

# CLIL modules and their affective impact on students with high English anxiety and low self-efficacy

Stephanie Ohlberger, Bielefeld University  
Claas Wegner, Bielefeld University

*Many studies in recent years focused on the efficiency of bilingual education in different school settings. If a school pursues a regular implementation of CLIL modules instead of the establishment of proper CLIL streams, the student group that is affected by this teaching approach is much more diverse when compared to the preselected CLIL stream groups. In this regard, also students with high English anxiety and low self-efficacy will have to participate in the modules and it is particularly interesting to see how these student types deal with teaching units held in a foreign language. The study at hand presents results from an intervention at German schools with two biology units taught in English. Students aged 15 to 16 rated the affective effects of these modules by filling in a pre-post-follow up questionnaire. Results show that English anxiety can indeed be lowered, while students' self-efficacy increases due to this intervention. Obviously, more research has to confirm whether this is the case for other age groups and subjects as well. Nonetheless, it is a good start to recommend employing CLIL modules on a greater scale since the effects are substantial even if students do not willingly choose to take part in CLIL.*

*Keywords:* anxiety, upper secondary school, self-efficacy, CLIL, CLIL modules

## 1 Introduction

Despite the success of bilingual education as a teaching concept (see e.g. Admiraal, Westhoff, & de Bot, 2006; Dallinger & Jonkmann, 2015; Dalton-Puffer, 2007; Lasagabaster, 2008; Osterhage, 2009; Rumlich, 2016, for different contexts) and its European equivalent called Content and Language Integrated Learning (Coyle, Hood, & Marsh, 2010), the question remains whether this approach suits all students or if some reported effects do not entirely hold true for every type of student. As previous studies at least in the German context too often research students in specific CLIL streams, positive results in terms of linguistic and affective advantages seem to come easy for these students who voluntarily chose CLIL as an option. However, as CLIL advocates (Marsh, 2002, as cited in Fernández Sanjurjo, Arias Blanco & Fernández-Costales, 2018, p. 18) regularly

---

Corresponding author's email: [stephanie.ohlberger@uni-bielefeld.de](mailto:stephanie.ohlberger@uni-bielefeld.de)

ISSN: 1457-9863

Publisher: Centre for Applied Language Studies

University of Jyväskylä

© 2019: The authors

<http://apples.jyu.fi>

<http://dx.doi.org/10.17011/apples/urn.201906253409>

state all the benefits that every participant in CLIL will have, our results will help shed light on students who do not typically enrol in bilingual programmes, namely those who experience high levels of anxiety and low levels of self-efficacy. Can we also help these students to gain more self-confidence when studying biology in a foreign language? And will this effectively reduce their language anxiety?

## 2 Theoretical background

### 2.1 CLIL modules

CLIL, short for Content and Language Integrated Learning, is the European equivalent of bilingual education, which is distinguished from immersion programmes in Canada and the United States (Brown & Bradford, 2016, p. 330; Cenoz, Genesee, & Gorter, 2013, p. 247; Coyle, 2007, p. 544, 2008, p. 97; Garcia, 2009, as cited in Costa & D'Angelo, 2011, p. 2) as well as other forms of bilingual education worldwide. In Germany, most bilingual school programmes are organised as CLIL streams. These streams can be taken as profile courses, in which participating students are usually selected using the criteria of high marks and motivation for learning foreign languages (Dalton-Puffer, 2011, as cited in Devos, 2016, p. 17). Apart from additional preparatory language lessons, content subjects that are taught bilingually comprise of one more weekly lesson in comparison to regular monolingual teaching. Commonly, CLIL teaching takes place parallel with two to three subjects, and these can change from term to term. The target language, in Germany this is most often English, is used as a vehicle to teach the content normally used in the regular (monolingual) syllabus. However, since the focus is explicitly on content and not on language, it is acceptable to use the L1; after all it is called bilingual education and should not turn out to be teaching the usual content in just another language (Marsh, Marsland, & Nikula, 1999, p. 36; Ohlberger & Wegner, 2018, pp. 47–48).

CLIL modules, on the other hand, are a rather short-term implementation of teaching a content subject in a second or foreign language. A teacher might decide that he wants to teach a particular unit, often due to cultural proximity to the target language and country, in another language. This undertaking can range from just a few lessons to complete units over the period of several weeks. In this case, no additional language lessons as preparatory means are offered and students cannot willingly choose or drop the class; however, it is advisable to discuss this enterprise with the participating students beforehand. There are basically no regulations on the practice of bilingual modules, but it seems to be best if students have a good language command already and the topic is somewhat connected to the other culture. Also, one might not see such a strict usage of the foreign language as it might be pursued in regular CLIL streams, where the students are accustomed to using the foreign language in teaching contexts. According to Costa and D'Angelo (2011, p. 10), this concept represents a far more “inclusive, unifying and democratic instrument” than other realisations of CLIL. Following this approach, one circumvents the so-called *creaming effect*, i.e. CLIL is mostly shown as having favourable effects since a positively selected group of students takes part in these classes and succeeds in them (Küppers & Trautmann, 2013, p. 291; Rumlich, 2016, p. 89).

## 2.2 Foreign Language Anxiety

Anxiety is a well-known and commonly researched emotion in second language learning (MacIntyre, 2017, p. 12). Language anxiety, however, is to be differentiated from other types of anxiety and can be considered a separate, situation-specific concept (MacIntyre, 2017, p. 16). Research has found various indications of how anxiety influences language learning (MacIntyre, 2017, p. 17) and most often it has been “described as one of the strongest predictors of success or failure in [foreign language] learning” (MacIntyre, 1999, as cited in Dewaele & MacIntyre, 2014, p. 238). Even though anxiety is commonly associated with harmful effects for learning, Pekrun, Goetz, Frenzel, Barchfeld and Perry (2011) point out that it can indeed undermine intrinsic motivation, “but can induce strong extrinsic motivation to invest effort to avoid failure” (p. 38), suggesting that the overall impact on learning motivation is not entirely negative (see also Alpert & Haber, 1960; Ellis, 2008; Scovel, 1978). However, works on changing levels of anxiety in language learning are still rare (Kruk, 2018, p. 33) and thus need further examination.

