Journal of Innovation in Psychology, Education and Didactics	Vol. 21, No. 2 2017	149 - 176

MOTIVATIONAL CHANGES DUE TO THE IMPLEMENTATION OF A BILINGUAL MODULE IN BIOLOGY

Stephanie OHLBERGER [a]*, Claas WEGNER [a]

[a] Bielefeld University, Germany

Abstract

Across Europe, language policies of recent years established different models of bilingual education, also known as "Content and Language Integrated Learning" (CLIL), in order to promote multilingualism. The study presented in this paper focuses on a bilingual project (English-German) in the subject of Biology in Germany. Commonly, two types of bilingual concepts in schools can be distinguished, namely bilingual branches and modules. Contrary to branches, bilingual modules include a whole class with all its students and do not choose participants based on marks and abilities. The modular concept comprises a unit of 10 to 14 lessons on a specific topic being taught in a foreign language, in this case English. The study's underlying question is whether motivation for English and/or Biology can be increased for certain subgroups of students due to participation in bilingual modules.

Key words*:* bilingual teaching, Biology, Content and Language Integrated Learning, English, motivation

Introduction

Bilingual education offers students an excellent opportunity to combine disciplines, but many aspects have to be taken into consideration in order to prove that this is a successful in practice. Breidbach (2002) denotes bilingual education as one of the most important changes in schools over recent years (p. 11), however, its effects are still disputed. Most studies have centred around the (lacking) acquisition of linguistic and contentual competences (for linguistic competences see e.g. Admiraal, Westhoff & de Bot, 2006; Dalton-Puffer, 2007; Lasagabaster, 2008; for contentual competences see e.g. Kondring & Ewig, 2005; Osterhage, 2009; Haagen-Schützenhöfer, Methelitsch & Hopf, 2011), just after the CLIL concept gained ground nationwide, but after

* Corresponding Author.
E-mail address: stephanie.ohlberger@uni-bielefeld.de

predominantly finding that CLIL students profit both language- and content-wise, the focus recently shifted towards affective-motivational effects of the new teaching trend (e.g. Dallinger & Jonkmann, 2015; Rumlich, 2016). The positive effects of bilingual teaching are most commonly reported by teachers' estimations, but are rarely empirically proven (Rumlich, 2015, p. 309). In order to evaluate the usefulness of bilingual modules, meaning a short-time bilingual intervention for a diverse group of students, the current study compares motivational changes in bilingual modules in a test group of students (n = 82) without any prior experience in bilingual education with a control group (n = 31) that already continuously takes part in bilingual classes in different subjects. The module examined consists of 12 lessons on the topic of 'enzymes' and was taught in six classes of year 10 students at grammar schools. Both before and directly after the module, students had to fill in a questionnaire to obtain their opinions about the subjects of Biology and English. Our results do not show any changes in students' interest and subject affinity. More specifically, the test group showed a general trend of a decrease in the motivation and interest for English and Biology, whereas the control group demonstrated a trend towards a gained interest and motivation for Biology.

Theoretical background

Bilingual modules

In Europe, the term Content and Language Integrated Learning (short: CLIL) is used for a wide variety of bilingual teaching concepts. The aim of CLIL is to teach subject content through the use of a foreign language in order to achieve indirect linguistic gains, even though the foreign language itself is only used as a means of communication and does not serve as a basis for assessment. In Germany, the language of choice often is English although it is not uncommon to encounter bilingual programmes using French and Spanish. Despite the large majority of CLIL found at grammar schools, the concept has broadened its scope to other school types in recent years, which causes the concept to lose its elitist character (Fehling, 2005, p. 27). Particularly in the case of primary and middle schools, it can be helpful to choose a 'gentle' introduction to bilingual education by using modules. Unlike bilingual branches, modules can be launched without major restructuring, and initially limit bilingual education to a few lessons or a unit. This strategy enables both teachers and students to get used to the concept and discover both advantages and disadvantages.

