This study will present the data on a group of Gaelic/English speakers on both the Gaelic and the English version of the test and of a group of English speakers on the English version of the reading test, aiming at comparing reading abilities in children attending a Gaelic medium education (GME) and children in English medium education (EME) living in the same urban area. The paper reports two studies. The first study presents data on 77 children bilingual in Gaelic/English recruited across four levels of primary school on reading in Gaelic. The second study looks at the performance on a version of the test adapted for English, comparing the performance of two groups of children (bilinguals Gaelic/English and monolinguals English) on several linguistic skills, including sentence comprehension and reading. About 40 monolingual English subjects in EME, living in the same urban area, were administered the English version. The reading abilities of the children attending EME and GME schools were comparable, supporting the idea of no disadvantage on reading from attending a school with the medium of a minority language. If differences were found, these were in favour of the bilingual Gaelic/English children, who attained better results in all linguistic tasks in English in the older groups.

**Keywords:** biliteracy; minority languages; Gaelic; Gaelic medium education; reading errors; reading comprehension; reading speed; macrostructures; sentence comprehension

## 1. Introduction

Reading is an intellectual activity that comprises the interplay of a linguistic dimension and a cognitive dimension with profound alterations in the brain circuitry during its acquisition in primary school (Dehaene 2009). Assessing reading among primary school children plays an important role in the field of education. It is estimated up to 1 in every 10 people in the UK has some form of dyslexia with long-term effects that can present challenges daily (British Dyslexia Association 2018). According to Bishop and Snowling (2004), the common broader characteristics of dyslexia include a mix of linguistic and cognitive difficulties, including memory problems, writing difficulties, organizational and time management difficulties. Other factors, such as socioeconomic causes, have recently emerged as having an impact on dyslexia opening interesting investigations in communities that are more isolated or are learning reading in a minority language with few opportunities for consistent daily reading in the minority language.

In children with learning difficulties, reading can be affected in different ways, with problems dealing with the linguistic level (e.g., understanding the meaning of a sentence)

and the cognitive level (e.g., remembering or focusing on the information), making crucial tracking reading abilities in school and promoting testing the of reading for understanding information and not just reading aloud.

Reading assessments are a crucial part of national curricula and are relevant also for the diagnosis and treatment of language and learning disorders. There are various types of assessments available for English, but the standardized norms do not consider bilingual speakers or speakers enrolled in programs where English is not the main literacy components, as in Gaelic medium education in Scotland (see Chondrogianni et al. 2021). Furthermore, many of the reading assessments available do not address discourse comprehension with a specific theoretical model, focusing mainly on reading errors in a reading aloud modality.

Learning to read in two languages simultaneously—for example within a school medium education program—is still considered by stakeholders in education to be a factor that can cause a delay in the acquiring literacy skills. This opinion is not supported by the current state of research on biliteracy. Recent studies are actually collecting evidence in favour of bilingual school programmes as an optimal environment for developing a fluent reading competence in more than one language. A number of studies and meta-analyses have now compared the various approaches to dual-language education, although more studies are needed with minority languages in particular (see Baker 2001; Greene 1998; Rolstad et al. 2005a, 2005b; Slavin and Cheung 2003, 2005). For example, in a study carried out with children attending a Spanish–English dual-language school at Grade 2 and Grade 3 in the USA and comparing monolingual English children living in the same area and matched for age, gender and SES, students in the dual language system performed significantly better in a series of tasks targeting phonological awareness, reading decoding, irregular word reading, passage comprehension and expressive language (Berens et al. 2013). This result is remarkable if we consider that the students placed in dual-language programmes received instruction in English for 50% or less of their time at school yet still managed to outperform children who were fully instructed in English. Berens et al. (2013) concluded that a dual-language approach to education not only does not hinder the development of language and reading skills in L1, but, on the contrary, seems to provide an advantage in the linguistic abilities at the core of learning how to read. Similar results have been reported in a study on bilingual Italian–English children (Costa et al. 2018). Bilingual participants who had been exposed to Italian since birth and to English within the first three years of life did not lag behind monolingual in the reading performance in Italian. Moreover, even when a difference between the two groups was present in their Italian oral performance, this disappeared by Grade 3. Overall, it seems that when dual-language exposure occurs at an early age, learning to read in an L1 and an L2 does not have a negative impact on reading performance in the L1. A number of other studies have reached similar results suggesting that, when compared to a later exposure to a second language, a bilingual exposure starting within the first three or four years of life yields the best linguistic results (e.g., Flege et al. 1999; Perani et al. 2003). Generally speaking, the earlier one becomes bilingual, the better, and this is true not just for oral language competence, but—as the two previously described studies show—for literacy proficiency as well. In fact, numerous authors have identified the first three to four years of life as a sensitive period for the acquisition of a second language; in other words, during their first few years of life, children experience a heightened sensitivity towards certain aspects of linguistic input and the learning they gain from this type of linguistic stimuli reaches its full potential (Kovelman et al. 2008). In spite of these encouraging results, the belief that learning to read in two languages simultaneously may confuse the child and slow down literacy acquisition is still widespread among stakeholders in education.

The main aim of this study is to introduce a new dimension (reading comprehension) to assess reading in children attending a biliteracy school program with the medium of a minority language (Scottish Gaelic), with few opportunities for practicing the language

outside the school and not many materials available to read in a natural context. This new dimension will be implemented in the Reading for All test.

The article presents two studies. The first study looks at the linguistic properties of the Reading for All test developed for Scottish Gaelic readers, reporting its linguistic properties and the normative data collected for children in Gaelic medium education (GME). The second study looks at the effects of biliteracy in both Gaelic and English, comparing the performance on a version of the test adapted for English in both children attending a Gaelic medium education school and children enrolled in an English medium school. Before discussing the data, we will introduce some features of the Gaelic medium education in Scotland and the main factors reported in the literature affecting reading for understanding.

### 1.1. Scottish Gaelic Medium Education

Gaelic medium education (GME) is a system of education in Scotland in which pupils acquire bilingualism and biliteracy in both Scottish Gaelic, henceforth Gaelic, and English by being educated through Gaelic immersion. It is a model of bilingual education adapted from other bilingual immersion education programmes that have grown in numerous countries and regions, primarily across Canada and Europe (Baker 2001). In most examples of bilingual immersion education programmes, the children, many of whom are L2 speakers or sequential learners of the target language, are taught almost exclusively in the target language in order to achieve full fluency and literacy in this target language (Cenoz and Gorter 2017). Thus, although GME does not exclude use of the dominant language— English—entirely from the curriculum, all GME children receive total immersion in Gaelic in at least the first three years of primary school (P1–P3) and, although English is gradually introduced in later primary, Gaelic remains the dominant language of instruction across the curriculum (Learning and Teaching Scotland 2010; O'Hanlon et al. 2013).

The results of the most recent UK census in 2011 confirmed that the number of speakers of Gaelic has continued to fall from 92,400 individuals reporting some Gaelic language skills—or some 59,000 speakers—in 2001 to 87,100 individuals—or some 58,000 speakers-in 2011 (National Records of Scotland 2015, p. 9). Gaelic speakers tend to be older on average in comparison with the general Scottish population as the census results demonstrated that 24% of Gaelic speakers were older than 65 years old in 2011. In comparison, only 17% of the population who could not speak Gaelic were older than 65 years old (ibid: p. 13). Nonetheless, the 2011 census has shown that there has been a slight increase in the proportion of the number of children under 18 years old living in Scotland who can speak Gaelic, particularly for children aged 5–11, for which the proportion of speakers grew from 0.91 per cent to 1.13 per cent (National Records of Scotland 2015, p. 15). This increase in younger speakers has most likely been the result of the increase in GME provision in recent decades as pupil enrolment figures have increased exponentially from 24 pupils in 1985 (O'Hanlon et al. 2013, p. 708) to 4631 pupils enrolled in the school year 2019–2020. It is important to note that most of the children attending GME are from English speaking homes (O'Hanlon et al. 2013), indicating that this growth in GME is in a large part driven by demand from new speakers or from parents who are interested in Gaelic for their children. Interest in bilingual education has grown in part due to the growing evidence of the inherent cognitive benefits of bilingualism such as greater executive control of language functions (Bialystok 2011; Baker 2001) and phonological awareness (Bialystok et al. 2005). Therefore, it was not unexpected that a study interviewing teachers and parents of children in GME found that many parents had chosen to send their children to GME because of the cognitive benefits attributed to bilingualism (O'Hanlon et al. 2010, pp. 51–52). A positive effect of being bilingual via GME in language and cognition has been recorded in a study on young adults attending GME since primary school, where both English grammar and attentional abilities were higher compared to English monolingual young adults living in the same area (Garraffa et al. 2020).

