Promoting bilingualism at the primary and secondary level:

The role of intelligence, motivation and anxiety

Verena Möller

Institut Langage et Communication, Université catholique de Louvain, Louvain-la-Neuve, Belgium

Dr. Verena Möller
Université catholique de Louvain
Institut Langage et Communication
Centre for English Corpus Linguistics
Place Blaise Pascal 1/L3.03.33
B-1348 Louvain-la-Neuve

– PREPRINT –

The Version of Record of this manuscript has been published and made available in *International Journal of Bilingual Education and Bilingualism*, 28 December 2018, https://www.tandfonline.com/doi/full/10.1080/13670050.2018.1559795

Promoting bilingualism at the primary and secondary level: The role of intelligence, motivation and anxiety

Following recommendations by the European Commission, changes have been implemented in the educational system of the German federal state of Baden-Württemberg in order to promote bilingualism. At the primary level, language classes and the concept of immersion were gradually introduced, while, at the secondary level, an increasing number of *Content and Language Integrated Learning* (CLIL) programmes were established. The present study analyses how these programmes interact with intelligence, motivation and anxiety. Data were collected from more than 400 learners in Year 11 by means of two psychometric tests. The effects of primary-level language lessons on intelligence proved to be minor. However, motivation tended to be higher and inhibiting exam anxiety lower the earlier pupils had started learning English. At the secondary level, a number of highly significant differences were revealed with respect to intelligence, particularly verbal aspects of intelligence, with CLIL participants outperforming their non-CLIL peers. Likewise, motivation proved to be greater in CLIL learners. Our results show that programmes promoting bilingualism not only influence L2 proficiency, but are related to cognitive and affective variables in ways that should be made explicit by curricula and taken into account by research on the linguistic outcomes of selective programmes such as CLIL.

Keywords: bilingual learning; immersion; CLIL; intelligence; motivation; anxiety

1 Introduction

In the past decades, a variety of assumptions have been voiced regarding language learning and, in particular, bilingual learning and its interaction with cognitive and affective factors. Regarding *Content and Language Integrated Learning* (CLIL), for instance, the European Commission states that

> Content and Language Integrated Learning (CLIL), in which pupils learn a subject through the medium of a foreign language, [...] opens doors on languages for a broader range of learners, nurturing self-confidence in

young learners and those who have not responded well to formal language instruction in general education. (European Commission 2003, 8)

This quote entails that both cognitive and affective variables play a part when methodologies are implemented: On the one hand, bilingual learning is assumed to foster self-confidence, on the other hand, it is viewed as compensating for a lack of cognitive capacities which might present an obstacle to succeeding in the traditional, academically oriented foreign language classroom. Section 2 illustrates how such assumptions impact the organization of and curricula for bilingual learning in Germany, and, specifically, the federal state of Baden-Württemberg, making it all the more surprising that cognitive and affective variables have rarely been investigated as dependent variables when assessing the impact of different methodologies. Even though recent studies from other countries (cf., for instance, De Smet et al. 2018) are beginning to close that gap, the overview of research literature for Germany in Section 3 shows that evidence is frequently no more than anecdotal.

In the present study, we seek to determine how intelligence, motivation and anxiety interact with educational settings promoting bilingualism at the primary and secondary level. Section 4 describes the participants, measures and statistical methods which were used. Section 5 details the results of the quantitative analysis of our data, which are discussed in Section 6.

2 Bilingual learning at the primary and secondary level

In its White Paper on education and training, the European Commission (1995: 47) states that '[i]t seems essential for [language] teaching to be placed on a systematic footing in primary education'. Following these requirements, most federal states in Germany decided to introduce language lessons in Year 3 (cf. KMK 2013a, 5). Like five other states, Baden-Württemberg opted for a different path, implementing language

learning programmes in the first year of primary education. Primary school curricula were therefore supplemented with syllabi for English or French in 2001, with English prevailing. In the same year, primary-level language learning programmes, which were allocated two lessons per week, commenced in 470 pilot schools (MKJS 2002, 2). In 2003, all other primary schools followed suit, with language learning being integrated into the regular curriculum in 2004 (cf. MKJS 2004).

Unfortunately, official communications do not always make a clear distinction between foreign language programmes and bilingual learning. A look at the two most recent primary school curricula in Baden-Württemberg confirms Dalton-Puffer et al.'s (2014, 214) statement that '[t]he words immersion and CLIL live the lives of words in natural languages: they have histories, migrate from one discourse to another, acquire connotations, and generally have fuzzy boundaries'. The syllabus in use with the participants in our study describes primary-level language lessons as *immersive-reflective*. While immersion is interpreted as exposure to authentic language referring to a situationally determined context, *immersive-reflective* is defined as allowing for the discussion of linguistic structures in a way that is appropriate to learners' age and proficiency (cf. MKJS 2004, 72). Very clearly, this use of the term *immersive* is not in accordance with earlier definitions, for instance by Genesee (1987, 1), who maintains that 'at least 50 percent of instruction during a given academic year must be provided through the second language for the program to be regarded as immersion'. Hence, the most recent primary school curriculum for Baden-Württemberg (cf. MKJS 2016) does not use the term *immersive* any longer. Instead, the new curriculum promotes the implementation of CLIL:

> Da sich Themen mit konkretem Lebensweltbezug in allen Sachfächern der Grundschule finden, wird die Fremdsprache so oft wie möglich nicht nur als

> Unterrichtssprache im Fremdsprachenunterricht genutzt, sondern auch in Sachfächern verwendet. [...] Fächerübergreifendes Arbeiten – die Integration des Fremdsprachenunterrichts in verschiedene Sachfächer in Form von CLIL – bietet sich ebenfalls an, um die lebensweltliche Relevanz der Unterrichtsinhalte zu verstärken. (MKJS 2016, 7)
>
> *(As topics that make concrete reference to students' immediate surroundings may be found in all content subjects in primary school, the foreign language is used as frequently as possible not only in the foreign language classroom, but also in content subjects. [...] Interdisciplinary lessons – the integration of foreign language teaching with various content subjects via CLIL – constitute a further way of enhancing the immediate relevance of content presented in lessons.)*

Interestingly, it is the syllabus of the language lesson that promotes the use of English in content subjects, while syllabi for content subjects do not take up this idea. In addition, the curriculum does not describe the extent to and the ways in which CLIL is to be integrated into primary education.

Despite the fuzziness regarding terminology, both the 2004 and the 2016 curriculum make clear that the ultimate goal of all measures adopted at the primary level is to lay a foundation for the development of multilingualism (MKJS 2004, 68 and MKJS 2016, 3). The present study uses the term *immersive-reflective lessons* (IRL), which is derived from the syllabus valid with our participants.

Apart from specifying requirements for the primary level, the White Paper on education and training states that 'secondary school pupils should study certain subjects in the first foreign language learned' (European Commission 1995, 47). Thus, more and more CLIL programmes were implemented in German schools. While in 1999, there

were 366 schools with a CLIL strand in Germany, their number had increased to more than 1,500 in 2013 (cf. KMK 2013b, 4), with the secondary level and, in particular, the *Gymnasium*, i.e. the type of school leading to qualification for university entrance, leading the way (cf. KMK 2013b, 10). Publications by the Baden-Württemberg Ministry of Education, Youth and Sports, however, add to the confusion regarding terminology. At the secondary level, programmes promoting bilingualism are called neither *immersion* nor *CLIL*, but *bilinguales Lehren und Lernen*, which translates as *bilingual teaching and learning*. In contrast to the programmes described for the primary level, the goal of bilingual teaching and learning at the secondary level is not the promotion of multilingualism in general, but, specifically, bilingualism and near-nativeness in the language which is used to teach content subjects:

> Bilinguales, anwendungsorientiertes Lernen eröffnet unseren Schülerinnen und Schülern den Weg zu einer annähernden Zweisprachigkeit. (MKJS 2008, 5)
>
> *(Bilingual, application-oriented learning opens the way to virtually complete bilingualism for our pupils.)*

The present study uses the term CLIL as, at the European level, this is the name which is generally adopted to refer to programmes which are organized in the same way as bilingual teaching and learning at the secondary level in Baden-Württemberg.