Diverse support in the respect of language and content is provided by anchoring phases in the mother tongue, bilingual teaching material, and visualisation aids (MSW NRW, 2011; MSW NRW, 2014; KMK, 2013). Although Biology appears difficult to teach as a bilingual subject with its challenging and complex topics and massive amount of technical terms (Kircher, 2004, as cited in Piesche et al., 2016, p. 109), it offers many possibilities of compensation: real-life objects and models can be used to facilitate illustration and understanding (Bohn & Doff, 2010, as cited in Piesche et al., 2016, p. 109). Further, the language is standardised (Crystal, 1993, as cited in Piesche et al., 2016, p. 109) and the scientific language style in the Anglo-American area is communication-oriented and geared towards comprehensibility; it resembles everyday language (Richter & Zimmermann, 2003, p. 116). Technical terms can often be related to the German terms (Preisfeld, 2016, p. 107), since they are of Greek or Latin origin (Richter & Zimmermann, 2003, p. 116).

The purpose of bilingual education is to encourage the acquisition of knowledge in both languages, to increase language competence and self-efficacy, and to stress the importance of English in everyday and work-related life (KMK, 2013). Consequently, participating in bilingual courses provides an advantage concerning general language competence. It prepares students for university studies and international jobs with a scientific approach, as English is the academic language in Biology and the remaining sciences (Richter & Zimmermann, 2003, p. 116).

Interest and motivation in CLIL

Various studies support evidence that motivation, interest, and the resulting learning objective of students decreases significantly over the course of the school years (Schiefele & Schaffner, 2015, p. 168/171; Fuchs, 2013, p. 189). The interest in science is particularly affected by negative development, which drops as students get older (Preisfeld, 2016, p. 104). By including a foreign language in regular subjects, however, interest can be increased again (Preisfeld, 2016, p. 104).

Before reporting potential changes, one must clarify that affective-motivational constructs incorporate all the individual determinants of school performance, which are motivation, emotion, self-concept, and attitude (Rumlich, 2015, p. 310). In this context, interest is commonly researched as a field of the educational sciences (Rumlich, 2015, p. 311). The following illustration attempts to present an overview of the numerous constructs and their relations with a focus on the constructs relevant for the present study. Results on motivation and interest will be described even though certain connections have not completely been explained and can overlap at some points (Schiefele & Schaffner, 2015, p. 170).

Interests display the central motivational component in educational and extracurricular settings and are closely linked with intrinsic learning motivation. However, interests should not be considered as a personality trait but are always in relation to a person's interactions with his/her environment (Krapp, 1998, p. 186).