Despite this relative success, the current provision of GME has been criticised due to the lack of sufficient assessment material published in Gaelic and the subsequent failure

to effectively identify pupils who are struggling with Gaelic literacy (Lyon and Quarrie 2013). In a paper discussing early years intervention in GME, Lyon illustrates this lack of adequate assessment material by drawing a comparison between the use of phonological screenings tools in English medium schools and GME schools, with all English medium schools in 2003 using "some form of baseline assessment" in the former "whereas 44% of Gaelic-medium schools did not use any screening tool at all; 20% used a screening tool in English; 36% translated existing tests into Gaelic or made up their own assessment" (2011, p. 14). A consequence of translating assessments designed for use in English medium schools into Gaelic can often be that the translations are written using vocabulary and language unfamiliar to the children (ibid.). There is a similar lack of adequate test material created for use in GME in other forms of assessment, such as for older pupils, and for aspects of attainment such as literacy. Thus, previous research on attainment in Gaelic has been reliant on imprecise indicators for measuring pupils' progress, such as teachers' judgements, as in a recent report analysing attainment in GME (O'Hanlon et al. 2010, 2013). This project found that GME pupils in both P5 and P7 were believed to perform comparatively or better than English monolinguals in assessments for Maths, Science, and crucially outperformed monolinguals in the English reading and writing assessments (O'Hanlon et al. 2013), indicating that there was sufficient catch up in English reading. Yet, the same study indicated that the GME pupils had higher attainments in English reading and writing than in Gaelic reading and writing, although the attainment gap was narrower in P7 then P5 (ibid, pp. 714–15). As stated above, aside from the comparison of results in science, these findings were based solely on teachers' estimations on pupil attainment and it was found that the teachers were more optimistic of their pupils' levels of attainment in science than the results obtained in the objective science tests reflected (O'Hanlon et al. 2010, p. 12). Similar inaccuracies could exist in the teachers' estimations of the GME pupils' attainment in English and Gaelic literacy. Thus, in order to develop a more exact understanding of attainment in later primary school, a new reading assessment designed for use in GME is required.

In the next section we will introduce the main factors reported in the literature affecting reading for understanding and adopted to develop the reading test used in this study, the Reading for All test. Discourse comprehension was the main target of the study aiming at gathering data on the reading competence at text level and not a single word level in both Gaelic and English.

## 1.2. Factors Affecting Reading for Understanding

The purpose of discourse comprehension is for a reader to extract global meaning from a passage of text (Graesser et al. 1997). That is, to specifically understand written or spoken text beyond the level of words and sentences. Discourse comprehension, or reading for meaning, is an everyday activity practiced, for example, when a student reads a story for enjoyment or studies a technical text for an examination. In fact, reading for meaning underlies our participation in many aspects of life including our social, work and leisure activities (Webster et al. 2018).

In their chapter on discourse comprehension, Graesser et al. (1997) report that both offline and online measures can be applied to measure reading ability. Offline measures include subsequently answering questions on a passage, whereas online measures typically consider reading speed or number of reading errors. However, reading ability is also impacted by several variables related to the individual reader and to the text. For the purposes of this study, we focus on text-related variables.

A model for discourse comprehension proposed by Kintsch (1988) describes how the reader extracts meaning from discourse through processing the structure of the text. At a conceptual level, discourse is comprised of the macrostructure and microstructure representing the structure of the text and meaning at global and local levels. Microstructure refers to the surface features of sentences including lexical items, whereas the macrostructure is formed from propositions that reoccur across the text base comprising the main

ideas (Van Dijk and Kintsch 1983). The construction–integration model (Kintsch 1988) posits that the reader must first process the linguistic form of words and sentences during the construction phase. This is then followed by integrating the information into a coherent representation through a process of returning to information stored in episodic memory and modifying any ambiguous or incorrect inferences.

The memory and recall of a passage are influenced by the relative importance of the information collected from the main ideas and details of the text. According to Haberlandt (1994) the most durable representations are those that are formed on a global level; representations based on surface features are more short-lived. This conveys that main ideas are better remembered because they are repeated and elaborated on; details, on the other hand, although related to the main ideas, are not repeated throughout and thus are more easily forgotten.

Meaning can be also be extracted directly from the text or from making inferences. Inferences are the derivation of additional knowledge from facts already known; this might involve going beyond the text to maintain coherence or to elaborate on what was actually presented. The ability to make inferences is a crucial aspect of discourse comprehension. Poor readers have been shown to have trouble constructing coherent representations and filling in the gaps due to specific challenges with integrating information in the text and incorporating general knowledge (Oakhill 1984; Cain and Oakhill 1999). There are three types of inferences. Logical inferences are made from the semantic meaning of the words in the text. Bridging inferences require relating new material found in the text to previously stored knowledge. Elaborative inferences are the most advanced and require world knowledge to extend what is in the text. According to Kintsch (1988) inferences are required at the global and local level, meaning that main ideas and details can be both stated in the text or implied.

Likewise impacting reading ability is the difficulty of the text. In a review study, Amendum et al. (2017) found that comprehension is affected by more complex texts with low familiarity of words. Students who were able to read above their grade level with 90% comprehension accuracy showed significantly poorer comprehension than students who read texts near their grade level, indicating that discourse comprehension is a complex activity independent of text processing. Readability formulas are commonly applied to ensure a text is at the appropriate level for a person of a given age. The Dale–Chall readability formula provides a rapid grading of text difficulty by considering mean sentence length (syntactic complexity) and number of unfamiliar words. Passages that contain longer sentences, longer words and rare words would be considered less readable than passages containing shorter sentences and words and more frequent words (Benjamin 2011).

## 2. Study 1: The Development of a Reading Assessment in a Minority Language

The main aims of study 1 are related to the development process of the reading assessment and its results. The study reports data on the reading abilities of a group of young Gaelic speakers and adult Gaelic learners. Study 1 explores:

i.     The trajectory of the reading abilities across four school levels (P4–P7) in a cross-sectional study of the children attending a minority language program.
ii.    The factors affecting reading comprehension, considering variables such as the implicit nature of the information or the capacity to remember details.
iii.   Other reading factors at microstructure level, such as reading errors and reading speed.

### 2.1. Methods

#### 2.1.1. Participants

The sample for this study consisted of 77 primary school children attending two Gaelic medium education schools and divided into four age groups (23 from P4, 19 from P5, 22 from P6 and 13 from P7) aged 7 to 11. This is a cross-sectional study, developed as a pilot study to assess the validity of the Reading for All test aiming at investigating reading

abilities in a minority language. The study does not represent an exhaustive picture of the reading competence of children attending GME but is a psycholinguistics study conceived for the development of a reading assessment in Gaelic. Participants were randomly selected from two data location sites providers for GME in Scotland. They were controlled for socioeconomic status (SIMD: Scottish index of multiple deprivation) and selected based on the following criteria: attending a GME program since the start of primary school, reported by their teachers to have typical language development, aged between 7 and 11, consent obtained from parents for their child to participate in this study. Only children who completed the test were included in the analysis, for a final total of 57 children (P4: 12, P5: 14, P6: 19, P7: 12).

Ethical approval was obtained from the Education Research Steering Group of the two municipalities involved in the project and from the School of Social Sciences Heriot-Watt University Ethics Committee.

2.1.2. Materials

Together with the reading test developed for the study (*Reading for All*) one other test assessed a baseline of the language development in a reading paragraph task (York Assessment of Reading Competence, YARC; Snowling et al. 2009).