Usually, language anxiety leads to lower school performance, limited test performance, and a decreasing self-evaluation of one’s second language competency (MacIntyre, 2017, p. 17). Regarding cognitive effects, heightened language anxiety correlates with self-critical thoughts, performance worries, and the fear of failure (MacIntyre, 2017, p. 17). Furthermore, cognitive performance at all stages of language production is restrictive and in general, it takes more time for anxious people to take in information and carry out tasks, thus limiting the ability to acquire and produce the target language (Gkonou, 2017, p. 136; MacIntyre, 2017, p. 17). Due to limited linguistic self-confidence, people tend to communicate less, emphasizing the influence language anxiety can have on a social level (MacIntyre, 2017, p. 17).

In the CLIL context, levels of language learning anxiety should presumably be lower, as communicating content is more of a priority than focusing on language form (Thompson & Sylvén, 2015, p. 5). This also fits Gregersen’s finding (2003) that language learning is “particularly prone to anxiety-arousal, [...as] error correction is both an indispensable part of learning and a key source of anxiety” (as cited in Dewaele & MacIntyre, 2014, p. 239), which is not as prominent in CLIL lessons. Muñoz (2002) even goes so far as to state that CLIL is “a relatively anxiety-free environment” (p. 32). This case goes hand in hand with findings from Thompson and Sylvén’s study, namely that non-CLIL-students have a higher English class performance anxiety than CLIL-students (Thompson & Sylvén, 2015, p. 14). This difference was found even before students had been separated into CLIL and non-CLIL groups – on an important note, this grouping happened as a result of students’ choice and not via screening tests (Thompson & Sylvén, 2015, p. 7). If groups are further differentiated, girls show higher levels of anxiety than boys (Thompson & Sylvén, 2015, p. 14).

Constituting CLIL as an environment for reducing language anxiety is also supported by Pihko (2007), who claims that CLIL increases students’ willingness to use English as a foreign language even during the critical phase of puberty, where the “linguistic self-confidence” of learners is often fragile (p. 137). Correspondingly, CLIL-students have more self-confidence when compared to non-CLIL students (Thompson & Sylvén, 2015, p. 14). In general, one can note that

CLIL-students possess personal skills beneficial for learning. However, Doiz, Lasagabaster and Sierra (2014, p. 216) found that the degree of anxiety is similar for CLIL and non-CLIL students during their first year of secondary education.

### 2.3 Subject-specific self-concepts and self-efficacy

Several studies have found that both self-concept and self-efficacy affect academic performance (e.g. Bandura, 1997; Bong & Skaalvik, 2003; Lent, Brown, & Gore, 1997; Multon, Brown & Lent, 1991; Skaalvik & Skaalvik, 2009; Zimmerman, Bandura & Martinez-Pons, 1992) in some way and usually both concepts can be contrasted quite clearly, however when narrowing it down to a specific domain, the boundaries become less obvious (Skaalvik & Skaalvik, 2009, p. 272). This is why both self-concept and self-efficacy will be depicted briefly in the following.

Self-concept, meaning the beliefs one has about oneself (Baumeister, 1999), is in most research considered in terms of individual domains. In the school context, therefore, a prominent question is how a subject-specific self-concept relates to a student's academic performance. Self-ascriptions on an affective level denote interest (Skaalvik, 1997; Skaalvik & Skaalvik, 2004) and have to be distinguished from evaluative-comparative statements; thus, someone can *like* a certain subject at school, but *being good or bad* in a subject is a matter of contrasting one's performance with others'. The latter is a person's self-concept *sensu stricto* (Donat, Radant, & Dalbert, 2008, p. 177). By means of social comparisons amongst other students (Jansen, Scherer & Schroeders, 2015, p. 14), self-concept decreases throughout the school career (Donat et al., 2008, p. 179). However, students also compare their performance at one time with performances in the same domain at a different point in time and comparisons among different domains also take place (Jansen et al., 2015, p. 14). It is therefore necessary to support students' autonomy, show an honest interest in them, and retain the structures needed to strengthen students' self-concept (Donat et al., 2008, p. 179), particularly, since "teachers' and parents' appraisal or stereotype endorsement" also influence students' self-concept (Jansen et al., 2008, p. 14).

As opposed to self-concept, self-efficacy beliefs are defined as their beliefs in their own capability to achieve a certain aim, thus these beliefs influence people's feelings, thoughts, motivation and behaviour (Bandura, 1994, p. 71). Scholastic self-efficacy, more specifically, can be described as a student's certainty about being able to complete new and difficult tasks solely based on their own ability (Jerusalem, 2016, p. 169). It does not rely on social comparisons (Bandura, 1986) and increases over the school years (Donat et al., 2008, p. 181), thus it is generally less stable (Jansen et al., 2015, p. 14). Academic success that is internally attributed – therefore perceptions of mastery – will enhance one's perceived self-efficacy (Pietsch, Walker, & Chapman, 2003, p. 590); in contrast, emotional and physiological states such as stress, anxiety, and fatigue can negatively influence self-efficacy (Donat et al., 2008, p. 182). High levels of self-efficacy in subjects are associated with a favourable set of self-regulating learning strategies (Donat et al., 2008, p. 183), which are often noticed in well-performing students. Furthermore, highly self-efficacious students tend to actively participate in lessons, leading to a better overall performance (Donat et al., 2008, p. 183). Thus, efficacy beliefs "contribute to academic performance over and above actual ability" (Zimmerman, 1995, p. 213), and other authors support the assumption that self-efficacy is one of the most important psychological constructs in predicting academic outcomes

(Jaekel, 2018, p. 7; Jansen et al., 2015, p. 13). However, it is difficult to separate this from other self constructs and therefore is often found in a “variety of larger, multi-dimensional frameworks such as self-regulation or motivation” (Jaekel, 2018, p. 7). While high levels of self-efficacy correlate with effective strategy use and good academic performance, low self-efficacy results in a heightened anxiety and the misjudgement of difficult tasks (Schunk & Pajares, 2010, as cited in Jaekel, 2018, p. 7). As found in Jaekel’s study by means of structural equation modelling, “self-efficacy exerted the strongest positive effect on language proficiency”; in particular, CLIL students had “higher levels of self-efficacy than their peers in regular EFL classes” (Jaekel, 2018, pp. 16–17).