Despite the supposed 'CLIL for all' approach suggested by the European Commission (cf. Section 1), Germany has opted for a different path, which is why results obtained from studies of German CLIL are not necessarily comparable to results from other EU member states. The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany conceived of CLIL as a way of offering a challenge to the most gifted students in a cohort (cf. KMK

2013b, 5). In Baden-Württemberg, CLIL is considered to form part of a framework of options for pupils with above-average motivation and aptitude, who are willing to invest a greater amount of time and accept greater intellectual challenges (cf. MKJS 2008, 6). Hence, any selectivity occurring in these strands is fully intentional.

3 Bilingual learning and intelligence, motivation and anxiety

According to the European Commission's (1995, 47) White Paper on education and training, '[c]ontact with another language is not only compatible with becoming proficient in one's mother tongue, it also makes it easier'. The IRL syllabus makes explicit reference to this assumption:

> Der Aufbau von Fremdsprachenkenntnissen unterstützt die Entwicklung der Muttersprache und umgekehrt. (MKJS 2004, 72)
>
> *(The development of foreign language skills supports the development of a child's mother tongue and vice versa.)*

Along similar lines, Böttger (cf. 2010, 60) states that the challenge of communicating in English at an early age will enhance linguistic flexibility and skill, which will exert a positive effect on learners' L1 development. This claim might lead us to assume that language learning at the primary level will bring about a higher level of verbal cognitive capacities. However, Doyé and Lüttge (cf. 1977, 95) show that this is not the case, having conducted an intelligence test with learners in Year 5, some of whom had participated in primary-level English lessons from Year 1 to 4, while others had not. Significant differences were found with respect to neither verbal nor non-verbal cognitive capacities.

Apart from aims related to cognitive capacities, the IRL syllabus lists a number of non-linguistic goals, e.g. expanding personal competencies, forging an identity, and developing aspects of personality (cf. MKJS 2004, 69). Sarter (cf. 2002, 20-21) is more

explicit about what this entails. In a more episodic account of experiences, she reports that children in language learning programmes at the primary level display a high level of motivation, particularly those who are faced with difficulties in other school subjects. Furthermore, she observes that English is not an anxiety-ridden subject for those students as the development of linguistic proficiency at a young age seems to be unrelated to cognitive capacities. Wallaschek (cf. 2003, 21) confirms that children accept the focus on performance that is prevalent in primary-level language learning and are eager to fulfil expectations.

Unfortunately, empirical evidence supporting the aforementioned claims is scarce. While the implementation of language learning at the primary level has been evaluated with respect to proficiency as well as pedagogical aspects (cf. Bos and Pietsch 2007; Engel 2009; BIG-Kreis 2015), data assessing affective variables are few and far between. However, the BIG study, in which 2,148 children of about ten years of age were presented with a series of questions, allows for indirect insight into these factors. More than 80 % of all participants described English lessons as 'cool' (BIG-Kreis 2015, 14). This expression of attitudes may lead us to assume that children will also display a high level of motivation for language learning. In addition, the study (cf. BIG-Kreis 2015, 13) reports that children who experience a low level of anxiety when speaking in front of their classmates tend to have better marks. At a mean grade of 1.9, learners were found to do better in English than they did in other school subjects, which may be due to learners experiencing less anxiety. An alternative explanation which is offered, however, is that teachers are simply less strict with their marking in language lessons than in other subjects.

Even though there are an extensive number of German studies in which the effects of CLIL at the secondary level are assessed, the selectivity of the system

described in the previous section has rarely been taken into account. Some studies, however, suggest that cognitive and affective variables are indeed relevant.

Zydatiß (2007, 126), for instance, reports that CLIL participants receive better overall marks. In addition, CLIL learners display greater linguistic competence in their L1 German (cf. Klieme et al. 2006, 59), as well as possessing a higher level of L2 competence even prior to the onset of CLIL (cf. Bos et al. 2009, 40-41). This suggests that CLIL participants, a priori, have better general and verbal cognitive capacities, which is confirmed by Fehling (cf. 2009, 59), who tested CLIL participants and non-participants for verbal intelligence. In her study, it was shown that CLIL participants, at the beginning of the CLIL programme, performed significantly better than their non-CLIL peers. No significant increase was found over a spell of two years in either group, which suggests that it is not CLIL that brings about higher levels of verbal intelligence, but that CLIL programmes are selective.

Empirical evidence regarding the role of motivation in the German CLIL system is scarce. Rumlich (cf. 2014, 86) reports that CLIL participants display higher levels of both language-related and subject-related interest, which suggests that these learners might also possess greater motivation, at least for language lessons and CLIL lessons.

To our knowledge, no studies have been undertaken with respect to anxiety in German CLIL and non-CLIL learners. Zydatiß (cf. 2007, 135), however, reports that CLIL participants generally displayed a higher level of confidence regarding their future success in language education. In other words, they might experience a lower level of language-related anxiety. On the other hand, Zydatiß (cf. 2007, 147) observes that educational success is not respected at the secondary level, in particular amongst CLIL participants' non-CLIL peers. Hence, CLIL learners might experience greater fear of stigmatization.

4 Method

The data on which the present study is based are derived from an L2 database comprising the learner corpus *SCooLE (Secondary-Level Corpus of Learner English)*, experimental data on passive constructions and a collection of metadata on more than fifty learner variables (cf. Möller 2017, 93-99).

4.1 Participants

Data were collected from 419 learners who were attending Year 11 in 2010/11 and 2011/12 at various secondary schools *(Gymnasien)* across the federal state of Baden-Württemberg. The fact that data collection took place during the year preceding students' final examinations entails that, when assessing how educational settings at the primary level and cognitive and affective learner variables are interrelated, we are faced with long-term effects. The majority of participants had not attended the pilot schools which had implemented IRL from Year 1 onwards (cf. Section 2). While most learners had not taken part in IRL at all, there was a sizable group of students who had had the chance of participating in IRL, though not for the full four years envisaged by the 2004 curriculum (cf. MKJS 2004, 74). We therefore contrast the following groups:

(1) IRL0 (n = 228): These students had started attending language lessons in Year 5, i.e. on entering secondary school.

(2) IRL2 (n = 116): This group were able to attend IRL, though only for about two years of their primary education.

(3) IRL4 (n = 71): These are learners who attended IRL pilot schools, thus benefitting from four years of IRL.

Regarding CLIL and non-CLIL settings at the secondary level, one of the issues of past studies has frequently revolved around the fact that both CLIL and non-CLIL

groups came from the same schools, which reinforced the impact of selectivity and has since led to the complaint that '[i]t is very convenient to select and then demonstrate that the selected perform better than the non-selected with additional language exposure' (Bruton 2011, 530). This is the reason why we opted for data collection in two different non-CLIL groups, leading to the following division with respect to CLIL:

(1) CLIL+ (*n* = 161): These students had participated in CLIL strands commencing in Year 7, with reinforced language teaching taking place in Year 5 and 6. To qualify as a CLIL participant in the present study, learners needed to have taken part in CLIL up to the end of Year 10, i.e. the year preceding data collection.