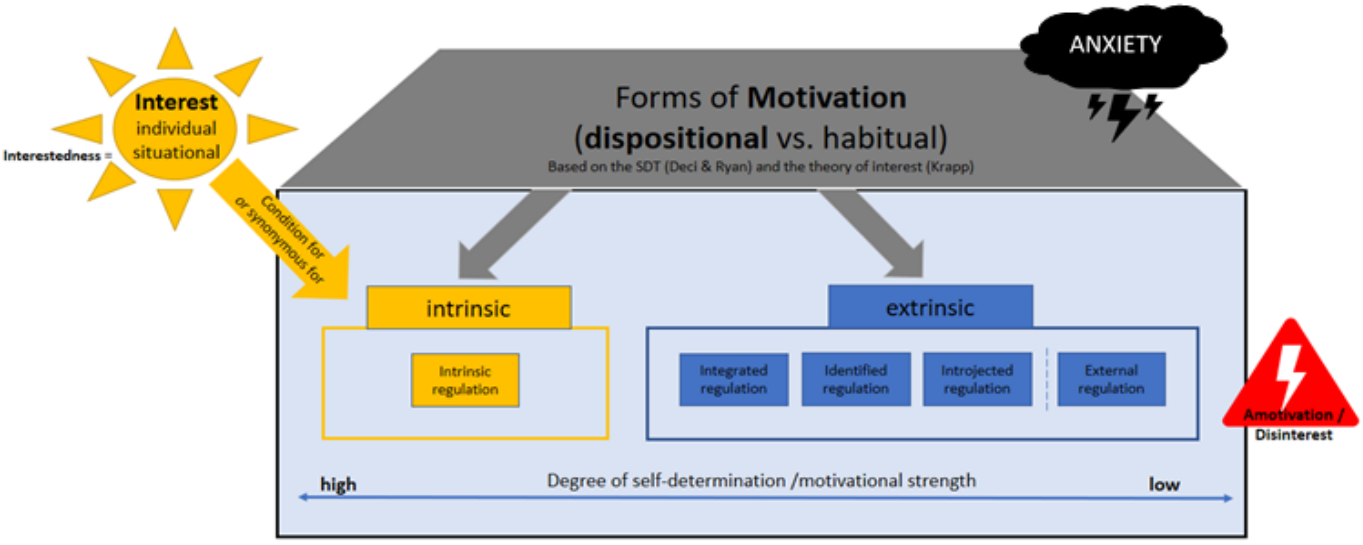

Figure 1. Overview of central constructs in motivation research

A "sparked" interest that is gained during a certain action is called situational interest or interestedness (Müller, 2006, p. 52). In school, this happens when a topic is presented in an appealing way (Krapp, 1998, p. 190). Situational interest can be seen as a starting point for the development of individual interest, which goes deeper into a person's interest (Krapp, 1998, p. 190). It denotes a relatively stable, dispositional characteristic that can be refreshed in appropriate situations (Müller, 2006, p. 52; Brandstätter et al., 2013, p. 96). It is comparatively easy to spark temporary interest or curiosity about a topic, but it is considerably more difficult to establish a long-term willingness to learn (Krapp, 1998, p. 191/ p. 198). According to Rheinberg and Vollmeyer (2012), motivation is an active change in someone's life towards a positively valued target state (as cited in Schiefele & Schaffner, 2015, p. 154). Motivation and interest represent a basis for students to engage with contents for longer periods of time (Schiefele & Schaffner, 2015, p. 154). In recent times, motivational process theories such as the Self Determination Theory (short: SDT) by Deci and Ryan (2002) are systematically linked to Krapp's theory of interest (Müller, 2006, p. 53). The SDT considers three basic human needs as a prerequisite for a high level of intrinsic motivation: autonomy, competence and relatedness (Müller, 2006, p. 53; Schiefele & Schaffner, 2015, p. 157). *Autonomy* means that a person can act self-determined without any external pressure (Schiefele & Schaffner, 2015, p. 157).

"Active" people need to have the feeling that they engage with their environment in an effective and *competent* manner (White, 1959, as cited in Schiefele & Schaffner, 2015, p. 157). Lastly, *relatedness* implies that both appreciation and trustful relationships are necessary and wield a big influence on intrinsic motivation (Schiefele & Schaffner, 2015, p. 157).

In terms of an action's purpose, one can distinguish between intrinsically and extrinsically motivated actions (Schiefele & Schaffner, 2015, p. 155). Schiefele (1996) defines extrinsic learning motivation as a state where an action is performed to induce positive consequences or avoid negative consequences (Schiefele & Schaffner, 2015, p. 155). In an intrinsically motivated action, however, the action itself is valued as a positive experience (Schiefele & Schaffner, 2015, p. 155). Intrinsically and extrinsically motivated actions are by no means mutually exclusive (Schiefele & Schaffner, 2015, p. 155). While intrinsic motivation is considered to be a consistent character trait, extrinsic motivation can be decomposed into the following categories: external, introjected, identified and integrated regulation; in this context, 'external' depicts the other-directed and 'integrated' represents the self-determined extreme of extrinsic motivation (Schiefele & Schaffner, 2015, p. 158f.). *External regulation* implies that an action is caused by rewards or punishment, while *introjected regulation* alludes to the first stage of internalisation and suggests that an action is performed in order to avoid feelings of anxiety, guilt, and having a bad conscience (Schiefele & Schaffner, 2015, p. 159; Brandstätter et al., 2013, p. 93). When an action has a personal, meaningful aim attached, this is called an *identified regulation* (Schiefele & Schaffner, 2015, p. 159). Finally, *integrated regulation* is achieved if this personal aim is not in conflict with other aims and/or activities, as the person identifies and prioritises the overall goal (Schiefele & Schaffner, 2015, p. 159).

When relating these underlying theories to CLIL, previous studies have found that the motivational effects arising from CLIL programmes are far lesser than assumed (Rumlich, 2015, p. 315). Current results should also be evaluated against the backdrop of bilingual modules, which is why a longitudinal study is recommended to track long-term motivational changes for a wide range of students (Rumlich, 2015, p. 316). Interest is also measured as it is highly related to intrinsic motivation (Schiefele & Schaffner, 2015, p. 158) and can even be equated with intrinsic learning motivation (Brandstätter et al., 2013, p. 96).