York Assessment of Reading for Comprehension (YARC; Snowling et al. 2009). The YARC is a commonly used assessment to measure the progression of children's discourse comprehension ability. The test contains 6 passages that each increase in length and complexity (level 1 to level 6). Subjects read the first passage aloud and verbally answer questions related to the passage. As the child reads, reading errors are noted and questions responded to are marked on a scoring sheet. The YARC standard instructions specify that the child progress to the next passage only if the number of errors do not exceed 15–20, otherwise the test should be discontinued. However, an adapted version of the YARC was administered in this study to be used as a validity measurement for the *Reading for All* tests. Three passages were selected (level 1, 3 and 6), which every participant completed in full regardless of the number of errors made. For each of these three passages, reading accuracy was measured by number of errors, reading rate was estimated as the time taken to read each passage and a comprehension score determined by the number of correct question responses.

Reading for All (Gaelic Version) The Reading for All test was designed for Gaelic by a native speaker of the language. It was conceived to be culturally neutral with no specific references to the Celtic culture but adapted to a contemporary society. The stories developed for the test were piloted with a group of adult fluent speakers from different backgrounds and with different reading habits to ensure no a priori knowledge was required. The test, conceived in Gaelic, was then adapted to other languages, including the English version of this study. The test was implemented using PsychoPy (Peirce et al. 2019) computer software specialized for psychological research. The design was like the YARC described above but instead of a paper and pencil task it was presented digitally, meaning that participants read passages and questions on a screen and selected specified keys to respond. This allowed measurement of reading times and eliminated the need for any verbal input to respond to the questions. Questions were in the style of sentence completion with three multiple-choice answers.

The story passages were devised based on the method used in Webster et al. (2018), which ensured sentences were both semantically and syntactically constrained. This eliminated the possibility of comprehension being affected by reversible sentences or noncanonical word order. The content of the stories depicted neutral events intended to appeal to a variety of readers independent of specific or general knowledge.

Two versions of the test were created (A and B). In each version the stories follow either "Anna and Calum" or "Iain and Dorota" as they embark on several naturalistic, everyday activities such as riding the bus or purchasing a new computer. Both versions of the task

contained one practice paragraph with two questions and seven test paragraphs followed by four questions. Participants were randomly allocated to one of the two versions.

Based on discourse models, the questions measured comprehension of four dimensions: stated and implied information in both the main ideas and details of the text. Sentence completion questions assessed factual and inferential understanding of each paragraph. Passage dependency effects (PDE), the reliance on the passage to answer the questions, were considered when devising the questions. Answers were selected by pressing a specified key from a choice of three answers (target and two distractor items).

### 2.1.3. Procedures

The Reading for All assessment was conducted individually in a quiet room in the school. For the children the group tasks were presented in the order of the YARC followed by Reading for All. Each test had two to three practice items prior to the administration of the actual subtest items. During the practice items, repetition of the target question or sentence was allowed. However, no repetition of items was given during the actual test. After the practice items, the tests were administered in one session. Each session lasted on average 30–40 min.

### 2.2. Results and Interim Discussion of Study 1

The raw data from the Reading for All Gaelic is reported in Table 1 for children divided by class group. The data reported are divided across the three reading components: reading comprehension score, reading errors and the reading time (the average reading times of all seven paragraphs).

**Table 1.** Raw data on the Gaelic version of the Reading for All test across the three dimensions. Note that while the score includes all participants, the errors and speed numbers only include participants who completed the seven paragraphs of the test.

| Group (N/Complete) | Reading Score Mean (SD) | Reading Errors Mean (SD) | Reading Time Mean (SD) |
|---|---|---|---|
| P4 (23/12) | 11.05 (5.13) | 130.45 (68.59) | 1050.43 (368.78) |
| P5 (19/14) | 17.00 (4.78) | 97.36 (68.30) | 678.57 (258.94) |
| P6 (22/19) | 17.41 (4.78) | 87.79 (55.25) | 671.09 (155.32) |
| P7 (13/12) | 20.54 (3.69) | 56.25 (49.12) | 490.82 (82.20) |

Generally, it shows participants improving with age. An upward trend can be seen for the score, while a downward trend can be seen for reading errors and reading times. It is also worth noticing that the variability among participants becomes smaller when they improve; this is shown by the SDs for each group of participants declining in the later age groups.

Figure 1 offers a more detailed exploration of the data reported in Table 1.

Overall, the results reported in the raw data on the Gaelic reading test provide interesting information on the ability to read in Gaelic from a group of children attending a Gaelic medium education and studying reading in English as a subject. First it was reported that the reading test score increases in later classes and with paragraph length. This is an interesting finding as it supports the idea that the length of a text is not the main cause of difficulty in reading comprehension and a longer passage offers more opportunity to understand and retain the information. A second finding was related to the decreases in the reading time in later classes, as naturally supported by the higher reading fluency expected in older children. An important result is the equivalence of the two versions of the Gaelic

test (version A and version B), with no difference in any of the reading components across the two groups. This finding makes the reading test a good tool for testing and retesting reading abilities at different time intervals.
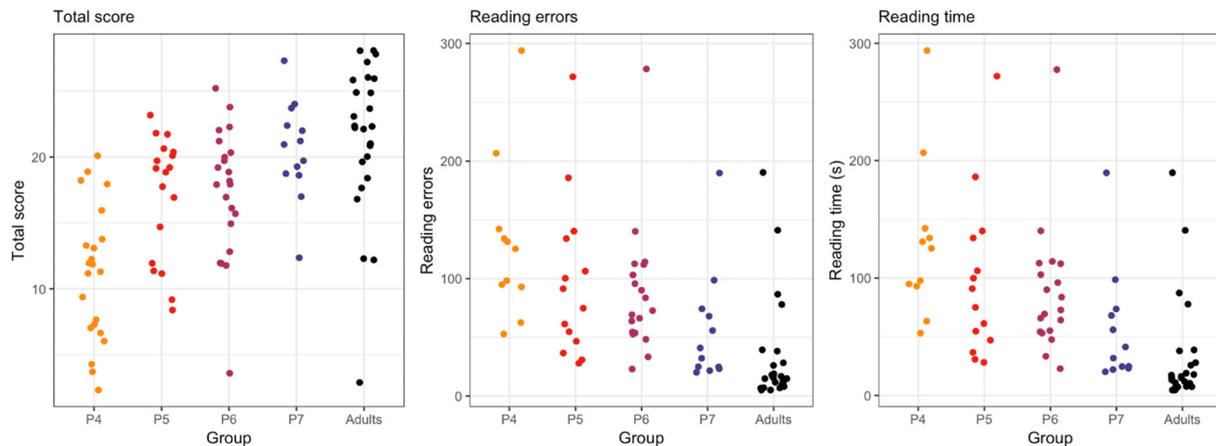


**Figure 1.** Scatterplot for reading test across reading dimensions: reading score, reading errors and reading times by participant groups. The reading errors and times only include participants who completed the test to give a more realistic view of the data.

### 2.2.1. Statistical Analyses of the Reading Components

The analyses modelled the outcome of the comprehension questions using a generalised mixed-effects logistic regression (Bates et al. 2015) for the total reading score, the number of reading errors and the reading times with a mixed-effect linear regression (Bates et al. 2015), including SIMD, school class, the length of paragraphs (in hundreds of words) and the version of the test as fixed effects, as well as a random slope for paragraph length by participant nested within school.

The modelling of the total score shows significant positive effects of school class and paragraph length wherein children answer more questions correctly with every school year ($\beta = 0.26$, $p = 0.001$) and with longer paragraphs ($\beta = 0.21$, $p = 0.03$).

School class and word count also had significant effects for reading times: children take shorter times to read a paragraph with each school year ($\beta = -12.32$, $p < 0.001$) and, as expected, take a longer time to read longer paragraphs ($\beta = 69.67$, $p < 0.001$). This means that, on average, children's reading times are about 12 s shorter for each year and 70 s longer for every 100 words.

As for the number of reading errors, only the word count of a paragraph influences children: longer paragraphs yield more errors ($\beta = 11.29$, $p < 0.001$). This means that for every 100 words, 11 more reading errors are committed on average.

The fixed effect for the version of the Reading for All never reached significance (respectively $p = 0.42$, $p = 0.28$ and $p = 0.34$ for the three analyses), meaning that the two versions of the test are equivalent in the scores, reading errors and reading times they yield.