(2) CLIL0 (*n* = 200): Learners in this group were attending a regular *Gymnasium*, i.e. schools which had not implemented a CLIL programme. Therefore, this is the only group in the design of the study which was entirely unaffected by the selectivity inherent in the CLIL system.

(3) CLIL- (*n* = 57): This group comprises non-CLIL learners from schools with CLIL strands.

4.2 Measures

In Second Language Acquisition (SLA) research, cognitive and affective variables have mostly been discussed in terms of concepts referring uniquely to the foreign language (FL) classroom. Cognitive capacities have been assessed referring to language aptitude (cf., for instance, the Modern Languages Aptitude Test, MLAT, Carroll & Sapon 1959) rather than intelligence, even though it has been argued that 'the contention that aptitude and intelligence are independent entities is [...] invalid' (Teepen 2006, 5). Likewise, motivation for language learning has often been viewed as being different from motivation governing other learning processes (cf., for instance, Gardner and Lambert 1959, Kruidenier and Clément 1986, Dörnyei 2009). In addition, rather than discussing

trait or state anxiety, SLA research has used the notion of FL anxiety (cf., for instance, the Foreign Language Classroom Anxiety Scale, FLCAS, Horwitz, Horwitz, and Cope 1986), which is 'an umbrella term for general FL anxiety and skill-specific anxieties like communicative, reading, and even pronunciation anxiety in a FL' (Jin and Dewaele 2018, 150). Unfortunately, these SLA-specific notions are often lacking in relevance in educational contexts as this type of data is not usually available, and is thus unable to serve as a basis for interventions in the classroom or in educational counselling. For this reason, we have used data which are available within the educational system where our data were collected. At the beginning of each data collection session, we conducted two psychometric tests, the intelligence test PSB-R 6-13[1] (cf. Horn 2003) and a test on motivation and anxiety, FLM 7-13[2] (cf. Petermann and Winkel 2007).

PSB-R 6-13 is a test which was designed for Year 6 to 13 in secondary education, i.e. for subjects who are generally between 11 and 19 years of age. It comprises nine subtests, each of which is timed. The individual subtests are merged to form clustered scales. All scales are listed in Table 1 with their Cronbach's α values.[3]

---

1. PSB-R 6-13 stands for *Prüfsystem für Schul- und Bildungsberatung für 6. bis 13. Klassen, Revidierte Fassung*, i.e. test system for educational counselling for Year 6 to 13, revised version.

2. FLM 7-13 stands for *Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse*, i.e. achievement motivation questionnaire for pupils attending Year 7 to 13.

3. As PSB-R 6-13 was constructed specifically for the German system of education, the German terminology was translated on the basis of Rost's summary of Thurstone's theory of primary mental abilities (cf. Rost 2009, 32-36). Cronbach's α values indicate reliability. Field (2009, 675) notes that, for Cronbach's α, 'the generally accepted value of .8 is appropriate for cognitive tests such as intelligence tests'. A caveat is in place as 'the value of α depends on the number of

Raw values for each scale are converted into a standard score with a mean of 100 and a standard deviation of 10, thus deviating slightly from the more common IQ scale, which uses a standard deviation of 15. As the test relies on the German language, corrective measures are applied for non-native speakers of German.

[Table 1 near here]

The *overall verbal* scale is composed of subtests on verbal comprehension, word fluency and verbal relations. The subtest on *verbal comprehension* is divided into five tasks, each of which comprises a content-related heading with a list of twenty words. In each of these words, one letter is substituted with a false one, and students have to recognize the words that fit the respective headings and indicate where the misspelling occurs. The subtest therefore measures general knowledge along with verbal comprehension (cf. Horn 2003, 7). The *word fluency* scale assesses how many words with predetermined initial letters a subject is able to produce during a limited period of time. The subtest on *verbal relations* aims at understanding the meaning of words. Within groups of five nouns, four are semantically related. The student has to find the fifth noun, which is the odd one out. This subtest is assumed to be of particular relevance to educational success (cf. Horn 2003, 8).

---

items on the scale' (cf. Field 2009, 675). Subtests on reasoning (numbers, letters, figures) include a comparatively low number of items (15 each, as opposed to, for instance, 100 in the verbal comprehension subscale). It is therefore natural that the reported Cronbach's α is lower for these subtests. The reliability of clustered scales is higher for the same reason. According to Horn (cf. 2003, 16), Cronbach's α cannot be calculated for numerical facility and perceptual speed as these tests are speed tests. The same is true for all clustered scales comprising results obtained from these two subtests.

The *overall reasoning* scale comprises three subtests on non-verbal reasoning and one on spatial perception. In the subtests on *reasoning (numbers, letters, figures)* students are presented with sequences of numbers, letters and geometric figures and are required to find the rule according to which the order of these sequences was established. One item does not match the rule and has to be found by the student. To measure *spatial perception*, students are presented with drawings of three-dimensional geometric figures which have a varying number of surfaces. Students have to indicate how many surfaces the three-dimensional object corresponding to the drawing has.

In the *overall concentration* scale, results regarding numerical facility and perceptual speed are merged. To assess *numerical facility*, students are presented with sequences of numbers for which they have to calculate sums. To measure *perceptual speed*, students are asked to compare two sequences of numbers which differ by one digit, and then have to find the diverging digit.

The results of all aforementioned scales are merged in a clustered scale concerned with *overall cognitive capacities*.

FLM 7-13 is a questionnaire which assesses factors relevant to achievement motivation. It comprises five individual scales, three of which are related to aspects of anxiety. All five scales are listed in Table 2 with their Cronbach's α values.[4] Raw values are transformed into a standard score with a mean of 50 and a standard deviation of 10. FLM 7-13 comes with separate norms for male and female subjects.

[Table 2 near here]

4. Field (2009, 675) states that 'when dealing with psychological constructs values below even .7 can, realistically, be expected because of the diversity of the constructs being measured'. In fact, Petermann and Winkel (cf. 2007, 29) rate results which are higher than .7 as satisfactory, while considering a Cronbach's α value between .6 and .7 as acceptable.

Orientation towards performance and success combines the concept of hope of success, i.e. a belief on the part of students that they will be able to achieve their aims, with the desire to expand the scope of their competence and to make this competence visible via performance. The dimension of *perseverance and effort* combines information about self-control and self-discipline with the idea of hope of success. These components are assumed to make students persevere in their efforts despite possibly adverse circumstances.

Fear of success can be described as 'a psychological barrier to achievement, aroused by the expectation that success in certain contexts may result in negative consequences, including [...] social rejection' (André and Metzler 2011, 416). In other words, students with high scores will display some conviction that success in education will lead to jealousy in their classmates or even bullying. Regarding the ways in which anxiety may be related to performance, Alpert and Haber (1960, 212) list three possibilities: '[I]t will facilitate it, debilitate it, or perhaps have no effect on it at all.' The concept of *activating exam anxiety* is based on the first of these possibilities. Dewaele et al. (2016, 57) define facilitating anxiety as follows: '[F]acilitating anxiety should not be considered a lower level or amount of anxiety (below a hypothetical "threshold") but rather a qualitatively different [...], positive emotion related to excitement, risktaking, arousal, eagerness, and so on.' Thus, it is assumed that exam anxiety may lead to students experiencing themselves as being particularly efficient as a result of a challenging situation. *Inhibiting exam anxiety*, by contrast, describes the debilitating aspect of anxiety. In contrast to activating exam anxiety, students experience a passive fear of failure.