The present study uses a questionnaire based on scales from Noels et al. (2000) and Doiz et al. (2014) to measure different types of **motivation**. The authors designed their scales based on the SDT, the Socio Educational Model (Gardner, 1985) and the L2 Motivation Self System (Dörnyei, 2005) and can be applied to the foreign language context (Noels et al., 2000, p. 34;

Doiz et al., 2014, p. 215). More specifically, the performance in the foreign language (L2 performance) correlates highly with intrinsic motivation (Noels et al., 2000, p. 36). Apart from intrinsic and subcategories of extrinsic motivation, amotivation is also tested, as it can be contrasted with intrinsic and extrinsic motivation (Deci & Ryan, as cited in Noels et al., 2000, p. 40). It is labelled as **disinterest** in the questionnaire. A further negative emotion to be investigated is **anxiety** (Doiz et al., 2014; Pekrun et al., 2011), which is related to English lessons and most commonly seen in oral examinations and presentations. **Interest** in foreign language acquisition is surveyed on the basis of items by Rumlich (2016), which mostly stem from bigger tests or evaluations such as DESI, KESS and PISA (Rumlich, 2016, p. 289). Questions about **interest, intrinsic, and extrinsic motivation** in the context of Biology lessons come from Wegner (2009). **Scholastic self-efficacy** (Jerusalem & Satow, 1999) denotes the subjective certainty of being able to cope with new or difficult situations based on one's own competence (Schwarzer & Jerusalem, 2002, p. 35). Therefore, the questions of this construct are related explicitly to the bilingual modules in the questionnaire of the present study. Lessons in this context do not count among routine tasks, but present a unique 'problem' that requires effort and persistence (Schwarzer & Jerusalem, 2002, p. 35). Students with a high self-efficacy stand out due to their heightened willingness for achievement, effective time management, better performance, a more difficult level of work and a realistic assessment of their own achievements (Schwarzer & Jerusalem, 2002, p. 37/ p. 38).

Legitimation of the research project

When previous studies compared CLIL students to students taught in monolingual settings, most CLIL students were recruited from firmly established bilingual classes at schools with one of their school profiles being "language". This created a bias, as it meant that there were different starting points of the two samples: the selection process for the admission to the bilingual programme often favoured particularly motivated, interested students that were eager to learn (Dallinger & Jonkmann, 2015, p. 75/76; Rumlich, 2014, p. 75). Hence the question remains as to whether the positive effects reported by CLIL programmes can really be ascribed to the teaching concept or to the group of students (Piesche et al., 2016, p. 108). Far more interesting, however, is to examine which effects a bilingual module has on a mixed group of students. On this account, the current study uses students from bilingual branches as the control group and participants of a bilingual module as the test group. Differences between the groups are able to be measured more accurately; i.e., to what extent the motivational effects caused by the foreign language can be transferred to the bilingually taught subject content in Biology (Pirner, 2007, p. 48), or to what

extent students with an interest in Biology can get enthusiastic about English by taking part in a bilingual module (Prüfer, 2012, p. 151; Preisfeld, 2016, p. 113). Bonnet summarises these cases as compensation and reinforcement effects (Bonnet, 2012a, as cited in Rumlich, 2015, p. 312). It is thus imaginable that bilingual modules are better suited than bilingual branches to promote specific talents or preferences in a mixed learning group (Preisfeld, 2016, p. 116).

Research question and hypotheses

In order to evaluate the effects of a temporary bilingual teaching situation and to relate these changes to the mixture of students that a usual class is composed of, the following research questions will be taken into focus: Which effect does a bilingual module have on the motivational preferences of students interested in the English language?; Which effect does a bilingual module have on the motivational preferences of students interested in the subject of Biology?.

The research questions were developed in light of the fact that traditional bilingual classes explicitly choose students interested in the language; this creates a highly selective class and by no means represents a regular class. If, however, bilingual modules result in positive effects on students who are not interested in languages, this can be considered a huge success which bypasses the trouble of organising bilingual classes at schools and further expands the usefulness of the concept to a broader group of students.