### 2.2.2. Statistical Analyses for Concurrent Validity

For concurrent validity, it was suggested that the group of child participants' scores on the Reading for All be compared with scores from a comparable test widely adopted in schools to measure reading in English, such as the York Assessment of Reading for comprehension, YARC (Snowling et al. 2009).

Chi-squared tests were run comparing the tests to see whether the results in one of them correlate with the others. P values were computed by Monte Carlo simulation with 2000 replicates. The overall score of the YARC correlated with the reading comprehension score of Reading for All ($\chi^2 = 382.54$, $p = 0.01$).

### 2.2.3. Statistical Analyses of the Reading Macrostructure Components

In terms of the properties of the test, the effect of the reading comprehension variables was also considered: stated vs. implied information, as well as main idea vs. detail. The results of the modelling show that implied information is harder to comprehend when the questions are about implied details.

A generalised mixed effect model was made predicting correct answers with main/detail, stated/implied and their interaction as fixed effects, and a random intercept for participant ID nested within the group.

The results for the children show no difference between stated information and stated details ($\beta = -0.02$, $p = 0.43$), but an effect is found where implied information is understood less than stated information ($\beta = -0.5$, $p < 0.001$). An interaction of the two factors (details/idea and stated/implied) is significant, meaning that implied details are the most difficult components ($\beta = -0.46$, $p < 0.001$): the difference in difficulty between stated and implied information is bigger for details than for main information.

### 2.2.4. Statistical Comparison between English and Gaelic Reading Tests

Comparing the total score each participant had in the English and Gaelic reading tests is also informative to consider if an immersion program in a minority language does affect reading abilities in the other language. The comparison was made through a linear model, predicting the score in the Gaelic test from that of the English test. The other predictors were chosen through model comparison: adding Class significantly improved the model ($p < 0.001$), while adding the SIMD or the interaction of the English score with Class did not (respectively $p = 0.26$ and $p = 0.07$). The results show that for each point increase in the English test score, there is a 0.6 increase in the Gaelic test score: this means that, in general, children score better in the English test ($\beta = 0.63$, $p < 0.001$). Moreover, the Gaelic score increases by approximately 1.8 points for every class ($\beta = 1.83$, $p < 0.001$). This result is shown in Figure 2.
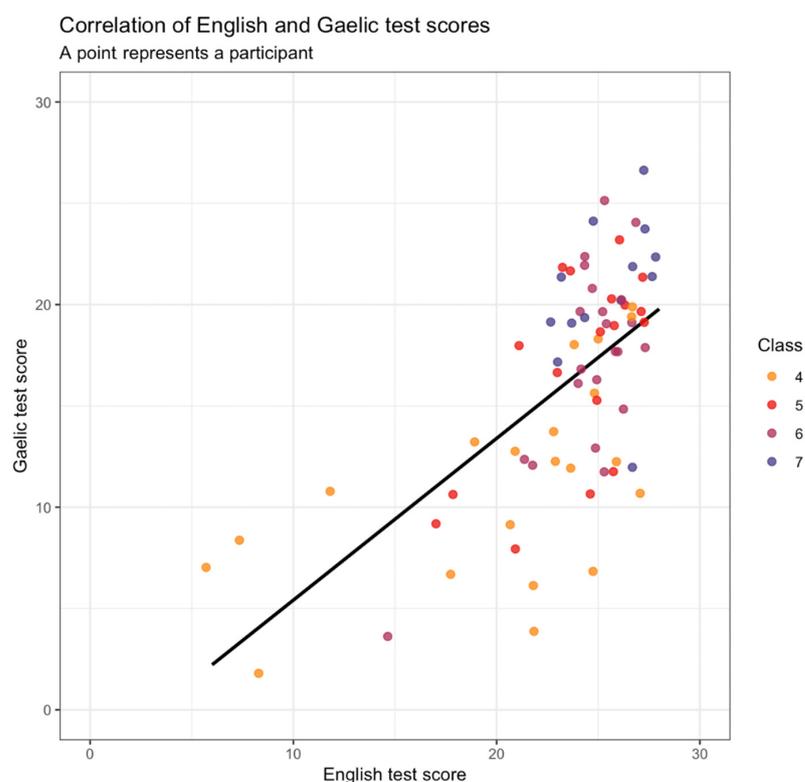


**Figure 2.** Correlation of the English and Gaelic test scores. The regression line is plotted in black, while the individual scores are shown as points colour-coded by school class.

### 3. Study 2: A Comparison on Reading in English between Gaelic/English Speakers and English Speakers

In study 2 we aimed at comparing the performance of the Gaelic/English speakers (presented above) with a sample of English-speaking children living in the same urban area on the English version of the Reading for All test. We also tested both groups on reading in English using the YARC and an additional measure of comprehension in English, TROG-2 (described below in materials).

The main aims for this study were:

- To compare the performance of Gaelic/English speaking children and adults with English speaking monolinguals on the English versions of the Reading for All test, looking at all reading components and macrostructures.
- To investigate if other linguistic measures such as sentence comprehension have a similar development in both Gaelic/English and English children.

### 3.1. Methods

#### 3.1.1. Participants

The same group of Gaelic/English bilinguals that participated in study 1 (described above) were also tested on the same measures in English for study 2.

English speaking monolinguals were recruited from an English Medium Education School in Scotland. The control group consisted of 20 participants from primary 4 to primary 7 (6 from p4, 9 from p5, 2 from p6 and 3 from p7) aged between 8 and 11. Participants were randomly selected from one site on the basis that they spoke no additional languages and had no diagnosed language impairment.

#### 3.1.2. Materials

English versions of the same materials used in Study 1 were adopted for use in Study 2. This includes Reading for All (English Version) and YARC (English version). An additional sentence comprehension task was included in study 2, Test for Reception of Grammar-2 (TROG-2; Bishop 2003).

TROG-2 is a standardized test of receptive grammar measuring understanding of simple to complex grammatical structures. Participants are presented with sentences read aloud by the investigator. The aim is to select the picture that corresponds correctly to the sentence from a set of images. Each trial consists of one sentence coupled with four pictures. The pictures comprise the target and three distractor images that manipulate the verb, subject or object of the sentence. For example, for the item "the girl pushes the box", the distractor images show a girl jumping on a box, an elephant pushing a box and a girl pushing a tree, as well as a girl pushing a box. The test is divided into 20 blocks, measuring increasingly complex grammatical constructs four times each (e.g., Block A—simple active sentence with two elements; Block T—centre embedded sentence). Subjects must give a correct response to all four items within a block in order to pass the block. Two scores are produced: the total score results in a number out of 20 consisting of the number of blocks passed. The standard score comprises the total number of correct individual items resulting in a number out of 80.

#### 3.1.3. Procedures

The Reading for All assessment in English was conducted individually in a quiet room at the school. The tasks were administered in a fixed order: TROG-2, YARC and Reading for All. Each test had two practice items prior to the administration of the actual subtest items. During the practice items, the target question or sentence could be repeated if necessary. During the actual test, no repetition of items was allowed. All tasks were administered in one session lasting approximately 45 min.

### 3.2. Results for Study 2 and Interim Discussion

Table 2 summarises the raw data on the English version of the Reading for All test and the other linguistics measure selected for English. Groups are divided by language (Gaelic are the bilingual children and English the monolinguals) and by class (P4–P7).