4.3 Statistical analysis

To compare the scores obtained by students from the predefined educational settings (cf. Section 4.1), we used one-way independent ANOVA. Our data were checked for homogeneity of variance via Levene's test and for normality via visual assessment and the calculation of *Z*-scores for skewness and kurtosis. Most, though not all, of our data were found to be normally distributed. Whenever this was not the case, we followed Ghasemi and Zahediasl (2012, 486), who state that 'in large samples (> 30 or 40), the sampling distribution tends to be normal, regardless of the shape of the data', and conclude that 'this implies that we can use parametric procedures even when the data are not normally distributed'. As a post hoc test, we used Gabriel's pairwise test, which is 'designed to cope with situations in which sample sizes are different' (Field 2009, 374). Effect sizes are reported as ω (cf. Field 2009, 389).[5]

Even though outliers may cause a small bias, they were retained rather than changed or eliminated. Regarding the elimination of a case, Field (2009, 153) states that 'this should be done only if you have good reason to believe that this case is not from the population that you intended to sample' (Field 2009, 153). Hence, outliers reflecting, for instance, exceptionally high intelligence, were not deleted from the data set as they reflect diversity in the classroom.

In several cases, we refer to individual items of FLM 7-13 in order to shed light on the most relevant types of cognitions and behaviours associated with individual scales. As responses were given on a 5-point Likert scale, we consider these data as

---

5. Whenever ω^2 is negative because the *F*-ratio falls between 0 and 1, we follow Keppel and Wickens (2004, 164): 'Under these circumstances, it is best to set $\omega^2 = 0$. This value does not prove that treatment effects are absent (nothing can do that), only that what differences there are can reasonably be attributed to random variation.' We therefore report $\omega = 0$.

ordinal rather than interval data (cf. Field 2009, 8), which is why we have chosen to adopt non-parametric procedures in this part of the analysis. To assess the impact of IRL in these cases, we used the Jonckheere trend test. For the secondary level, the Kruskal-Wallis test was used, with Mann-Whitney tests with a Bonferroni correction acting as post hoc procedures. Several individual items presented a problem in terms of homogeneity of variance. In these cases, log or cube transformations were used.

5 Results

When analysing the data from the L2 database, it is essential to bear in mind that data collection took place when students were in Year 11. For the analysis of educational settings at the primary level, this means that we are confronted with the long-term effects of treatment that students were subjected to at least seven years prior to the two psychometric tests being administered. Educational settings at the secondary level, by contrast, varied until the year preceding data collection, i.e. Year 10. It is therefore unsurprising that results for the secondary level are, in many cases, clearer and more significant than those obtained for the primary level.

When applying one-way independent ANOVA to data obtained from the intelligence test, PSB-R 6-13, virtually no significant differences between the three educational settings at the primary level were detected (cf. Table A1 in the Appendix for the complete set of results), with the sole exception being the *verbal comprehension* scale, $F(2, 406) = 3.17$, $p < .05$, $\omega = .10$. A look at group means shows that it is the group who took part in IRL for four years (IRL4) which, at $M_{IRL4} = 105.93$, differs from both IRL0 ($M_{IRL0} = 108.29$) and IRL2 ($M_{IRL2} = 108.40$) at a descriptive level, even though only the difference between IRL4 and IRL0 is statistically significant. As was described in Section 4.2, students had to recognize words in their L1, German, in this subtest. While an onset of language education around Year 3 does not seem to affect

this ability at all, an early onset in the first year of primary education is associated with lower word recognition abilities in students' mother tongue.

Regarding motivation, results from FLM 7-13 reveal a significant difference between students from different educational settings regarding *orientation towards performance and success*, $F(2, 412) = 4.00$, $p < .05$, $\omega = .12$, with the IRL4 group, at $M_{IRL4} = 56.18$, attaining a significantly higher level than IRL0 students ($M_{IRL0} = 52.88$). In particular, the Jonckheere trend test revealed significant trends regarding the following items:

- *I aspire to be one of the best students at school:* $T_{JT} = 28{,}876.00$, $z = 2.86$, $p < .01$.
- *I strive to perform better than most other students:* $T_{JT} = 28{,}458.00$, $z = 2.48$, $p < .05$.

In other words, the longer students are subjected to English lessons at the primary level, the more they will strive to become good students. Even though no significant difference was revealed when applying ANOVA to data on *perseverance and effort*, group means (cf. Table A2 in the Appendix) suggest that students will display this phenomenon to a greater extent the longer they have taken part in IRL. In particular, a significant trend was revealed for the item *When I cannot perform a task immediately, I try everything to find a solution*, $T_{JT} = 28{,}477.50$, $z = 2.53$, $p < .05$.

While levels of motivation increase the longer students have taken part in IRL, *inhibiting exam anxiety* decreases, $F(2, 412) = 3.07$, $p < .05$, $\omega = .10$. Gabriel's pairwise test detected a significant difference between IRL0 ($M_{IRL0} = 53.59$) and IRL4 students ($M_{IRL4} = 50.62$). At a descriptive level, it becomes clear that the IRL2 group, at $M_{IRL2} = 52.06$, is situated in between. This means that the longer students have taken

part in IRL, the less they are afraid of failing in test situations. Significant trends were detected for the following two items:

- *When I have to perform a difficult task, I am afraid of failing:* T_{JT} = 22,330.00, z = -2.52, $p < .05$.
- *I only take part in classroom discussions when I am certain that my answer is correct:* T_{JT} = 22,791.00, z = -2.15, $p < .05$.

Results concerning the aforementioned subscales of FLM 7-13 are illustrated in Figure 1.

[Figure 1 near here]

The application of one-way independent ANOVA to data on intelligence revealed manifold significant differences between CLIL+ learners and their peers from non-CLIL strands (cf. Table A3 in the Appendix). These discrepancies centre around verbal aspects of intelligence, while for the non-verbal scales of PSB-R 6-13, significant differences were found only with respect to reasoning/figures and numerical facility. For verbal aspects of intelligence, however, all subscales proved to be relevant (cf. Figure 2):

- *Verbal comprehension:*

  A highly significant interaction was found between educational setting and verbal comprehension, $F(2, 408) = 10.33$, $p < .0001$, $\omega = .21$. CLIL+ learners, at $M_{CLIL+} = 109.54$, were found to possess a higher level of verbal comprehension capacities than the CLIL- group ($M_{CLIL-} = 104.43$), who were also outperformed by CLIL0 learners ($M_{CLIL0} = 107.72$).
- *Word fluency:*

  Even though the application of ANOVA suggests a significant difference

between group means, $F(2, 408) = 3.09$, $p < .05$, $\omega = .10$, this effect seems to be minor. Gabriel's pairwise test reports the lowest *p*-value for the comparison between CLIL+ ($M_{CLIL+} = 100.01$) and CLIL0 ($M_{CLIL0} = 98.29$) learners. However, this effect remains insignificant.

- *Verbal relations:*

  A significant effect was found for verbal relations, $F(2, 408) = 6.63$, $p < .01$, $\omega = .16$, with CLIL+ students, at $M_{CLIL+} = 109.46$, outperforming both CLIL0 ($M_{CLIL0} = 107.20$) and CLIL- learners ($M_{CLIL-} = 105.09$).

- *Overall verbal:*

  Results obtained for the aforementioned subscales led to a highly significant effect being revealed for the overall verbal scale, $F(2, 408) = 10.92$, $p < .0001$, $\omega = .21$. Significant differences were found between the CLIL+ group ($M_{CLIL+} = 107.10$) and both CLIL0 ($M_{CLIL0} = 104.93$) and CLIL- learners ($M_{CLIL-} = 103.00$).