As a means of contrasting self-efficacy and self-concept, one can ask different questions to address the issues: “Can I cope with this science problem?” relates to self-efficacy, while “Am I good at science?” refers to the self-concept (Pietsch et al., 2003, p. 590). A combining element of both is their positive relation to educational and career choices, even though this factor is not equally well proven with less empirical evidence in the case of self-efficacy (Jansen et al., 2015, p. 16).

There are reports stating that the academic self-concept is a “better predictor of academic performance than self-efficacy” (Choi, 2005, as cited in Bong, Cho, Seon Ahn, & Jin Kim, 2012, p. 337), but since we were more interested in evaluating how students dealt with a CLIL module as a challenge and did not assess performance, we just included items on self-efficacy instead of self-concept.

### 3 Study

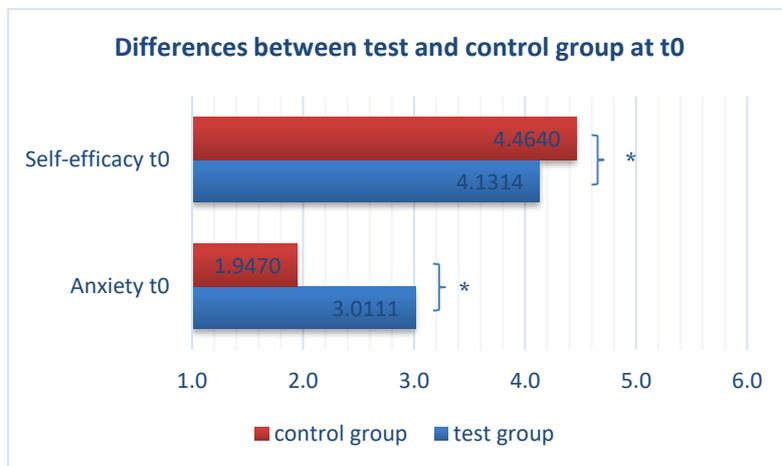
#### 3.1 Sample

Between 2017 and 2019, a total of 330 students took part in the study; 75.5% (249 cases) of data sets were complete and could be used for data analysis over four points of measurement ( $t_0$  to  $t_3$ ). The test group ( $n = 166$ , 49.4 % male, mean age = 15.66 years) consisted of students that took part in two bilingual modules, whereas the control group was made up of students in regular bilingual streams of nearby schools ( $n = 83$ , 43.4 % male (mean age = 15.82 years). 166 questionnaires from the test group and 83 from the control group were used for further analysis. Students from all participating schools had a comparable socioeconomic status. All students, both in the test and control group had received English instruction since the transition of primary and secondary school, thus amounting to having had at least five years of consecutive lessons in English as a foreign language. According to the European framework of reference for languages, this corresponds to the level B1 (MSW NRW, 2014, p. 12). In case of the control group, the students had been used to CLIL teaching in different subjects for two years.

It should be noted that there are consequences arising from this selection process (for more detail on that, see Ohlberger & Wegner, 2018, p. 80); as calculated via independent t-tests, students in the control group tended to be high achievers as they had significantly better marks in English ( $\bar{x}_{\text{control}} = 2.04^1$  vs.  $\bar{x}_{\text{test}} = 2.75$ ;  $t(198.88) = 7.15$ ,  $p = 0.000$ ,  $d = 0.946$ ) and biology ( $\bar{x}_{\text{control}} = 2.12$  vs.  $\bar{x}_{\text{test}} = 2.52$ ;  $t(173.81) = 3.502$ ,  $p = 0.001$ ,  $d = 0.463$ ). Further significant differences were observable regarding English anxiety and self-efficacy at test point  $t_0$  ( $p < 0.05$ ). While the control group showed significantly less anxiety ( $\bar{x} = 1.95$ ;  $t(219.4) = 7.722$ ,  $p = 0.000$ ,  $d = 1.038$ ) and higher levels of self-efficacy before the two

teaching units, test group students generally felt less self-efficacious ( $\bar{x} = 4.13$ ;  $t(247) = -3.033$ ,  $p = 0.003$ ,  $d = -0.408$ ) and more anxious ( $\bar{x} = 3.01$ ) (see figure 1).

For further evaluation, only the test group will be put into focus since we consider it particularly interesting what CLIL does to students who are not accustomed to being taught in a foreign language.



**Figure 1.** Differences between test and control group students before starting bilingual intervention with test group students ( $t_0$ ) regarding self-efficacy and anxiety. (\*) indicates significant differences ( $p < 0.05$ ).

### 3.2 Test instrument

Since a longitudinal evaluation of the implications of bilingual modules was sought, questionnaires had to be used in a pre-post-follow up design. Further, as many students took part, a qualitative way of researching was not an option and the subgroups analysed for this partial study would have been even smaller if no quantitative approach had been chosen. Very similar study designs that have proven successful in parts also inspired this study, e.g. Rumlich (2016), Rodenhauer and Preisfeld (2015), Dallinger, Jonkmann, Hollm, & Fiege (2016) and Verriere (2014).

In the study at hand, an identical questionnaire was used to survey students' opinions before ( $t_0$ ), right after the first ( $t_1$ ) and after the second ( $t_2$ ) bilingual module. Apart from sociodemographic data, questions covered biology and English motivation and interest, English disinterest and anxiety, and self-efficacy beliefs concerning the bilingual intervention. Further, students stated their favourite subject(s) choosing between biology, English, both, or neither. Subject preference was correlated with the index variable 'affinity', which is calculated from the means of intrinsic motivation and interest for each of the subjects, providing a consolidated perspective on subject preference.

Since anxiety and self-efficacy are of major importance, a few sample questions are provided below (see table 1); each were answered on a six-point Likert scale.

**Table 1.** Sample questions for anxiety and self-efficacy (Cronbach's alpha for complete scale).

<b>Anxiety (<math>\alpha = .853</math>, <math>n = 246</math>, 5 items)</b>
<i>I feel more tense and nervous in my English class than in my other classes. (based on Doiz et al., 2014)</i>
<i>I always have the feeling that my classmates speak English better than I do / speak better English than me. (based on Doiz et al., 2014)</i>
<i>I am very nervous during English exams. (based on Pekrun et al., 2011)</i>
<b>Self-efficacy (<math>\alpha = .835</math>, <math>n = 219</math>, 7 items)</b>
<i>Even if I am sick for a longer time, I will still be able to perform well. (based on Jerusalem &amp; Satow, 1999)</i>
<i>I am able to solve difficult tasks in school when I make an effort. (based on Jerusalem &amp; Satow, 1999)</i>
<i>Even if my teacher doubts my abilities, I am sure I can still perform well. (based on Jerusalem &amp; Satow, 1999)</i>

### 3.3 Data analysis

For our data analysis, the statistics software IBM SPSS 25.0 was used. To compare the test and control groups, or pre-post-follow up data from one of the groups, independent and dependent t-tests were applied. Since group sizes did not always measure up to numbers that are robust enough for t-test statistics, the non-parametric equivalents of Wilcoxon and Mann Whitney U were used accordingly.