Based on the theoretical background, the following hypotheses are proclaimed:

- H1: After having completed the bilingual module, the test group's motivation towards the use of English will be significantly higher than before the module.
- H2: After having completed the bilingual module, the control group's motivation towards the use of English will not result in significant differences when compared to before the module.
- H3: The test group's self-efficacy with regard to the bilingual module will be significantly higher after having completed the module.
- H4: The control group's self-efficacy will be significantly higher than the test group's self-efficacy before the bilingual module.
- H5: Students in the test group who are English-orientated will have a significantly higher affinity for Biology after having completed the bilingual module when compared to before the module/t0.

- H6: Students in the test group who are Biology-orientated will have a significantly higher affinity for English after having completed the bilingual module when compared to before the module/t0.
- H7: After three months, the students in the test group will have a significantly higher motivation towards the use of English than before the bilingual module.

Methodology

Sample

During the 2016-2017 school term, students in four different classes in year 10 for the test group (n = 82) and two different classes in year 10 for the control group (n = 31) were observed. The students in the test group had never experienced bilingual education before, whereas the control group consisted of two regular bilingual classes. All students attended grammar schools in the region of Eastern North Rhine-Westphalia. The average age of the test group was 15.66 years, and 15.81 years in the control group. Girls accounted for 51.2% of the test group and 64.5% of the control group. The average marks for the test group were 2.72 in English (German grading system: from 1 = very good; to 6 = fail) and 2.35 in Biology, whereas they were 2.26 and 2.29 in the control group for English and Biology respectively.

Intervention

The intervention involved a bilingually (German-English) taught unit about enzymes, which comprised of 10-12 lessons. Firstly, the structure and function of proteins was taught before experiments were done on the temperature, pH, and concentration dependency of enzymatic reactions. After having pointed out the importance of enzymes as biocatalysts, students had a closer look at certain enzymes in the human body, e.g. in the context of digestion. With the example of a diet pill, different types of enzyme inhibition were addressed.

The control group was taught with the same material and worked on the same topics, with the only difference being that the students were used to bilingual learning environments and thus did not regard the foreign language as a new obstacle. With the test group's as well as the control group's mother tongue being German (L1), the test group was offered all materials in both German and English (MSW NRW, 2011) in order to support the complexity of the subject-specific terminology in a suitable manner (Preisfeld, 2016, p. 103). More specifically, anchoring phases were used in both languages, L1 and L2, (MSW NRW, 2014; KMK, 2013) in order to make sure that all content could be explained and prevent any misunderstandings caused by using a foreign language. It was stressed in the beginning that students could talk in whichever

language they choose. Few used their mother tongue exclusively, as a measure of the material in the folders resulted in 15% of German worksheets left at the end of each lesson. In order to ease the first steps into the bilingual module, a document with useful phrases in English was provided. This table included phrases for describing diagrams, and terms to use in discussions and presentations. In addition, when asked to design a poster about the mode of action of Orlistat, a diet pill exemplifying enzymatic inhibition, 90% of the posters were created and presented in English with only a few that needed assistance from the teacher or fellow classmates.

Test instrument

A questionnaire was constructed to collect the students' assessments of their motivation and interest for the subjects of English and Biology; it comprised mostly of closed questions and answers could be indicated on a six-point Likert scale (1 = strongly disagree up to 6 = strongly agree). Apart from the constructs of *anxiety* and *disinterest*, high values are generally favourable. 25 students in the same age group used a test pilot that had the following set-up: questions about Biology lessons were subdivided into items for the constructs' interest, intrinsic, and extrinsic motivation (Wegner, 2009), while questions about English lessons included items on the constructs' interest (Rumlich, 2016), intrinsic and extrinsic motivation (Noels et al., 2000; Doiz et al., 2014), disinterest (Noels et al., 2000) and anxiety (Doiz et al., 2014; Pekrun et al., 2011). Further measures asked to determine the scholastic self-efficacy (Jerusalem & Satow, 1999) with regard to the bilingual module, as well as a few open questions to provide an overall evaluation of the bilingual module (Abendroth-Timmer, 2007). Internal consistency for each of the constructs was checked with a Cronbach's alpha reliability analysis and yielded the following values (see Table 1).