**Table 2.** Raw data on the English Reading for All test, York Assessment of Reading for Comprehension (YARC) and Test for Reception of Grammar-2 (TROG-2) across age and linguistic groups. All cells report means and standard deviation.

| Group (N) | English Reading for All Score | English Reading for All Errors | English Reading for All Speed (msec) | YARC Score | YARC Errors | YARC Speed | TROG-2 |
|---|---|---|---|---|---|---|---|
| | | | **P4 level** | | | | |
| Gaelic (23) | 20.73 (6.55) | 45.82 (30.06) | 660.89 (323.44) | 3.27 (2.38) | 22.02 (29.19) | 188.33 (142.40) | 13.91 (2.45) |
| English (9) | 24.11 (2.76) | 29.89 (15.38) | 577.77 (117.01) | 5.00 (1.71) | 6.48 (7.28) | 138.26 (67.77) | 14.89 (2.57) |
| | | | **P5 level** | | | | |
| Gaelic (19) | 23.58 (3.59) | 34.76 (30.20) | 412.55 (112.45) | 4.86 (2.19) | 11.00 (13.50) | 100.75 (57.39) | 15.95 (1.54) |
| English (11) | 23.45 (2.21) | 21.45 (6.56) | 615.27 (304.08) | 5.25 (1.92) | 5.47 (4.93) | 130.88 (72.73) | 14.27 (2.50) |
| | | | **P6 level** | | | | |
| Gaelic (22) | 24.45 (2.58) | 20.10 (13.32) | 391.50 (86.40) | 5.88 (1.57) | 6.15 (7.37) | 93.65 (53.40) | 16.64 (1.15) |
| English (8) | 24.00 (2.98) | 32.38 (24.85) | 359.17 (76.60) | 6.13 (2.03) | 6.96 (7.79) | 88.77 (43.18) | 14.50 (2.56) |
| | | | **P7 level** | | | | |
| Gaelic (13) | 25.54 (1.94) | 15.23 (15.45) | 350.48 (102.19) | 6.18 (1.35) | 5.82 (6.95) | 82.82 (49.65) | 16.92 (1.98) |
| English (9) | 24.89 (2.09) | 22.22 (10.66) | 344.05 (65.83) | 6.30 (1.68) | 2.89 (4.72) | 74.56 (28.49) | 15.67 (1.16) |

We will now present the statistical analyses for each test, discussing first the data on the bilingual vs. monolingual children group and in a separate section the data from the adults.

### 3.2.1. Results from TROG-2 Test: Bilingual vs. Monolingual Children

The Test of Receptive grammar (TROG-2) analysis modelled the outcome of a single TROG block using a generalised mixed-effects logistic regression (Bates et al. 2015). The maximal random effects structure was used when supported by the data (Barr et al. 2013) and a principal components analysis of the random effect did not indicate any overspecification. SIMD, school class, group and the interaction of group with SIMD and with school class were used as fixed effects for the dataset of children. The model also included a random slope for school class and its interaction by participant ID nested within their school, and a random slope for group and its interaction by TROG block.

The results show an effect of class whereby the performance improves in later school years ($\beta = 0.25$, $p = 0.003$), and an interaction of class with group showing that children in Gaelic-language schools improve even more as they progress in school years ($\beta = 0.35$, $p = 0.04$). This interaction is visualised in Figure 3.
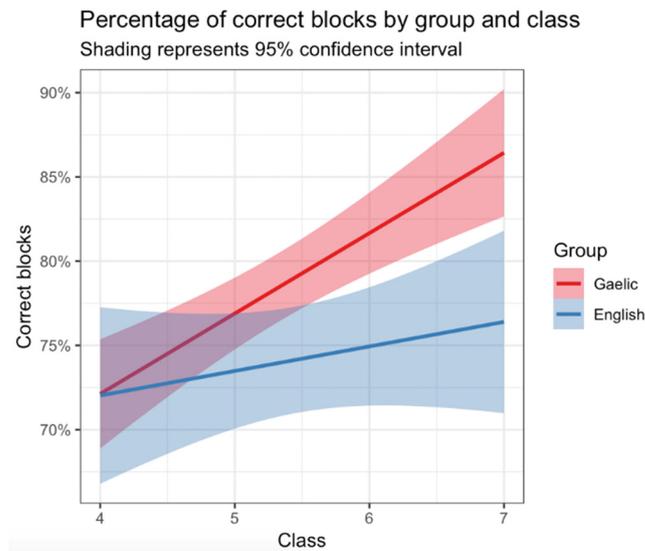
**Figure 3.** TROG-2 performance by group (Gaelic and English), by class.

### 3.2.2. Results on YARC Test: Bilingual vs. Monolingual Children

The YARC test data is composed of three measures: total score (comprehension), number of reading errors, and reading time. The three paragraphs composing the test are ordered by incremental difficulty and length. Analyses for the three measures are reported separately.

The total score analysis used mixed-effects linear modelling using group, SIMD, school class, the interaction of group with SIMD and school class, and the YARC paragraph as fixed effects, with a random intercept for participant ID nested within school. The results of the modelling show a significant effect of group whereby the total score is lower for children in Gaelic-language school than in non-Gaelic-language schools ($\beta = -1.69$, $p = 0.01$). As in the TROG test, there is a significant effect of school class such that students in higher classes perform better ($\beta = 0.78$, $p < 0.001$). Moreover, class and group show an interaction whereby the positive effect of being in a later class is even more pronounced in Gaelic-language schools ($\beta = 0.63$, $p = 0.02$). This is visualised in Figure 4. Finally, an effect of paragraph can be seen such that more difficult paragraphs produce lower scores ($\beta = -0.77$, $p < 0.001$).
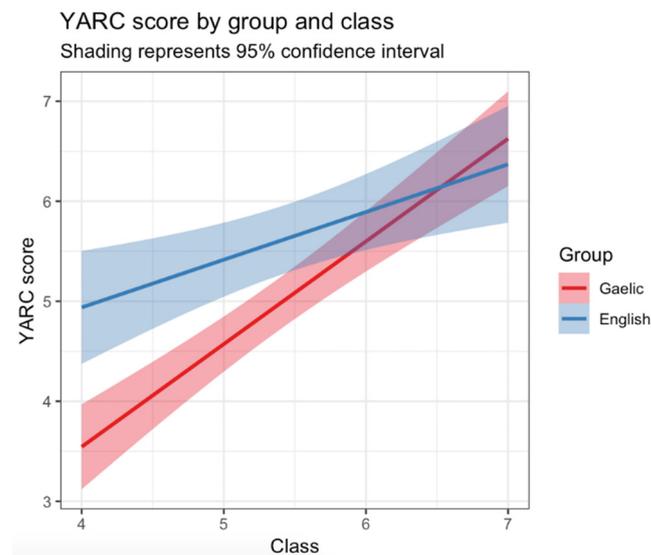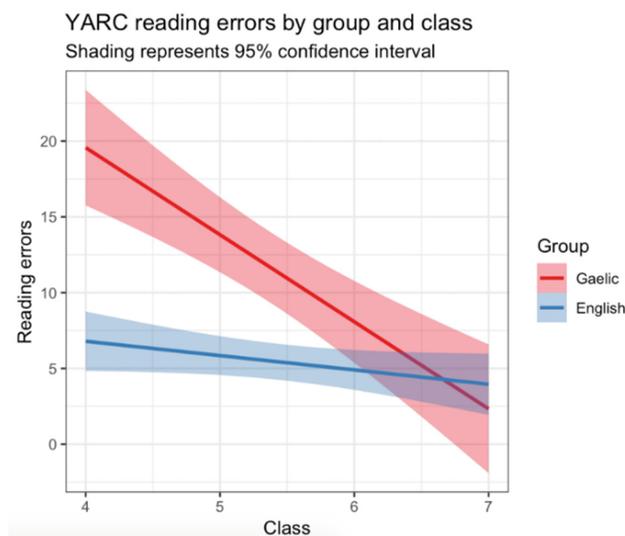


**Figure 4.** YARC performance by group (Gaelic and English), by class.

The reading errors analysis used the same methods and predictors as the score analysis. The results of reading errors show an effect of group, with Gaelic-language school students making more errors while reading ($\beta = 18.66$, $p = 0.001$) and an effect of class where, as expected, in the later school classes the students make fewer errors ($\beta = -4.08$, $p < 0.001$). Moreover, these two predictors interact: students from Gaelic-language schools improve more, meaning that they make proportionally even fewer errors in later classes ($\beta = -6.00$, $p = 0.01$). This is visualised in Figure 5. Finally, an effect of paragraph difficulty can be seen in which the more difficult paragraphs of the test produce more reading mistakes ($\beta = 6.34$, $p < 0.001$).



**Figure 5.** YARC reading errors by group (Gaelic and English), by class.

The reading times analysis also used the same methods and predictors. The modelling results show an effect of school class such that with each class the reading time goes down by approximately 33 s ($\beta = -33.16$, $p < 0.001$). An effect of paragraph can be seen where with each increment in complexity of paragraph the reading times increase by approximately 58 s ($\beta = 57.53$, $p < 0.001$). No effect of group was reported.