[Figure 2 near here]

Regarding aspects of motivation (cf. Table A4 in the Appendix), ANOVA showed a significant effect for *orientation towards performance and success*, $F(2, 413) = 5.27$, $p < .01$, $\omega = .14$. A significant difference was found to exist between CLIL+ ($M_{CLIL+} = 55.76$) and CLIL- learners ($M_{CLIL-} = 51.19$), with the former attaching more importance to their educational success than the latter. When looking at individual items, it becomes clear that CLIL+ learners have a tendency to be more ambitious than their non-CLIL peers in that they wish to be amongst the best students, they want their results to be above average, and they prefer tasks which present a challenge to them:

- *I aspire to be one of the best students at school*, $H(2) = 7.52$, $p < .05$: The CLIL+ group agreed to this statement significantly more than both the CLIL0 and the CLIL- group.
- *I strive for my performance to be better than average*, $H(2) = 8.18$, $p < .05$: CLIL+ students agreed to this statement to a significantly greater extent than CLIL- learners.
- *I prefer working on tasks which present a challenge to me*, $H(2) = 9.10$, $p < .05$: CLIL+ students reported this preference significantly more than the CLIL0 group.

Advantaged as CLIL+ students may seem, they pay a price by displaying significantly greater *fear of success* than their non-CLIL counterparts, $F(2, 413) = 11.26$, $p < .0001$, $\omega = .22$. At $M_{CLIL+} = 56.16$, CLIL participants, who, as was shown above, are both the most intelligent and the most motivated learners, are also more afraid of the consequences of their success than both the CLIL0 ($M_{CLIL0} = 51.42$) and the CLIL- group ($M_{CLIL-} = 52.04$). While in the non-CLIL groups, only about a quarter of all participants displayed above-average fear of success (i.e. they attain a standard score > 60), this was the case for more than 40 % in the CLIL+ group. An analysis of individual items reveals that CLIL participants have a greater tendency to believe that taking part in classroom discussions and making an effort with their work in general will make them unpopular. Therefore, they also have a greater tendency to hide good marks from their peers:

- *Students who make a greater effort than most of their classmates cause themselves to become unpopular with their fellow students*, $H(2) = 7.01$, $p < .05$:

CLIL+ learners agreed to this statement to a significantly greater extent than CLIL0 students.

- *If you participate a lot in classroom discussions, others will assume that you are a swot*, $H(2) = 24.24$, $p < .00001$: The CLIL+ group reported this belief significantly more than both CLIL0 and CLIL- learners.
- *When I have done exceptionally well in a class test, I do not want other students to know about it*, $H(2) = 20.34$, $p < .0001$: CLIL+ students agreed significantly more to this statement than both the CLIL0 and the CLIL- group.

6 Discussion

As was outlined in Section 3, programmes promoting bilingualism at the primary level have sometimes been said to support not only L2 acquisition, but also children's linguistic development in L1 and even linguistic flexibility and skill in general. Our results regarding differences in verbal cognitive capacities, however, are similar to those reported by Doyé and Lüttge (cf. 1977, 95), who did not find participation in early language classes to impact upon verbal (and also non-verbal) cognitive capacities at all. Our analysis of aspects of intelligence yielded a single significant result regarding primary-level language education, which, in addition, is opposed to the position assumed by the authors of the IRL syllabus, who argue that language classes at the primary level support L1 development (cf. MKJS 2004, 72). The results of the present study suggest that, to a certain extent, early language classes impair linguistic development in L1 as early learners of English mastered word recognition tasks in their L1 less successfully than learners whose language classes had started no sooner than Year 5. The effect size associated with this sole significant result is small, and the difference is unlikely to be detectable in real life at all. Nonetheless, our results suggest

that the claim put forward by the IRL syllabus, namely that early language learning aids L1 development, has yet to be validated by empirical evidence for the German context.

Other goals which have been defined with respect to primary school language education, however, are indeed supported by our data. The IRL syllabus (cf. MKJS 2004, 69) states that one of the goals of IRL is the development and expansion of personal competencies. While no examples are provided as to which competencies are targeted, the development of motivation certainly qualifies as one such personal competency. In the present study, orientation towards performance and success was found to increase the longer students had taken part in IRL. Inhibiting exam anxiety, on the other hand, decreased, which suggests that the aim of developing self-confidence in language lessons at the primary level (cf. Böttger 2010, 60) is, to a certain extent, a realistic one.

The present study calls for more research into the role of primary-level language education for cognitive and affective learner variables, such that curricula, in the future, will be able to be more specific about requirements regarding the development of personal competencies. Such research, in addition, should differentiate more between the impact of language education and that of immersive or CLIL programmes than is currently possible on the basis of existing curricula for the primary level.

Regarding bilingual learning at the secondary level, Küppers and Trautmann (2013, 294) state that 'the underlying secret of the outstanding CLIL reputation cannot only be attributed to its theoretical underpinnings or sophisticated teaching methodology. The biggest secret seems, quite simply: bright and enthusiastic learners'. Our data amply confirm this statement in that CLIL is revealed to be selective with respect to both intelligence, in particular verbal aspects of intelligence, and orientation towards performance and success as an aspect of motivation. Even beyond the highly

significant differences reported in the previous section, it is observable even at a purely descriptive level that selectivity is at work in schools with a CLIL section: The CLIL0 group, i.e. the group of students who attended schools without a CLIL strand, mostly attained means between the ones reported for CLIL+ and CLIL- students. In other words, while groups in schools without a CLIL section are diverse regarding both intelligence and motivation, schools with a CLIL programme sort the students with greater cognitive capacities and higher levels of motivation into CLIL classes, while the rest remain in the non-CLIL set. Seeing that this selectivity is intentional and even mentioned explicitly in official CLIL publications (cf. MKJS 2008, 6), our results are entirely unsurprising. What is surprising is that this selectivity has rarely been taken into account when assessing linguistic outcomes in CLIL strands. Even studies which do report such differences (cf., for instance, Zydatiß 2007) have yet to adopt statistical procedures to distinguish between the effect of CLIL itself and the effect of its selectivity, as has been done by Möller (2017), who showed that the impact of selectivity is greater than that of CLIL itself.

Our results regarding fear of success suggest that CLIL participants are very aware of having been subjected to a process of selection, and that consequences are not necessarily positive. This should lead us to question the German view of CLIL as an offer for a positively selected group. If a 'CLIL for all' approach is not envisaged, the current system needs to be supported by assessment and regulation of group processes as well as individual counselling support for both negatively selected students, who may feel inferior, and positively selected students, who may experience problems due to that selection. Using instruments for research that stem from the educational system itself, as has been done in the present study, will allow for both types of intervention.

Word count

6954 words

Acknowledgments

Use of the psychometric test FLM 7-13 was supported by Pearson Assessment & Information GmbH.

Declaration of interest statement

The author has no conflict of interest to disclose.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. Data are not publicly available due to their containing sensitive personal information (e.g. intelligence data) obtained from minors. The Baden-Württemberg Ministry of Education, Youth and Sports has approved use of the data set for research purposes, but not general availability.

References

Alpert, Richard, and Ralph Norman Haber. 1960. "Anxiety in Academic Achievement Situations." *Journal of Abnormal and Social Psychology* 61 (2): 207-215.