Table 1. Reliability of the individual constructs, calculated in the basis of N = 113 at time of testing t_0.

Construct (Number of items)	Cronbach's α
Interest Biology (5)	0.846
Intrinsic Motivation Biology (3)	0.872
Extrinsic Motivation Biology (4)	0.734
Interest English (7)	0.900
Intrinsic Motivation English (10)	0.921
Extrinsic Motivation English (3)	0.745
Anxiety English (5)	0.864
Disinterest English (4)	0.911
Self-efficacy bilingual module (7)	0.830

Additionally, sociodemographic data was collected in order to categorise students based on marks and language background, which could be used for further analyses. Students were grouped into the following categories: English-oriented, Biology-oriented, oriented towards both subjects, or no orientation. This was established based on the mean values of the constructs' interest, intrinsic, and extrinsic motivation, for both Biology and English, giving a new variable called 'affinity'. If these means were above 4 for both subjects, students were classified as being interested in *both subjects*. For means under 3 in both subjects, students were described as having *no orientation*. A mean value over 4 for one of the subjects divided the students into either *English-oriented* or *Biology-oriented*. Please note that orientation, interest focus and affinity are used interchangeably in this paper.

The questionnaire was employed at t_0 as a pre-test (before starting with the bilingual module) and at t_1 as a post-test (directly after the bilingual module). A follow-up test (t_2) roughly three months after completing the module enabled a realistic long-term assessment of the module's effects. However, this kind of longitudinal study is still uncommon in bilingual research contexts (Piesche et al., 2016, p. 109).

Statistical analysis

Various statistical methods were employed to interpret the test results in the statistics software SPSS 24.0. To compare the test and control group or pre-post-data from one of the groups, independent and dependent t-tests were respectively used. The significance level was determined as $p \leq 0.05$ as calculations included subgroups with sample sizes lower than 100 students (for recent discussions on the meaningfulness of p values, see Dahiru, 2008; Biau et al., 2010; Palesch, 2014). Differences between pre-, post- and the follow-up test were examined using a general linear model with repeated measures. In order to calculate correlations between learner variables, either a Pearson's correlation coefficient or a Chi-square test for independence was used depending on variable scale. Effect sizes were considered *high* if values were above 0.8 for Cohen's *d*, 0.14 for Eta², 0.5 for Cramer V, and 0.5 for correlations based on Pearson. *Medium* effects are represented by values between 0.5 to 0.8 for Cohen's *d*, 0.06 to 0.14 for Eta², 0.3 to 0.5 for Cramer V, and 0.3 to 0.5 for correlations based on Pearson. Only *small* effects are achieved if values are lower than 0.5 for Cohen's *d*, 0.06 for Eta², 0.1 for Cramer V, and 0.1 for correlations based on Pearson. (Field, 2013, p. 80/82).

Results

Distribution of orientations within the test and control groups

The questionnaire divided students into several groups based on their opinions towards whether they favoured Biology, English, both, or none of the subjects. Further examinations resulted in another variable named "affinity", which was calculated based on the means of interest, intrinsic, and extrinsic motivation for each of the subjects. In order to verify the categorisation based on the mean values, a Chi² test was run for the students' self-estimation and the calculated affinity at t_0. As can be seen in Table 2, there is a strong correlation of these variables in the test group but not in the control group, with a medium effect size (0.353) in the test group.

Table 2. Correlations of self-estimated and calculated affinity for test and control group

Chi-square test for independence			
	Value	df	Significance
Test group: Chi-square based on Pearson	26.896	9	0.001
Control group: Chi-square based on Pearson	5.318	2	0.070
Symmetrical measures			
		Value	Significance
Test group (n=72)	Cramer-V	0.353	0.001
Control group (n=13)	Cramer-V	0.640	0.070
Sample size		72	

The pre-test affinity foci in the test group resulted in a dominance for 'no affinity' (n = 27; 32.93%), with 23 students (28.00%) following at 'double affinity' and 22 students (26.80%) at 'English affinity'. The affinity for Biology was only calculated for ten students (12.20%) (see Table 3). Furthermore, it was observed that between pre- and post-test four students shifted from an English-affinity to the double-focus group, and as a contrasting development, six of formerly 23 double-focused students shifted to a pure English-focus. Over the period of the module, the number of Biology-oriented students decreased; from ten students, only five of them stay with the Biology-focus, and four did not show any particular affinity[2].