The significance level was determined as  $p \leq 0.05$  as calculations included subgroups with sample sizes lower than 100 pupils (for previous discussions on the meaningfulness of p values, see Biau, Jolles, & Porcher, 2010; Dahiru, 2008; Palesch, 2014). Effect sizes were considered *high* if values were above 0.8 for Cohen's *d*, *medium* for values between 0.5 to 0.8 and *small* if values were lower than 0.5 (Field, 2013, pp. 80, 82).

For most comparisons, students were categorised as having low or high levels of either anxiety or self-efficacy. The distinction was made based on the construct mean and extended to  $\pm 1$  point as a medium zone, where students did not show any particularly unusual values. Since the general mean of anxiety was 3, low anxiety was categorised as  $\leq 2$  points and high anxiety as  $\geq 4$  points. For self-efficacy, the same method applies; the mean for all students in the test group was 4, so students with low self-efficacy were those with construct means of  $\leq 3$  points and with high self-efficacy with construct means of  $\geq 5$  points.

## 4 Hypotheses

Based on the literature review, the following hypotheses were formulated are to be examined with the data obtained from this study:

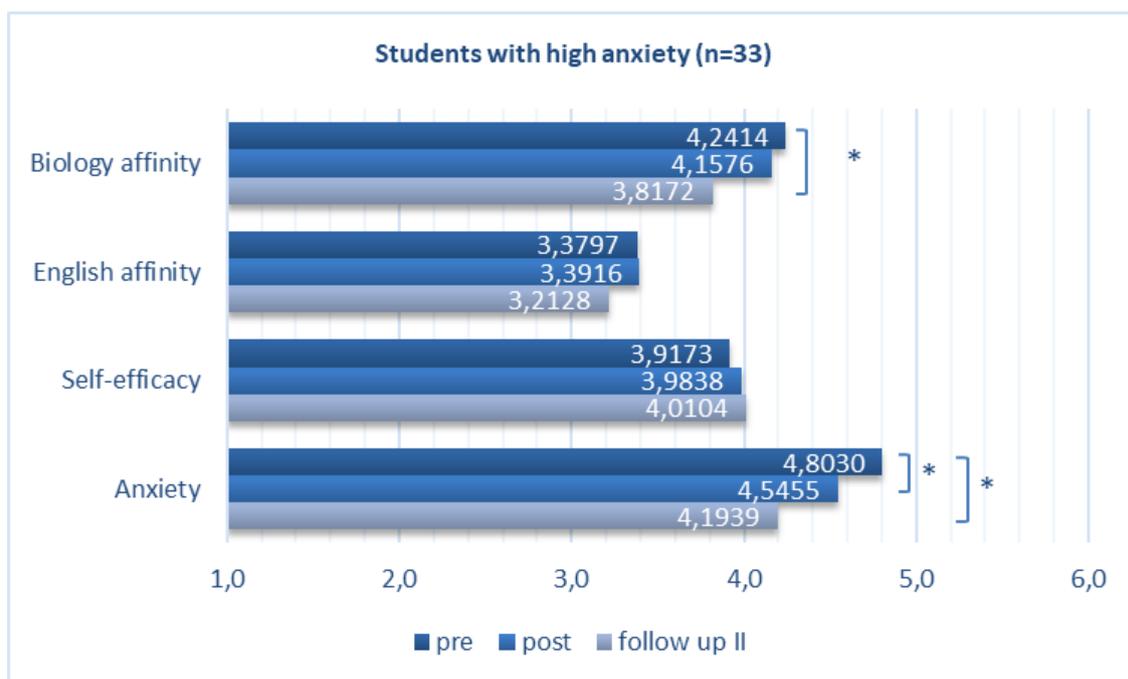
- 1) Students with high English anxiety at  $t_0$  will have lower levels of English anxiety at  $t_3$  after two bilingual modules.
- 2) Students with low levels of self-efficacy at  $t_0$  will have increased levels of self-efficacy at  $t_3$ .

## 5 Results

Anxiety and self-efficacy correlate negatively with each other ( $r = -.269$ ) in the test group, suggesting that students with high English anxiety have low levels of self-

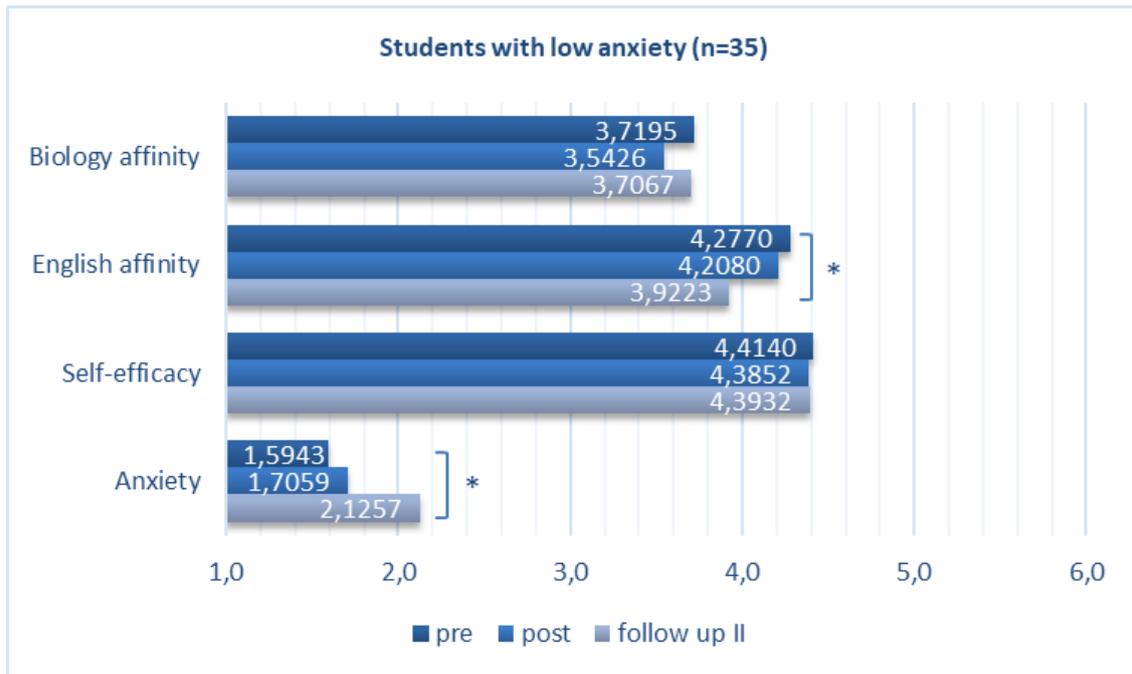
efficacy. In some cases, students will therefore be in the categories “high self-efficacy” and “low anxiety” as well as “low self-efficacy” and “high anxiety” at the same time.