André, Nathalie, and Jonathan N. Metzler. 2011. "Gender Differences in Fear of Success: A Preliminary Validation of the Performance Success Threat Appraisal Inventory." *Psychology of Sport and Exercise* 12: 415-422.

BIG-Kreis, eds. 2015. *Der Lernstand im Englischunterricht am Ende von Klasse 4. Ergebnisse der BIG-Studie.* http://www.fff-konferenz.de/tl_files/BIG-Studie-2015/Big-Studie-2015_Lernstand-Englischunterricht-Klasse-4.pdf

Böttger, Heiner. 2010. *Englisch lernen in der Grundschule.* 2nd ed. Bad Heilbrunn: Julius Klinkhardt.

Bos, Wilfried, and Marcus Pietsch. 2007. *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen.* Münster: Waxmann.

Bos, Wilfried, Martin Bonsen, Carola Gröhlich, Karin Guill, Peter May, Anna Rau, Tobias C. Stubbe, Ulrich Vielus, and Hans Wocken. 2009. *KESS 7. Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 7.* http://bildungsserver.hamburg.de/contentblob/2627296/data/pdf-kess-7.pdf

Bruton, Anthony. 2011. "Is CLIL So Beneficial, Or Just Selective? Re-evaluating Some of the Research." *System* 39: 523-531.

Carroll, John B., and Stanley M. Sapon. 1959. *Modern Languages Aptitude Test – Form A.* New York: Psychological Corporation.

Dalton-Puffer, Christiane, Ana Llinares, Francisco Lorenzo, and Tarja Nikula. 2014. "'You Can Stand Under My Umbrella': Immersion, CLIL and Bilingual Education.

A Response to Cenoz, Genesee & Gorter (2013)." *Applied Linguistics* 35 (2): 213-218.

De Smet, Audrey, Laurence Mettewie, Benoit Galand, Philippe Hiligsmann, and Luk Van Mensel. 2018. "Classroom Anxiety and Enjoyment in CLIL and Non-CLIL: Does the Target Language Matter?" *Studies in Second Language Learning and Teaching* 8 (1): 47-71.

Dewaele, Jean-Marc, Peter MacIntyre, Carmen Boudreau, and Livia Dewaele. 2016. "Do Girls Have All the Fun? Anxiety and Enjoyment in the Foreign Language Classroom." *Theory and Practice of Second Language Acquisition* 2 (1): 41-63.

Dörnyei, Zoltán. 2009. "The L2 Motivational Self System". In *Motivation, Language Identity and the L2 Self,* edited by Zoltán Dörnyei, and Ema Ushioda, 9-42. Bristol: Multilingual Matters.

Doyé, Peter, and Dieter Lüttge. 1977. *Untersuchungen zum Englischunterricht in der Grundschule. Bericht über das Braunschweiger Forschungsprojekt 'Frühbeginn des Englischunterrichts', FEU.* Braunschweig: Westermann.

Engel, Gaby. 2009. "EVENING – Konsequenzen für die Weiterentwicklung des Englischunterrichts in der Grundschule." In *Englisch in der Primarstufe – Chancen und Herausforderungen. Evaluation und Erfahrungen aus der Praxis,* edited by Gaby Engel, Bernd Groot-Wilken, and Eike Thürmann, 197-215. Berlin: Cornelsen.

European Commission. 1995. *White Paper on Education and Learning: Teaching and Learning. Towards the Learning Society.* http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf

European Commission. 2003. *Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee of the Regions.*

Promoting Language Learning and Diversity: An Action Plan 2004 – 2006. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52003DC0449&from= EN

Fehling, Sylvia. 2009. "Lernprozesse und kognitive Entwicklung im bilingualen Unterricht: Bericht aus einer zweijährigen Longitudinalstudie." In *Bilingualer Unterricht macht Schule. Beiträge aus der Praxisforschung.* 2nd ed., edited by Daniela Caspari, Wolfgang Hallet, Anke Wegner, and Wolfgang Zydatiß, 51-63. Frankfurt am Main: Peter Lang.

Field, Andy. 2009. *Discovering Statistics Using SPSS (and Sex and Drugs and Rock 'n' Roll).* 3rd ed. Los Angeles: Sage.

Gardner, Robert C., and Wallace E. Lambert. 1959. "Motivational Variables in Second Language Acquisition." *Canadian Journal of Psychology* 13 (4): 266-272.

Genesee, Fred. 1987. *Learning through Two Languages: Studies in Immersion and Bilingual Education.* Cambridge, MA: Newbury House.

Ghasemi, Asghar, and Saleh Zahediasl. 2012. "Normality Tests for Statistical Analysis: A Guide for Non-Statisticians." *International Journal of Endocrinology & Metabolism* 10 (2): 486-489.

Horn, Wolfgang. 2003. *PSB-R 6-13. Prüfsystem für Schul- und Bildungsberatung für 6. bis 13. Klassen – revidierte Fassung.* Göttingen: Hogrefe.

Horwitz, Elaine K., Michael B. Horwitz, and Joann Cope. 1986. "Foreign Language Classroom Anxiety." *The Modern Language Journal* 70 (2): 125-132.

Jin, Yin Xing, and Jean-Marc Dewaele. 2018. "The Effect of Positive Orientation and Perceived Social Support on Foreign Language Classroom Anxiety." *System* 74: 149-157.

Keppel, Geoffrey, and Thomas D. Wickens. 2004. *Design and Analysis: A Researcher's Handbook.* 4th ed. Upper Saddle River, NJ: Pearson Prentice Hall.

Klieme, Eckhard, Wolfgang Eichler, Andreas Helmke, Rainer H. Lehmann, Günter Nold, Hans-Günter Rolff, Konrad Schröder, Günther Thomé, and Heiner Willenberg. 2006. *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistungen International (DESI).* www.dipf.de/de/forschung/projekte/pdf/biqua/desi-zentrale-befunde

KMK (Kultusministerkonferenz). 2013a. *Bericht 'Fremdsprachen in der Grundschule – Sachstand und Konzeptionen 2013'. Beschluss der Kultusministerkonferenz vom 17.10.2013.* http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2013/2013_10_17-Fremdsprachen-in-der-Grundschule.pdf

KMK (Kultusministerkonferenz). 2013b. *Bericht 'Konzepte für den bilingualen Unterricht – Erfahrungsbericht und Vorschläge zur Weiterentwicklung'. Beschluss der Kultusministerkonferenz vom 17.10.2013.* http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/201_10_17-Konzepte-_bilingualer-_Unterricht.pdf

Kruidenier, Bastian G., and Richard Clément. 1986. *The Effect of Context on the Composition and Role of Orientations in Second Language Acquisition*. Québec: International Centre for Research on Bilingualism.

Küppers, Almut, and Matthias Trautmann. 2013. "It Is Not CLIL That Is a Success – CLIL Students Are! Some Critical Remarks on the Current CLIL Boom." In *Content and Language Integrated Learning (CLIL) in Europe. Research Perspectives on Policy and Practice,* edited by Stephan Breidbach, and Britta Viebrock, 285-296. Frankfurt am Main: Peter Lang.

MKJS (Ministerium für Kultus, Jugend und Sport Baden-Württemberg), eds. 2002. *Fremdsprachen in der Grundschule. Lehren und Lernen mit dem Konzept des kommunikativen Unterrichts.* http://www.km-bw.de/site/pbs-bw-new/get/documents/KULTUS.Dachmandant/KULTUS/kultusportal-bw/zzz_pdf/Brosch_Fremdsprachen%20-%200103 05.pdf

MKJS (Ministerium für Kultus, Jugend und Sport Baden-Württemberg), eds. 2004. *Bildungsplan 2004. Grundschule.* Ditzingen: Philipp Reclam Jun.