[2] More details can be read from table 3, which should be interpreted as follows: in the left column, the affinity foci at t_0 are listed. If one follows one affinity focus row-wise, one can identify how the affinity changes from t_0 to t_1. *Example for the test group*: a total of ten students is Biology-affinity at t_0, but after the bilingual module (t_1) out of the ten students only five remain Biology-affinity, one of the ten students shows a double-affinity and four of them have no affinity.

For 60.00% of the control group, an English-affinity was determined, with only 23.33% being interested in both English and Biology and 16.66% not showing any affinity focus (see Table 3). No pupil was predominantly interested in Biology at t_0. The number of students with a double focus remained stable over the period of the enzyme unit; five students that previously belonged to the group of English-affinity even shifted to the double focus at t_1.

Table 3. Cross tab on the development of affinity foci from pre (t_0; to be read in rows) to post (t_1; to be read in columns) within test (n=82) and control (n=30) group

Test group		**Affinity focus post**				**Total pre**
		Biology	English	Biology and English	none	
Affinity focus pre	Biology	5	0	1	4	**10**
	English	1	15	4	2	**22**
	Biology and English	3	6	12	2	**23**
	none	3	1	1	22	**27**
Total post		**12**	**22**	**18**	**30**	**82**
Control group		**Affinity focus post**				**Total pre**
		Biology	English	Biology and English	none	
Affinity focus pre	English	0	10	5	3	**18**
	Biology and English	0	0	7	0	**7**
	none	1	1	0	3	**5**
Total post		**1**	**11**	**12**	**6**	**30**

With regard to hypotheses H5 and H6, the affinity for Biology and English was measured beforehand and compared between the points of measurement of t_0 and t_1. For the English-focused students (n=22) in the test group, a decrease from 3.50 (t_0) to 3.33 (t_1) points in their Biology-affinity could be detected. This, however, does not mean a significant difference, which is why H5 has to be discarded (see Table 4).

Table 4. t-test for the comparison of the Biology-affinity before and after the bilingual module in case of the English-oriented students of the test group

Construct	**N**	**Mean**	**SD**	**T**	**df**	**Sig.**	**Effect size *d***
Affinity Biology Pre	22	3.5028	0.37046	1.217	21	0.237	0.25
Affinity Biology Post	22	3.3306	0.78329				

For the Biology-focused students (n=10) in the test group, a slight increase from 3.05 to 3.12 points in their English-affinity was noticed, which still does not make for a significant difference. Therefore, also H6 has to be discarded (see Table 5).

Table 5. t-test for the comparison of the English-affinity before and after the bilingual module in case of the Biology-oriented students of the test group

Construct	N	Mean	SD	T	df	Sig.	Effect size *d*
Affinity English Pre	10	3.0503	0.89567	-0.297	9	0.773	-0.074
Affinity English Post	10	3.1170	0.90373				

Correlations between marks and subject affinity as well as mark differences between test and control group

Correlation calculations in the test group between the Biology affinity (mean of interest, intrinsic, and extrinsic motivation) and Biology marks resulted in a strong negative correlation (r = -0.535), which implies that a high affinity for Biology correlates with a good mark in the subject (adequate to a small value in the German grading system). Similar results were observed for English; the affinity for English and respective marks have a correlation coefficient of r = -0.508 (see Table 6).

Table 6. Correlations between English marks and English-affinity as well as Biology marks and Biology-affinity at t_0 in the test group.

		English mark	Affinity English Pre
English mark	Correlation based on Pearson	1	-.508**
	Significance		0.000
	N	81	81
Affinity English Pre	Correlation based on Pearson	-.508**	1
	Significance	0.000	
	N	81	82
		Biology mark	Affinity Biology Pre
Biology mark	Correlation based on Pearson	1	-.535**
	Significance		0.000
	N	81	81
Affinity Biology Pre	Correlation based on Pearson	-.535**	1
	Significance	0.000	
	N	81	82

**. The correlation is significant at a level of 0.01 (two tailed).

Regarding correlations between Biology marks and affinity in the control group, a stronger negative correlation with r = -0.700 was detected. In turn, this entails that a higher Biology mark correlates with a high motivation and interest for Biology (see Table 7). However, there is no correlation between English marks and affinity.