Students with high anxiety ( $n = 33$ , 4.80 points at  $t_0$ , see figure 2) were significantly less anxious after one module ( $\bar{x} = 4.55$ ;  $d = 0.37$ ) and after two modules ( $\bar{x} = 4.19$ ;  $d = 0.68$ ), confirming the first hypothesis. Self-efficacy also slightly increased, although the difference was insignificant. Over the course of two modules, affinity for biology dropped by 0.4 points to 3.82 ( $d = 0.38$ ), whereas there was no significant change in English affinity.



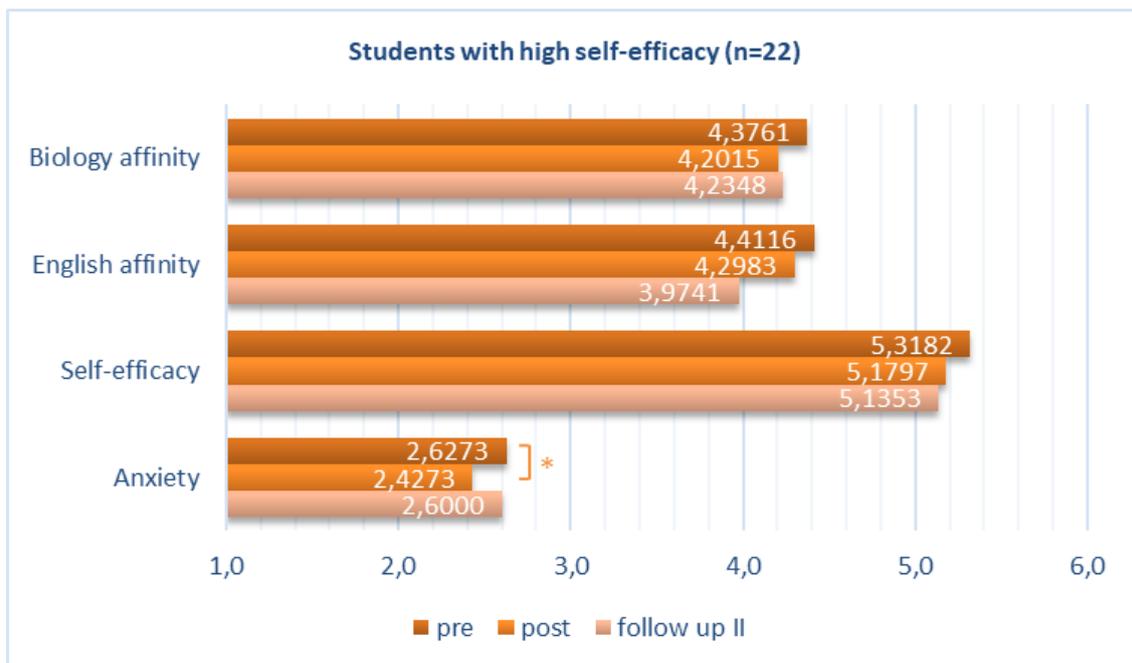
**Figure 2.** Mean differences of students with high anxiety with comparisons between  $t_0 - t_1$  and  $t_0 - t_3$ . (\*) indicates significant differences between either  $t_0$  and  $t_1$  or  $t_0$  and  $t_3$  ( $p < 0.05$ ).

When looking at students with low anxiety levels ( $n = 35$ ,  $\bar{x}_{pre} = 1.59$ , see figure 3), their anxiety increased by 0.1 points after the first module and by 0.5 points after the second bilingual module, where the latter change is of significant, medium-sized effect ( $d = 0.651$ ). While biology affinity dropped by 0.2 points after the first module, it came back to its initial value of 3.71 after two modules. English affinity, however, decreased over the complete course of the intervention, with a difference of 0.3 points (small effect,  $d = 0.34$ ). In both short- and long-term comparisons, self-efficacy remained largely unaffected at a value of 4.4.