MKJS (Ministerium für Kultus, Jugend und Sport Baden-Württemberg), eds. 2008. *Bilingualer Unterricht deutsch-englisch an allgemein bildenden Gymnasien.* http://www.kultusportal-bw.de/site/pbs-bw/get/documents/KULTUS.Dachmandant/KULTUS/kultusportal-bw/zzz_pdf/Bilingualer-Unterricht.pdf

MKJS (Ministerium für Kultus, Jugend und Sport Baden-Württemberg), eds. 2016. *Bildungsplan 2016. Bildungsplan der Grundschule. Englisch*. http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GS_E.pdf

Möller, Verena. 2017. *Language Acquisition in CLIL and Non-CLIL Settings. Learner Corpus and Experimental Evidence on Passive Constructions*. Amsterdam: John Benjamins.

Petermann, Franz, and Sandra Winkel. 2007. *FLM 7-13. Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse.* Frankfurt am Main: Harcourt.

Rost, Detlef H. 2009. *Intelligenz. Fakten und Mythen.* Weinheim: Beltz.

Sarter, Heidemarie. 2002. "Fremdsprachen in der Grundschule." In *Fremdsprachen in der Grundschule. Lehren und Lernen mit dem Konzept des kommunikativen*

Unterrichts, edited by Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 18-25. http://www.km-bw.de/site/pbs-bw-new/get/documents/KULTUS.Dachmandant/KULTUS/kultusportal-bw/zzz_pdf/Brosch_Fremdsprachen%20-%200103 05.pdf

Teepen, Jim. 2006. "On the Relationship between Aptitude and Intelligence in Second Language Acquisition. *Columbia University Working Papers in TESOL & Applied Linguistics* 4: 1-9.

Wallaschek, Uta. 2003. "Lernstandserhebung – Leistungsförderung – Leistungsmessung im Fremdsprachenunterricht der Grundschule." In *Fremdsprachen in der Grundschule. Handreichung mit den Ergebnissen der Erprobungsschulen Englisch und Französisch,* edited by Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 20-25. http://www.km-bw.de/site/pbs-bw/get/documents/KULTUS.Dachmandant/KULTUS/kultusportal-bw/zzz_pdf/fremdsprachen_grundschule.pdf

Zydatiß, Wolfgang. 2007. *Deutsch-Englische Züge in Berlin (DEZIBEL): Eine Evaluation des bilingualen Sachfachunterrichts an Gymnasien. Kontext, Kompetenzen, Konsequenzen.* Frankfurt: Peter Lang.

Appendix

Table A1. Primary level – PSB-R 6-13 (ANOVA, Gabriel's pairwise test)

PSB-R 6-13 Scale	**IRL0** ***n*=225** ***M*** ***SD***	**IRL2** ***n*=113** ***M*** ***SD***	**IRL4** ***n*=71** ***M*** ***SD***	***F*** ***p***	***ω***	**IRL0 vs. IRL2** ***p***	**IRL0 vs. IRL4** ***p***	**IRL2 vs. IRL4** ***p***
Overall verbal	105.54 6.04	105.83 6.49	104.63 6.22	0.85 .427	.00	.966	.605	.486
Verbal comprehension	108.29 7.12	108.40 7.70	105.93 7.24	**3.17** **.043**	**.10**	.999	**.041**	.074
Word fluency	98.54 7.10	99.42 7.53	100.66 6.14	2.54 .080	.09	.622	.066	.564
Verbal relations	108.04 8.26	107.95 8.52	106.94 8.69	0.47 .624	.00	1.000	.688	.813
Overall reasoning	105.32 8.71	105.33 7.85	104.13 8.58	0.58 .559	.00	1.000	.633	.720
Reasoning (numbers)	105.45 9.37	105.91 7.80	105.58 8.71	0.10 .904	.00	.957	.999	.992
Reasoning (letters)	104.37 9.41	103.97 8.39	102.61 9.09	1.02 .361	.01	.973	.360	.681
Reasoning (figures)	106.77 8.53	107.30 9.40	106.59 9.92	0.18 .840	.00	.939	.998	.937
Spatial perception	104.73 8.46	104.61 8.05	103.92 9.15	0.25 .777	.00	.999	.846	.929
Overall concentration	99.06 8.59	100.19 8.08	98.82 8.52	0.83 .439	.00	.564	.995	.629
Numerical facility	101.70 8.99	101.76 8.98	99.99 8.54	1.11 .330	.02	1.000	.371	.461
Perceptual speed	98.08 7.90	99.92 8.36	99.10 9.44	1.90 .150	.07	.149	.728	.883
Overall cognitive capacities	104.36 6.84	104.97 7.04	103.28 6.80	1.32 .269	.04	.815	.551	.279

Table A2. Primary level – FLM 7-13 (ANOVA, Gabriel's pairwise test)

FLM 7-13 Scale	**IRL0 *n*=228 *M* *SD***	**IRL2 *n*=116 *M* *SD***	**IRL4 *n*=71 *M* *SD***	***F* *p***	***ω***	**IRL0 vs. IRL2 *p***	**IRL0 vs. IRL4 *p***	**IRL2 vs. IRL4 *p***
Orientiation tow. perf. and success	52.88 10.03	55.07 9.59	56.18 8.94	**4.00** **.019**	**.12**	.131	**.029**	.828
Perseverance and effort	52.34 10.15	53.69 10.24	55.62 9.58	3.00 .051	.10	.550	**.039**	.490
Fear of success	53.96 9.86	52.94 9.91	51.59 10.51	1.60 .202	.05	.745	.198	.746
Exam anxiety/ activating	55.63 8.81	54.90 9.66	54.79 8.71	0.38 .686	.00	.853	.858	1.000
Exam anxiety/ inhibiting	53.59 9.28	52.06 9.55	50.62 9.18	**3.07** **.048**	**.10**	.376	**.045**	.661

Table A3. Secondary level – PSB-R 6-13 (ANOVA, Gabriel's pairwise test)

PSB-R 6-13 Scale	**CLIL+ *n*=158** ***M*** ***SD***	**CLIL0 *n*=199** ***M*** ***SD***	**CLIL- *n*=54** ***M*** ***SD***	***F*** ***p***	***ω***	**CLIL+ vs. CLIL0** ***p***	**CLIL+ vs. CLIL-** ***p***	**CLIL0 vs. CLIL-** ***p***
Overall verbal	107.10 5.78	104.93 6.44	103.00 5.50	**10.92** **.00002**	**.21**	**.003**	**.00004**	.088
Verbal comprehension	109.54 6.74	107.72 7.55	104.43 7.45	**10.33** **.00004**	**.21**	.054	**.00001**	**.006**
Word fluency	100.01 6.63	98.29 7.46	100.04 6.58	**3.09** **.046**	**.10**	.065	1.000	.248
Verbal relations	109.46 8.62	107.20 7.80	105.09 8.80	**6.63** **.001**	**.16**	**0.31**	**.002**	.225
Overall reasoning	106.96 8.14	103.96 8.60	104.04 8.13	**6.18** **.002**	**.16**	**.002**	.065	1.000
Reasoning (numbers)	106.95 8.77	104.79 8.82	104.78 8.66	2.94 .054	.10	.063	.284	1.000
Reasoning (letters)	104.74 8.78	103.36 9.15	103.74 9.56	1.03 .357	.01	.393	.851	.989
Reasoning (figures)	108.96 9.17	105.21 8.93	106.78 7.91	**7.80** **.0005**	**.18**	**.0003**	.293	.542
Spatial perception	105.82 8.11	103.90 8.46	103.56 9.19	2.75 .065	.09	.096	.219	.989
Overall concentration	100.40 8.53	98.66 8.22	98.67 8.61	2.09 .125	.07	.148	.441	1.000
Numerical facility	102.60 8.84	101.10 8.62	98.89 9.68	**3.74** **.024**	**.11**	.299	**.018**	.242
Perceptual speed	99.42 8.32	97.97 7.89	99.96 9.46	1.98 .139	.07	.269	.964	.272
Overall cognitive capacities	106.19 6.75	103.51 6.94	102.28 6.05	**9.95** **.00006**	**.20**	**.001**	**.001**	.514