### 3.2.3. Results of the Reading for All Test (English Version): Bilingual vs. Monolingual Children

The Reading for All test data is the most complex. The first three analyses correspond to those of the YARC test to provide an easier comparison, investigating total score (comprehension), number of reading errors, and reading time. The length and complexity of the paragraphs composing the test are used as predictors. The length of the paragraphs was measured in hundreds of words. The difficulty of the paragraphs was expressed in different ways: with the Dale–Chall Score and with the simple/complex binary of the experimental design. The fit of the reading time and reading errors models was much better (both $p < 0.001$) using the Dale–Chall Score as a measure of difficulty. The complexity binary was used in the total score analysis, as (unlike the Dale–Chall score) it considers things such as the type of structures, which may influence comprehension.

A further predictor was added to all models to check whether the version of the test (A vs. B) made a difference. Children who did not complete the test were excluded from the analyses.

The total score analysis modelled the outcome of the single comprehension questions using a generalised mixed-effects logistic regression (Bates et al. 2015), with group, SIMD, school class, the interaction of group with SIMD and with school class, paragraph length and test version as fixed effects. It included a random intercept by participant ID nested within their school.

Both SIMD and school class are significant effects (respectively $\beta = 0.06$, $p < 0.001$ and $\beta = 0.19$, $p = 0.008$), where participants in higher SIMD score and later classes produce more correct answers. These two effects are visualised in Figure 6. Greater paragraph lengths also increase the number of correct answers ($\beta = 0.48$, $p < 0.001$). On the other hand, complex paragraphs elicit fewer correct responses ($\beta = -0.28$, $p < 0.001$). The two versions of the test do not differ significantly.
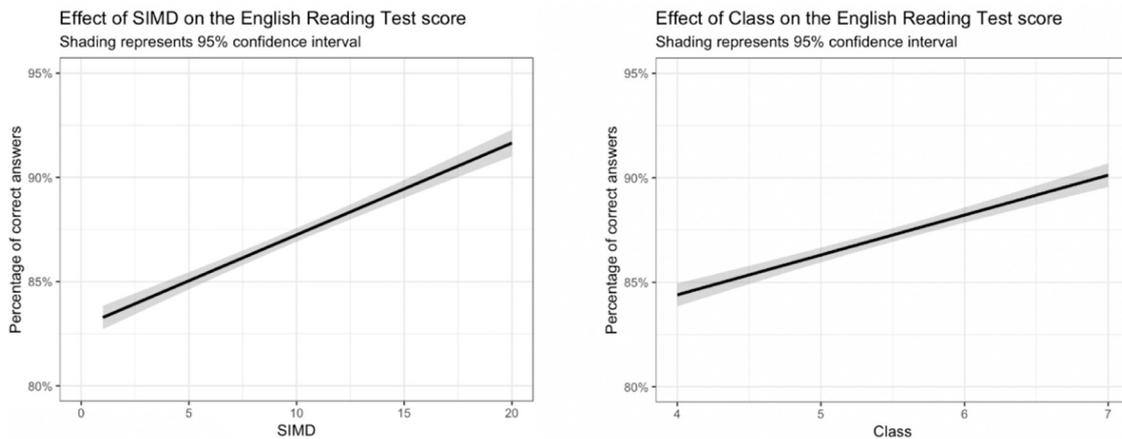


**Figure 6.** Effects of Scottish Index of Multiple Deprivation (SIMD) and class on total score of the English Reading for All.

The reading errors analysis used a mixed-effect linear regression (Bates et al. 2015), with SIMD, school class, group, the interaction of group with SIMD and with school class, the version of the test, the paragraph length and the Dale–Chall score as fixed effects. It included a random intercept by participant ID nested within their school.

The significant effects of group and class can be seen where children from Gaelic-language schools make more reading mistakes ($\beta = 3.66$, $p = 0.01$) and those from later classes make fewer mistakes ($\beta = -0.82$, $p = 0.004$). An interaction of group and class can be seen whereby on progressing to the later classes, children in Gaelic schools decrease their number of mistakes more steeply ($\beta = -1.30$, $p = 0.02$); this is visualised in Figure 7. The length and difficulty of a paragraph also have an influence, with participants making about 3.8 additional mistakes with each 100-word increase in length ($\beta = 3.80$, $p < 0.001$) and 1 additional mistake for each point of increase of the Dale–Chall score ($\beta = 1.01$, $p < 0.001$).
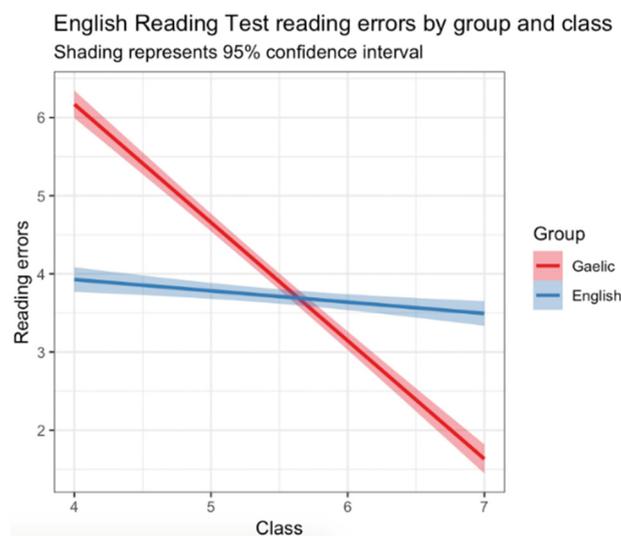


**Figure 7.** English Reading for All reading errors by group (Gaelic and English), by class.

The two versions of the test differed slightly regarding reading errors ($\beta$ = 1.29, $p$ = 0.03), with test B eliciting more mistakes.

The reading time analysis used the same method and predictors of the analysis of the number of errors. The group reaches a significant effect with children from English schools being faster than children from Gaelic schools ($\beta$ = −25.13, $p$ = 0.04). An effect of school class can again be seen in which with each year of progression in school, the reading time decreases by approximately 12s ($\beta$ = −12.39, $p$ < 0.001). Finally, an interaction between group and SIMD can be seen, whereby children from Gaelic schools and a lower deprivation area are slower ($\beta$ = 2.20, $p$ = 0.03). This effect may be due to the scarcity of the data. As expected, longer paragraphs take a longer time to read, with an increase of 53 s for every 100 words ($\beta$ = 53.31, $p$ < 0.001) and more complex paragraphs take also a longer time, with a 4s increase for every point of the Dale–Chall score ($\beta$ = 4.12, $p$ < 0.001).

To see whether some types of content are understood better than others, further modelling considered whether the comprehension questions were targeting main ideas or details from the text and stated or implied content. This was also compared across groups. Main vs. Detail and Stated vs. Implied were coded as binaries whereby main and stated had negative values and detail and implied had positive values. A generalised mixed effect model was made predicting correct answers with group, main/detail and stated/implied and their two way and three-way interactions as fixed effects, and a random intercept for participant ID nested within school. The model was compared to simpler model without interactions, but the best model proved to be the one with all interactions.