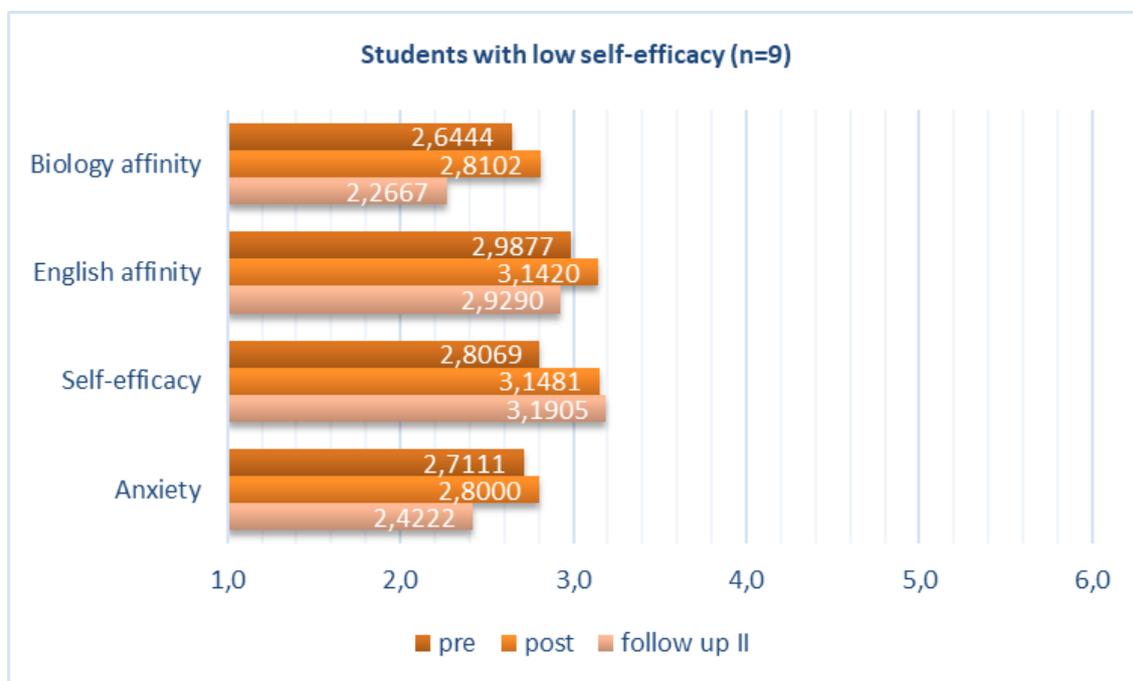
**Figure 3.** Mean differences for students with low anxiety with comparisons between  $t_0 - t_1$  and  $t_0 - t_3$ . (\*) indicates significant differences between either  $t_0$  and  $t_1$  or  $t_0$  and  $t_3$  ( $p < 0.05$ ).

Students with high self-efficacy ( $n = 22$ , see figure 4) experienced a decrease in self-efficacy to 5.14 points at  $t_3$ . Anxiety dropped significantly by 0.2 points during the first module ( $d = 1.03$ ), but rose during the second module, so that in the end the initial value was reached ( $\bar{x} = 2.60$ ). Both affinities decreased, by 0.1 and 0.5 points for biology and English, respectively.



**Figure 4.** Mean differences for students with high self-efficacy with comparisons between  $t_0 - t_1$  and  $t_0 - t_3$ . (\*) indicates significant differences between either  $t_0$  and  $t_1$  or  $t_0$  and  $t_3$  ( $p < 0.05$ ).

For students with low self-efficacy ( $n = 9$ , see figure 5), an increase in self-efficacy was observed, confirming the second hypothesis, however this difference was insignificant. Anxiety rose slightly during the first module, decreased by 0.3 points after both modules. Both subjects were liked a little more after the first module, but this did not hold true at the end of the intervention, particularly for biology (difference of 0.4 points between  $t_0$  and  $t_3$ ).



**Figure 5.** Mean differences for students with low self-efficacy with comparisons between  $t_0 - t_1$  and  $t_0 - t_3$ . (\*) indicates significant differences between either  $t_0$  and  $t_1$  or  $t_0$  and  $t_3$  ( $p < 0.05$ ).

## 6 Discussion

The study set out to research the impact of CLIL on students who would not typically choose CLIL in the first place; as for the results, many assumptions were confirmed and are likely explained by the preconditions of the groups. For **students with high levels of anxiety**, the modules helped significantly reduce their anxiety, as a focus on the English language was not present in the bilingual biology modules (Gregersen, 2003, as cited in Dewaele & MacIntyre, 2014, p. 239; Pihko, 2007, p. 137). Furthermore, their feelings of achievement induced an increase in self-efficacy (Jerusalem, 2016, p. 169, Pietsch et al., 2003, p. 590). Although it seems that bilingual modules are particularly effective regarding these developments, both the affinity for English and biology decreased over the course of two modules. Since 66% of the students in this subgroup considered themselves biology-prone, this change might be explained by the fact that they did not like shifting the usual focus on biology to include English.

Surprisingly, **students with very little anxiety** experienced a significant increase in anxiety from  $t_0$  to  $t_3$ . Even though they were able to cope with the challenge (self-efficacy remains at 4.4 points), using English in an unfamiliar context might have added more pressure. Additionally, affinity for English

significantly decreased, suggesting that it was more difficult to uphold their aspirations and expectations. As these students predominantly rated English as their favourite subject, it might have bothered them that there was no linguistic focus in bilingual biology.

Although **students with high self-efficacy** experienced no change in English anxiety, self-efficacy decreased after both bilingual modules. This is similar to students with low anxiety, and could be explained by the fact that these students are used to performing at a high level and put themselves under pressure to succeed. Nonetheless, they may struggle with the bilingual modules in the beginning, which may explain the more prominent decrease in self-efficacy (and the increase of English anxiety in the low anxiety students) between  $t_0$  and  $t_1$ . As for all groups, we also see a decrease in values for both subjects.

High self-efficacy students displayed the highest English affinity amongst all four groups with 4.4 points at  $t_0$ , which can be connected to Jaekel's (2018, p. 16) finding that self-efficacy positively influences language proficiency, since there is a close relation between affinity and performance.

The results obtained from **students with low self-efficacy** look encouraging in terms of positive changes, as they displayed a continuous increase in self-efficacy to a high of 3.1 points. Anxiety was reduced with the help of both modules; they seemed to realise that they were able to cope with the challenge. Since the English language was not focused on as much as it usually is in regular language lessons, they were able to freely express themselves without consequences, thus gaining confidence (Pihko, 2007, p. 137). An increase in both subject affinities was detected after the first module, which might be due to the novelty value of the intervention. However, there is a decrease after the second module, especially in biology. It is likely that the novelty wore off and that it was seen as a nuisance, requiring more effort without noticeable benefits for the students.

A negative correlation of self-efficacy and anxiety was seen in both the test and control groups ( $r = -0.269$ ), similar to previous studies (Donat et al., 2008, p. 182; Schunk & Pajares, 2010, as cited in Jaekel, 2018, p. 7). Regular CLIL students (control group) have higher levels of self-efficacy (Jaekel, 2018, p. 17), and indeed in our control group, this value amounts to 4.46 points at  $t_0$ , while the test group students have a significantly lower value of 4.13. Our findings on language anxiety also confirm past research (Thompson & Sylvén, 2015, pp. 5, 14), as students in the control group are significantly less anxious about English than their peers who do not take part in regular CLIL streams ( $\bar{x} = 3.01$ ). A short intervention of one or two bilingual modules can successfully reduce test students' anxiety levels, supporting Muñoz (2002) that CLIL is a stress- and anxiety-free setting for students that does not have to focus on linguistics.