Table A4. Secondary level – FLM 7-13 (ANOVA, Gabriel's pairwise test)

FLM 7-13 Scale	**CLIL+ *n*=161** ***M*** ***SD***	**CLIL0 *n*=198** ***M*** ***SD***	**CLIL- *n*=57** ***M*** ***SD***	***F*** ***p***	***ω***	**CLIL+ vs. CLIL0** ***p***	**CLIL+ vs. CLIL-** ***p***	**CLIL0 vs. CLIL-** ***p***
Orientiation tow. perf. and success	55.76 9.19	53.53 9.90	51.19 10.37	**5.27** **.005**	**.14**	.088	**.005**	.258
Perseverance and effort	53.94 10.38	53.32 9.64	51.02 11.37	1.75 .175	.06	.918	.156	.309
Fear of success	56.16 9.01	51.42 10.10	52.04 10.02	**11.26** **.00002**	**.22**	**.00002**	**.014**	.960
Exam anxiety/ activating	56.34 9.05	54.37 8.67	55.30 9.97	2.12 .121	.07	.115	.882	.856
Exam anxiety/ inhibiting	51.96 8.69	53.19 9.66	52.82 10.47	0.76 .468	.00	.526	.903	.991

Tables

Table 1. PSB-R 6-13 (Horn 2003) – Reliability of scales

PSB-R 6-13: Name of scale	Cronbach's α Form A	Cronbach's α Form B
Overall verbal	.91	.92
Verbal comprehension	.93	.94
Word fluency	.80	.77
Verbal relations	.72	.77
Overall reasoning	.89	.90
Reasoning (numbers)	.63	.63
Reasoning (letters)	.58	.68
Reasoning (figures)	.65	.55
Spatial perception	.88	.89
Overall concentration	–	–
Numerical facility	–	–
Perceptual speed	–	–
Overall cognitive capacities	–	–

Table 2. FLM 7-13 (Petermann and Winkel 2007) – Reliability of scales

FLM 7-13: Name of scale	Cronbach's α
Motivation	
Orientation towards performance and success	.73
Perseverance and effort	.74
Anxiety	
Fear of success	.69
Exam anxiety/activating	.67
Exam anxiety/inhibiting	.62

Figures

Figure 1. FLM 7-13 – Aspects of motivation and anxiety (primary level)

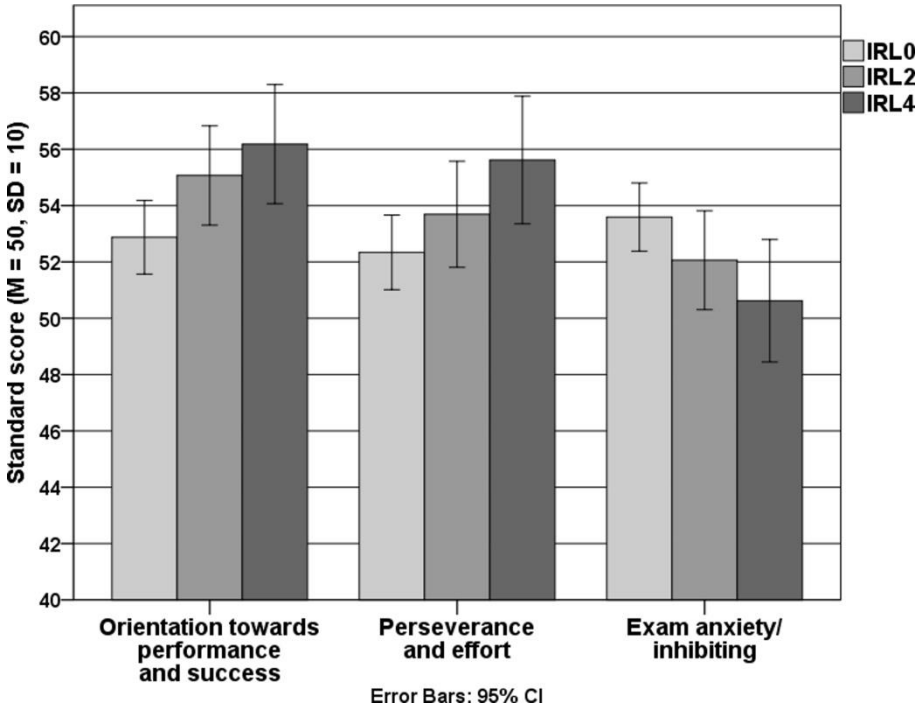

Figure 2. PSB-R 6-13 – Verbal aspects of intelligence (secondary level)

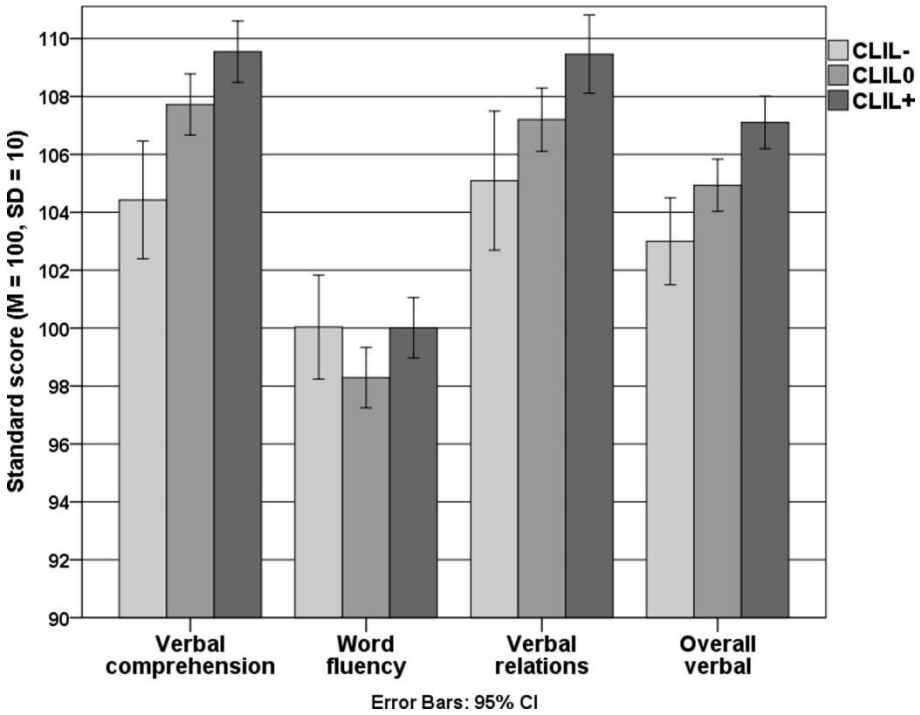

Figure captions

Figure 1. FLM 7-13 – Aspects of motivation and anxiety (primary level)

Figure 2. PSB-R 6-13 – Verbal aspects of intelligence (secondary level)