The results show no effect of group ($p$ = 0.14). The difference between main content and details does not seem to make a difference ($p$ = 0.54), but that between stated and implied content does, with implied content lowering the rate of correct answers ($\beta$ = −0.62, $p$ < 0.001). There is an interaction of group with the stated vs. implied condition, where children from Gaelic schools struggle even more with implied content ($\beta$ = −0.27, $p$ < 0.001). This is visualised in Figure 8. Moreover, if the content is both implied and a detail, the rate of correct answers goes down even more ($\beta$ = −0.75, $p$ < 0.001). Finally, a three-way interaction can be seen whereby for children in Gaelic schools, this last tendency is less strong, and questions targeting content that is an implied detail produce more correct answers than among English children ($\beta$ = 0.30, $p$ = 0.02).
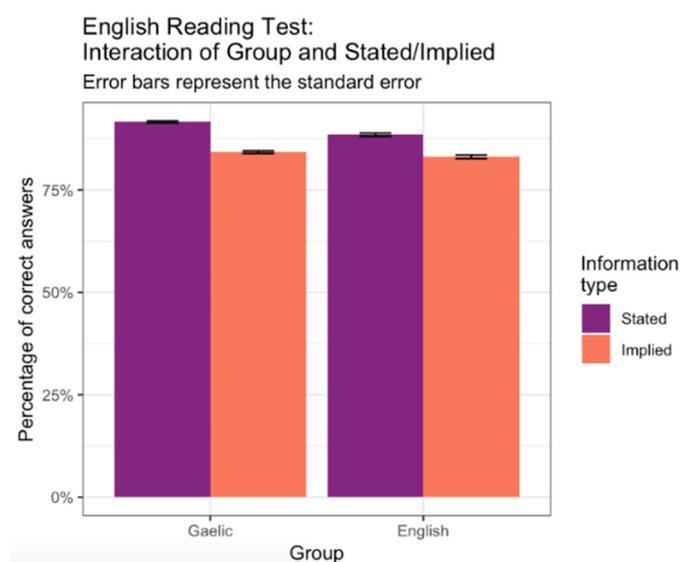


**Figure 8.** English Reading for All score by types of information and group.

## 4. Discussion

This paper presents data on reading in Gaelic/English bilingual children attending Gaelic Immersion Education. A first study was conducted to validate a reading assessment

for Gaelic and the main reading components (reading comprehension, reading errors and reading speed).

The study showed that the Gaelic version of the Reading for All assessment was an effective tool for analysing the effect of class and socioeconomic background, which was represented by the participant's SIMD number, on reading abilities for children. The results of the assessment demonstrate that school class was the most significant factor as children in later primary education performed better on the tests than children in earlier classes. Conversely, socioeconomic background had no significant effect on any of the test results. These findings matched expectations that the participants would become more fluent readers, i.e., scoring higher in the tests which measure passage comprehension while conversely having lower reading times and error rates, with each additional year of education. Furthermore, it was found that the variations in results decreased as class increased indicating that overall reading confidence increases across each year group.

Although the reading rate and error count increased with passage length and difficulty as expected, the children's reading comprehension scores in the test were found to increase as well, which contrasts with the hypothesis, found in Amendum et al. (2017), that a reader's comprehension would be affected by passages which are less readable in terms of length. Furthermore, although it might be assumed that the participants would find recalling short-lived information in our test more difficult, such as when answering questions relating to questions comprehending details (Haberlandt 1994), there were no significant differences found in the children's comprehension scores between correctly answering questions relating to main ideas and details.

Another finding of this study was that a direct correlation between the Gaelic and English versions of the assessment can be made for the children's comprehension scores. The analysis of this correlation demonstrated that Gaelic bilinguals performed better on the English adaptation than they did on the original Gaelic Reading for All assessment, as the pupils made fewer reading errors in a faster time in the English version. Furthermore, the pupils on average scored only another 0.6 points for reading comprehension in the Gaelic version for every additional correct point they achieved in the English. This finding was expected as O'Hanlon et al. (2013) reported that GME teachers believed that their bilingual Gaelic pupils were also performing better on average in English literacy than in Gaelic literacy. Given that these children receive considerably more Gaelic input at a GME school than they do English, it is important to consider that the vast majority of GME pupils in general and in this study are sequential bilinguals, often acquiring Gaelic in school. This means that they likely receive vastly reduced Gaelic input in other domains external to the school such as the home. Without further study, we are unable to say whether the dominance of English outside the classroom is resulting in the greater acquisition of English literacy skills, hampering the acquisition of Gaelic literacy skills, or both.

A second study reported a comparison of the linguistic competence in English of the participants involved in study 1 and an age-matched group of children living in or around Glasgow.

As was found in the first study, it was found that the participants' class or age had the greatest effect on performance across the reading variants: reading comprehension; reading errors; reading rate. Yet, for the purposes of the study, the most significant finding was that language group had a large effect on the child participants' performance, as English-speaking monolinguals performed statistically better than Gaelic bilinguals in both comprehension scores and reading errors in the YARC comprehension score and reading errors in the English adaptation of the Reading for All assessment. However, when the effects of both class and group are taken into consideration it is striking to see that the Gaelic bilingual children improved at a faster rate than the English monolinguals and that Gaelic children performed better than their monolingual peers between the classes primary 6 and 7 in the YARC comprehension score and reading errors and between the classes primary 5 and 6 in the English adaptation of the Reading for all assessment.

This accelerated improvement in tests for reading errors hint at the Gaelic bilingual pupils having greater phonological awareness, a finding which has also been evidenced in other studies, notably Bialystok et al. (2005). Despite widespread and long held concerns that bilingualism, particularly if one of these languages is Gaelic, would hamper a child's acquisition of English (McEwan-Fujita 2005), these results demonstrate that the children's literacy skills in Gaelic are easily transferred into English as predicted for example by Costa et al. (2018). Moreover, it was found from the analysis of the Reading for all results that Gaelic bilingual pupils were found systematically to be faster readers in all classes than their monolingual peers, suggesting a greater reading confidence. Furthermore, it was found that, despite receiving a greater amount of English language input in school, the English monolingual pupils performed systematically worse in the Trog-2 test than the bilingual pupils. This suggests the bilingual pupils have a greater degree of grammatical awareness of English even than their monolingual peers, supporting claims of superior executive language functions (Bialystok 2011; Baker 2001). Similar findings of greater grammatical awareness in the comprehension of complex sentences with the same methodology have been reported in other studies (for bilingual Sardinian/Italian children and adults see Garraffa et al. 2015, 2017; for bilingual Gaelic/English children Garraffa et al. 2020).

From analysing the comprehension scores for the type of content, bilingualism had no effect on the overall results for children. While comprehension of a passage's main ideas or its details had no effect on the children, our results suggest that the bilingual children did appear to struggle more with implicit content in comparison with stated content. It is possible that bilingualism does not give the reader an advantage or disadvantage over a monolingual in the inferring meaning overall. Gaelic bilingual children did appear to struggle more if the content was both implicit and related to a detail, suggesting that bilingual children are less adept at comprehending meaning at a local level in comparison with their monolingual peers. While this could be expected given that the bilingual pupils have less experience at reading in English, our data is not extensive enough at this stage to make any definitive assertions. Instead, more research into this area will need to be conducted.

In comparing the results of the two studies, it was observed that while the children are demonstrably less fluent in Gaelic than they are in English, by tracking their literacy development in each reading variant over the course of the upper years of primary school, it is clear that the children are improving in Gaelic literacy at a faster pace than they are in English literacy, so that this "fluency gap" was very slight. For example, by looking at the variant of reading comprehension in both the Gaelic and English versions of the Reading for all assessment, we see a larger fluency gap in primary 4, for which the children scored 11.05 in the Gaelic version and 20.73 in the English version, for a difference of 9.68, whereas this gap narrows significantly by primary 7, for which the children scored 20.54 in the Gaelic version and 25.54 in the English version, for a difference of 5.00.

## 5. Conclusions

This study tested the validity of a new reading assessment specifically developed for measuring literacy development in a Gaelic context. Moreover, as we have created both an English adaptation which can be used for gauging literacy development in English and two variations of the assessment in each language, the assessment can be used not only for comparing the development of a child's acquisition of literacy in both languages but also as an effective tool for tracking a child's literacy development at different intervals in their education.

Our results strongly support the assertion that the bilingual children achieve fluency in both Gaelic and English reading, as demonstrated through gauging their performance in the reading variants: reading comprehension, reading errors, and reading rate. While it is evident that the children are demonstrably less fluent in Gaelic than they are in English, it has been found that over the course of later primary school that their literacy development in Gaelic increases at a faster rate than it does in English. Furthermore, while in Primary

4 the bilingual children are less fluent than their monolingual peers of the same age in a few metrics, for example Gaelic bilinguals make more reading errors in English, the Gaelic bilinguals manage to catch up with the English reading fluency of their monolingual peers and a few reading variants surpass them by Primary 7. The only aspect of reading comprehension in which the English monolinguals appear to have a slight advantage is comprehending local or detailed information which is implicit. Further study of the comprehension of the different qualities of information in text between bilinguals and monolinguals is recommended.