## 7 Conclusion

Our study explored the question whether bilingual modules can positively impact students who do not represent the typical target group of bilingual education. This question was to be answered with the help of two hypotheses, namely that by attending two modules, (1) students' English anxiety will decrease and (2) students' self-efficacy will increase. Both hypotheses could be confirmed, thus we carefully claim that CLIL modules are helpful to reduce English anxiety and increase self-efficacy, at least on the basis of this small-scale study. The reported

differences of students in CLIL streams and modules confirm the so-called *creaming effect* (Rumlich, 2016), stating that selection processes of CLIL students have led to obscure comparisons in the past. Unfortunately, CLIL streams were often taken as the bilingual test groups and compared with non-CLIL peers at the same school, so that eventually significant differences between the groups existed due to the selection procedure, but not necessarily due to the CLIL teaching. For years, this circumstance was largely ignored and solely CLIL was made responsible for the positive outcomes of CLIL students. Keeping this in mind, our study focused on subgroups of module students instead of comparing the two groups with different performance preconditions, as CLIL is supposed “to be appropriate ‘for a broad range of learners, not only those from privileged or otherwise elite backgrounds’ as compared with the past when ‘learning content though an additional language was either limited to very specific social groups, or forced upon school populations from whom the language of instruction was a foreign language’” (Cenoz et al., 2013, p. 249). In our study, unfortunately, the affinity for both subjects decreased over the course of two bilingual modules. However, it is rarely the case that motivational changes occur within these short-term interventions (Dörnyei & Ushioda, 2011, p. 198-199); the lack of any significant changes could be explained by a general interest decline over the students’ school career (Daniels, 2008, pp. 221–237; Krapp, 1998, as cited in Spörhase, 2012, p. 101).

This study’s results definitely emphasise the educational and even more affective importance of (short-time) CLIL for a very heterogeneous group of students and follow-up studies will hopefully continue to verify the efficiency of CLIL modules, also for different age groups and a further range of subjects. In particular, it has to be identified how aptitude, anxiety and performance interrelate exactly (Dewaele, Petrides & Furnham, 2008, p. 918) and what we can learn from that for adapting CLIL lessons accordingly.

## Acknowledgements

This project is part of the "Qualitätsoffensive Lehrerbildung", a joint initiative of the Federal Government and the Länder which aims to improve the quality of teacher training. The programme is funded by the Federal Ministry of Education and Research. The authors are responsible for the content of this publication.

## Endnote

<sup>1</sup> German marking system: 1 (very good) to 6 (fail)

## References

- Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students’ language proficiency in English. *Educational Research and Evaluation*, 12(1), 75–93.
- Alpert, R., & Haber, R. (1960). Anxiety in academic achievement situations. *Journal of Abnormal and Social Psychology*, 61(2), 207–215.

- Ohlberger, S., & Wegner, C. (2018). Bilingualer Sachfachunterricht in Deutschland und Europa – Darstellung des Forschungsstands [Content and Language Integrated Learning in Germany and Europe – Presentation of the state of research]. *HLZ – Herausforderung Lehrer\_innenbildung*, 1(1), 45–89.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior*, Vol. 4 (pp. 71–81). New York: Academic Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Baumeister, R. F. (1999). *The self in social psychology*. Philadelphia, PA: Psychology Press.
- Biau, D. J., Jolles, B. M., & Porcher, R. (2010). P value and the theory of hypothesis testing: an explanation for new researchers. *Clinical orthopaedics and related research*, 468(3), 885–892.
- Bong, M., Cho, C., Seon Ahn, H., & Jin Kim, H. (2012). Comparison of self-beliefs for predicting student motivation and achievement. *The Journal of Educational Research*, 105(5), 336–352.
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educational Psychology Review*, 15(1), 1–40.
- Brown, H., & Bradford, A. (2016). EMI, CLIL, & CBI: Differing approaches and goals. In P. Clements, A. Krause, & H. Brown (Eds.), *Transformation in language education* (pp. 328–334). Tokyo: JALT.
- Cenoz, J., Genesee, F., & Gorter, D. (2013). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics*, 35(3), 243–262.
- Costa, F., & D'Angelo, L. (2011). CLIL: A suit for all seasons? *Latin American Journal of Content and Language Integrated Learning*, 4(1), 1–13.
- Coyle, D. (2007). Towards a connected research agenda for CLIL pedagogies. *The International Journal of Bilingual Education and Bilingualism*, 10(5), 543–62.
- Coyle, D. (2008). CLIL – a pedagogical approach from the European perspective. In N. Van Dusen-Scholl & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 97–111). Berlin: Springer.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL – Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Dahiru, T. (2008). P – Value, a true test of statistical significance? A cautionary note. *Annals of Ibadan Postgraduate Medicine*, 6(1), 21–26.
- Dallinger, S., & Jonkmann, K. (2015). Competences and Motivation in Bilingual Instruction in History (COMBIH). Eine Längsschnittstudie zu deutsch-englischem Geschichtsunterricht an Gymnasien [Competences and Motivation in Bilingual Instruction in History (COMBIH). A longitudinal study on German-English history teaching at grammar schools]. In D. Elsner & B. Viebrock (Eds.), *Triangulation in der Fremdsprachenforschung* [Triangulation in Foreign Language Research] (pp. 171–189). Frankfurt am Main: Peter Lang.
- Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C. (2016). The effect of content and language integrated learning on students' English and history competences – killing two birds with one stone? *Learning and Instruction*, 41, 23–31.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated (CLIL) classrooms*. Amsterdam: Johan Benjamins.
- Daniels, Z. (2008). *Entwicklung schulischer Interessen im Jugendalter* [Development of scholastic interests during adolescence]. Münster: Waxmann.
- Devos, N. J. (2016). *Peer interactions in new Content and Language Integrated settings*. Heidelberg: Springer International Publishing.
- Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237–274.
- Dewaele, J.-M., Petrides, K. V., & Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: a review and empirical investigation. *Language Learning*, 58(4), 911–960.

- Doiz, A., Lasagabaster, D., & Sierra, J. M. (2014). CLIL and motivation: the effect of individual and contextual variables. *The Language Learning Journal*, 42(2), 209–224.
- Donat, M., Radant, M., & Dalbert, C. (2008). Psychologie der Schülerpersönlichkeit [Psychology of student personality]. In M. K. W. Schweer (Ed.), *Lehrer-Schüler-Interaktion* [Teacher-Student-Interaction] (pp. 127–154). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. 2<sup>nd</sup> edition. Harlow: Pearson Education.
- Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.
- Fernández Sanjurjo, J., Arias Blanco, J. M., & Fernández-Costales, A. (2018). Assessing the influence of socio-economic status on students' performance in Content and Language Integrated Learning. *System*, 73, 16–26.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. 4<sup>th</sup> edition. Los Angeles, CA: Sage.
- Gkonou, C. (2017). Towards an ecological understanding of language anxiety. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety theory, research and educational implications* (pp. 135–155). Bristol: Multilingual Matters.
- Jaekel, N. (2018). Language learning strategy use in context: the effects of self-efficacy and CLIL on language proficiency. *International Review of Applied Linguistics in Language Teaching*. Advance online publication, available at <https://doi.org/10.1515/iral-2016-0102>
- Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology*, 41, 13–24.
- Jerusalem, M. (2016). Selbstwirksamkeit [Self-efficacy]. In H. W. Bierhoff & D. Frey (Eds.), *Selbst und soziale Kognition* [Self and social cognition] (pp. 169–240). Göttingen: Hogrefe.
- Jerusalem, M., & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung [Scholastic self-efficacy beliefs]. In R. Schwarzer & M. Jerusalem (Eds.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen* [Scales for the collection of teacher and student characteristics]. Berlin. Online publication, available at <http://www.psyc.de/skalendoku.pdf>
- Kruk, M. (2018). Changes in foreign language anxiety: A classroom perspective. *International Journal of Applied Linguistics*, 28(1), 31–57.
- Küppers, A., & Trautmann, M. (2013). It's not CLIL that is a success – CLIL students are! Some critical remarks on the current CLIL boom. In S. Breidbach & B. Viebrock (Eds.), *Content and Language Integrated Learning (CLIL) in Europe. Research perspectives on policy and practice* (pp. 285–296). Frankfurt am Main: Peter Lang.
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1(1), 31–42.
- Lent, R. W., Brown, S. D., & Gore, P. A. (1997). Discriminant and predictive validity of academic self-concept, academic self-efficacy, and mathematics-specific self-efficacy. *Journal of Counseling Psychology*, 44, 307–315.
- MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety. Theory, research and educational implications* (pp. 11–30). Bristol: Multilingual Matters.
- Marsh, D., Marsland, B., & Nikula, T. (1999). CLIL: a review of current thinking. In D. Marsh & B. Marsland (Eds.), *CLIL initiatives for the millenium* (pp. 34–45). Jyväskylä: University of Jyväskylä Continuing Education Centre.
- MSW NRW (2014). *Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen: Englisch* [Curriculum for upper secondary school 'Gymnasium/Gesamtschule' in North Rhine-Westphalia: English]. Retrieved from [http://www.schulentwicklung.nrw.de/lehrplaene/upload/klp\\_SII/e/KLP\\_GOST\\_Englisch.pdf](http://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_SII/e/KLP_GOST_Englisch.pdf)
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38, 30–38.

- Muñoz, C. (2002). Relevance and potential of CLIL. In D. Marsh (Ed.), *CLIL/EMILE – The European dimension. Actions, trends and foresight potential*. Strasbourg: European Commission.
- Osterhage, S. (2009). Sachfachkönnen (scientific literacy) bilingual und monolingual unterrichteter Biologieschüler: ein Kompetenzvergleich [Content knowledge (scientific literacy) of bilingually and monolingually taught biology pupils: a comparison of competency]. In D. Caspari, W. Hallet, A. Wegner, & W. Zydatiś (Eds.), *Bilingualer Unterricht macht Schule: Beiträge aus der Praxisforschung* [Bilingual education in schools: contributions from practice research] (pp. 41–50). Frankfurt: Peter Lang.
- Palesch, Y. Y. (2014). Some common misperceptions about p-values. *Stroke; a journal of cerebral circulation*, 45(12), 244–246.
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36(1), 36–48.
- Pietsch, J., Walker, R., & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology*, 95(3), 589–603.
- Pihko, M. K. (2007). Foreign language anxiety in content and language integrated learning (CLIL) and in traditional foreign language classes. In S. Tella (Ed.), *From brawn to brain: Strong signals in foreign language education* (pp. 129–142). Helsinki: University of Helsinki.
- Rodenhauser, A., & Preisfeld, A. (2015). Bilingual (German-English) molecular biology courses in an out-of-school lab on a university campus: Cognitive and affective evaluation. *International Journal of Environmental & Science Education*, 10(1), 99–110.
- Rumlich, D. (2016). *Evaluating bilingual education in Germany*: Peter Lang.
- Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning*, 28(1), 129–142.
- Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology*, 89(1), 71–81.
- Skaalvik, E. M., & Skaalvik, S. (2004). Self-concept and self-efficacy: A test of the internal/external frame of reference model and predictions of subsequent motivation and achievement. *Psychological Reports*, 95(3), 1187–1202.
- Skaalvik, E. M. & Skaalvik, S. (2009). Self-concept and self-efficacy in mathematics: relation with mathematics motivation and achievement. *Journal of Educational Research*, 3(3), 255–278.
- Spörhase, U. (2012). *Biologie-Didaktik* [Biology Didactics]. 5<sup>th</sup> edition. Berlin: Cornelsen.
- Thompson, A. S., & Sylvén, L. K. (2015). “Does English make you nervous?” Anxiety profiles of CLIL and non-CLIL students in Sweden. *Apples – Journal of Applied Language Studies*, 9(2), 1–23.
- Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92(1), 144–151.
- Verriere, K. (2014). *Bilinguale Module im Mathematikunterricht und ihr Einfluss auf die Lernbereitschaft der Schüler/innen für das Sachfach* [Bilingual Modules in Mathematics and their Influence on students' willingness to learn subject content]. Trier: WVT Wiss. Verl.
- Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202–231). New York, NY, US: Cambridge University Press.
- Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29(3), 663–676.

Received January 14, 2019

Revision received April 9, 2019

Accepted June 7, 2019