**Author Contributions:** Conceptualization, E.D. and M.G.; methodology, L.M., M.O. and M.G.; formal analysis, M.O.; investigation, E.D. and L.M.; data curation, E.D., L.M., M.O. and M.G.; writing—original draft preparation, E.D., L.M. and M.G.; writing—review and editing, M.G.; project administration, M.G.; funding acquisition, M.G. All authors have read and agreed to the published version of the manuscript.

## References

Amendum, Steve, Kristine Conradi, and Elfrida Hiebert. 2017. Does Text Complexity Matter in the Elementary Grades? A Research Synthesis of Text Difficulty and Elementary Students' Reading Fluency and Comprehension. *Educational Psychology Review* 30: 121–51. [CrossRef]

Baker, Colin. 2001. *Foundations of Bilingual Education and Bilingualism*, 3rd ed. Clevedon: Multilingual Matters.

Barr, Dale, Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255–78. [CrossRef] [PubMed]

Bates, Douglas, Marting Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]

Benjamin, Rebekan George. 2011. Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review* 24: 63–88. [CrossRef]

Berens, Melody S., Ioulia Kovelman, and Laura-Ann Petitto. 2013. Should bilingual children learn reading in two languages at the same time or in sequence? *Bilingual Research Journal* 36: 35–60. [CrossRef] [PubMed]

Bialystok, Ellen. 2011. Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale* 65: 229–35. [CrossRef] [PubMed]

Bialystok, Ellen, Gigi Luk, and Ernest Kwan. 2005. Bilingualism, Biliteracy, and Learning to Read: Interactions Among Languages and Writing Systems. *Scientific Studies of Reading* 9: 43–61. [CrossRef]

Bishop, Dorothy. 2003. *Test for Reception of Grammar version 2 (TROG-2)*. London: Pearson.

Bishop, Dorothy V. M., and Margaret J. Snowling. 2004. Developmental Dyslexia and Specific Language Impairment: Same or Different? *Psychological Bulletin* 130: 858–86. [CrossRef] [PubMed]

British Dyslexia Association. 2018. Available online: https://www.bdadyslexia.org.uk/ (accessed on 20 May 2020).

Cain, Kate, and Jane Oakhill. 1999. Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal* 11: 489–503. [CrossRef]

Cenoz, Jasone, and Durk Gorter. 2017. Minority languages and sustainable translanguaging: Threat or opportunity? *Journal of Multilingual and Multicultural Development* 38: 901–12. [CrossRef]

Costa, Francesca, Stefania Sharley, and Maria Teresa Guasti. 2018. Does learning to read in English delay primary school children's literacy acquisition in Italian? *Rivista di Grammatica Generativa* 40: 5.

Chondrogianni, Vicky, Morna Butcher, and Maria Garraffa. 2021. Supporting children with Developmental Language Disorder in Gaelic-medium primary education. In *RnAG-Proceedings 2018*. Dublin: RnAG.

Dehaene, Stanislas. 2009. *Reading in the Brain. The New Science of How We Read*. New York: Penguin Viking.

Flege, James Emil, Grace Yeni-Komshian, and Serena Liu. 1999. Age constraints on second-language acquisition. *Journal of Memory and Language* 41: 78–104. [CrossRef]

Garraffa, Maria, Magdelein Beveridge, and Antonella Sorace. 2015. Linguistic and cognitive skills in Sardinian-Italian bilingual children. *Frontiers in Psychology* 6: 1898. [CrossRef] [PubMed]

Garraffa, Maria, Mateo Obregon, and Antonella Sorace. 2017. Linguistic and Cognitive Effects of Bilingualism with Regional Minority Languages: A Study of Sardinian–Italian Adult Speakers. *Frontiers Psychology* 8: 1907. [CrossRef] [PubMed]

Garraffa, Maria, Mateo Obregon, Bernadette O'Rourke, and Antonella Sorace. 2020. Language and Cognition in Gaelic-English Young Adult Bilingual Speakers: A Positive Effect of School Immersion Program on Attentional and Grammatical Skills. *Frontiers Psychology* 11: 570587. [CrossRef] [PubMed]

Graesser, Arthur, Keith Millis, and Rolf Zwaan. 1997. Discourse Comprehension. *Annual Review of Psychology* 48: 163–89. [CrossRef]

Greene, Jay. 1998. *A Meta-Analysis of the Effectiveness of Bilingual Education*. Claremont: Tomas Rivera Policy Institute.

Haberlandt, Karl. 1994. Methods in reading research. In *Handbook of Psycholinguistics*. Edited by Morton A. Gernsbacher. Cambridge: Academic Press, pp. 1–31.

Kintsch, Walter. 1988. The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review* 95: 163–82. [CrossRef]

Kovelman, Ioulia, Stephanie A. Baker, and Laura-Ann Petitto. 2008. Bilingual and monolingual brains compared: A functional magnetic resonance imaging investigation of syntactic processing and a possible "neural signature" of bilingualism. *Journal of Cognitive Neuroscience* 1: 153–69. [CrossRef] [PubMed]

Learning and Teaching Scotland. 2010. *Curriculum for Excellence: Literacy and Gàidhlig–Principles and Practice*; Glasgow: Learning and Teaching Scotland.

Lyon, Fiona, and Sarah Mac Quarrie. 2013. Assessment of reading skills in Gaelic Medium Education: Exploring teachers' perceptions and present practice. In *Additional Support Needs: A Collective Resource for GME*. Stornoway: Stòrlann, pp. 29–43.

McEwan-Fujita, Emily. 2005. Neoliberalism and Minority-Language Planning. *International Journal of the Sociology of Language* 171: 155–71.

National Records of Scotland. 2015. Scotland's Census 2011: Gaelic Report (Part 1). Available online: https://www.scotlandscensus.gov.uk/documents/census2021/Scotlands_Census_2011_Gaelic_Report_Part_1.pdf (accessed on 15 May 2020).

O'Hanlon, Fiona, Wilson McLeod, and Lindsay Paterson. 2010. *Gaelic-Medium Education in Scotland: Choice and Attainment at the Primary and Early Secondary School Stages*. Inverness: Bòrd na Gàidhlig.

O'Hanlon, Fiona, Lindsay Paterson, and Wilson McLeod. 2013. The attainment of pupils in Gaelic-medium primary education in Scotland. *International Journal of Bilingual Education and Bilingualism* 16: 707–29. [CrossRef]

Oakhill, Jane. 1984. Inferential and Memory Skills in Childrens Comprehension of Stories. *British Journal of Educational Psychology* 54: 31–39. [CrossRef]

Peirce, Jonathan, Jeremy Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51: 195–203. [CrossRef] [PubMed]

Perani, Daniela, Jubin Abutalebi, Eraldo Paulesu, Simona Brambati, Paola Scifo, Stefano F. Cappa, and Ferruccio Fazio. 2003. The role of age of acquisition and language usage in early, high-proficient bilinguals: An fmri study during verbal fluency. *Human Brain Mapping* 19: 170–82. [CrossRef] [PubMed]

Rolstad, Kellie, Kate Mahoney, and Gene Glass. 2005a. The big picture: A metaanalysis of program effectiveness research on english language learners. *Educational Policy* 19: 572–94. [CrossRef]

Rolstad, Kellie, Kate Mahoney, and Gene Glass. 2005b. Weighing the evidence: A meta-analysis of bilingual education in Arizona. *Bilingual Research Journal* 29: 43–67. [CrossRef]

Slavin, Robert, and Alan Cheung. 2003. *Effective Programs for English Language Learners: A Best-Evidence Synthesis*. Baltimore: John Hopkins University, CRESPAR.

Slavin, Robert, and Alan Cheung. 2005. A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research* 75: 247–84. [CrossRef]

Snowling, Margaret, Susan E. Stothard, Paula Clarke, Claudine Bowyer-Crane, Angela Harrington, Emma Truelove, K. Nation, and Charles Hulme. 2009. *York Assessment of Reading for Comprehension-Passage Reading*. Brentford: GL Assessment.

Van Dijk, Teun Adrianus, and Walter Kintsch. 1983. *Strategies of Discourse Comprehension*. Cambridge: Academic Press.

Webster, Janet, Julie Morris, David Howard, and Maria Garraffa. 2018. Reading for meaning: What influences paragraph understanding in aphasia? *American Journal of Speech-Language Pathology* 27: 423–37. [CrossRef] [PubMed]