Table 7. Correlations between English marks and English-affinity as well as Biology marks and Biology-affinity at t_0 in the control group.

		English mark	Affinity English Pre
English mark	Correlation based on Pearson	1	-0.196
	Significance		0.292
	N	31	31
Affinity English Pre	Correlation based on Pearson	-0.196	1
	Significance	0.292	
	N	31	31
		Biology mark	Affinity Biology Pre
Biology mark	Correlation based on Pearson	1	-.700**
	Significance		0.000
	N	31	31
Affinity Biology Pre	Correlation based on Pearson	-.700**	1
	Significance	0.000	
	N	31	31

**. The correlation is significant at a level of 0.01 (two tailed).

Previously mentioned differences in marks for English and Biology show a significantly better performance in English on behalf of the control group with a medium effect size (d = 0.543) (see Table 8).

Table 8. T-test for the comparison of the test and control group concerning the English mark.

Group	N	Mean	SD	T	df	Sig.	Effect size *d*
Test group	81	2.72	0.884	2.569	110	0.012	0.543
Control group	31	2.26	0.729				

Comparison between the test and control group regarding construct means and pre, post, and follow-up testings

For the whole test group, the means decreased for the constructs *interest Biology*, *intrinsic* and *extrinsic motivation Biology*, *interest English*, *intrinsic* and *extrinsic motivation English*, *anxiety English* and *self-efficacy bilingual module*. All changes were insignificant (Table 14 in the appendix) leading hypothesis H3 to be falsified (see Table 9).

Table 9. t-test for the comparison of self-efficacy before and after the bilingual module within the test group

Construct	N	Mean	SD	T	df	Sig.	Effect size *d*
Self-efficacy bilingual module Pre	82	4.1738	0.86334	0.403	81	0.688	0.029
Self-efficacy bilingual module Post	82	4.1481	0.92742				

Disinterest in English slightly increased from t0 to t1. Both of the affinities for Biology and English decreased by approximately 0.1 points. All differences in both directions amount to a maximum of 0.2 points (Appendix A), which does not make any changes significant. Hence, hypothesis H1 has to be falsified as well.

Long-term motivational changes in the test group were measured using a general linear model (GLM). It became clear that interest Biology (small effect, η^2=0.058), intrinsic motivation Biology (medium effect, η^2=0.064), interest English (small effect, η^2=0.049) changed significantly from t0 to t2 (see Table 10 and Appendix B). Differences between the times of measurements are depicted graphically in figure 3. It is striking that all interest and motivation constructs for both subjects decrease in value from the pre-test to the follow-up test, apart from extrinsic motivation English (see Appendix C). While anxiety for English decreased by 0.14 points, disinterest increased slightly. The self-efficacy value, regarding bilingual modules, increased by 0.1 points between the pre-test and the follow-up test. Hypothesis H7 must be rejected as there is no significant increase for any motivational form of English from t0 to t2.

Table 10: GLM for the inspection of differences for all constructs over time within the test group. Based on data of students who were present at all three testings, thus n = 73

Construct	Sum of squares (Type III)	df	Mean squares	F	Sig.	Partial Eta²
Interest Biology	2.841	2	1.421	4.412	0.014	0.058
Intrinsic Motivation Biology	5.118	2	2.559	4.909	0.009	0.064
Extrinsic Motivation Biology	0.760	2	0.380	1.242	0.292	0.017
Interest English	1.760	2	0.880	3.688	0.027	0.049
Intrinsic Motivation English	1.109	2	0.555	2.502	0.085	0.034
Extrinsic Motivation English	0.251	2	0.125	0.260	0.772	0.004
Anxiety English	1.659	2	0.830	2.957	0.055	0.039
Disinterest English	0.816	2	0.408	2.803	0.064	0.037
Self-efficacy bilingual module	0.310	2	0.155	0.627	0.536	0.009

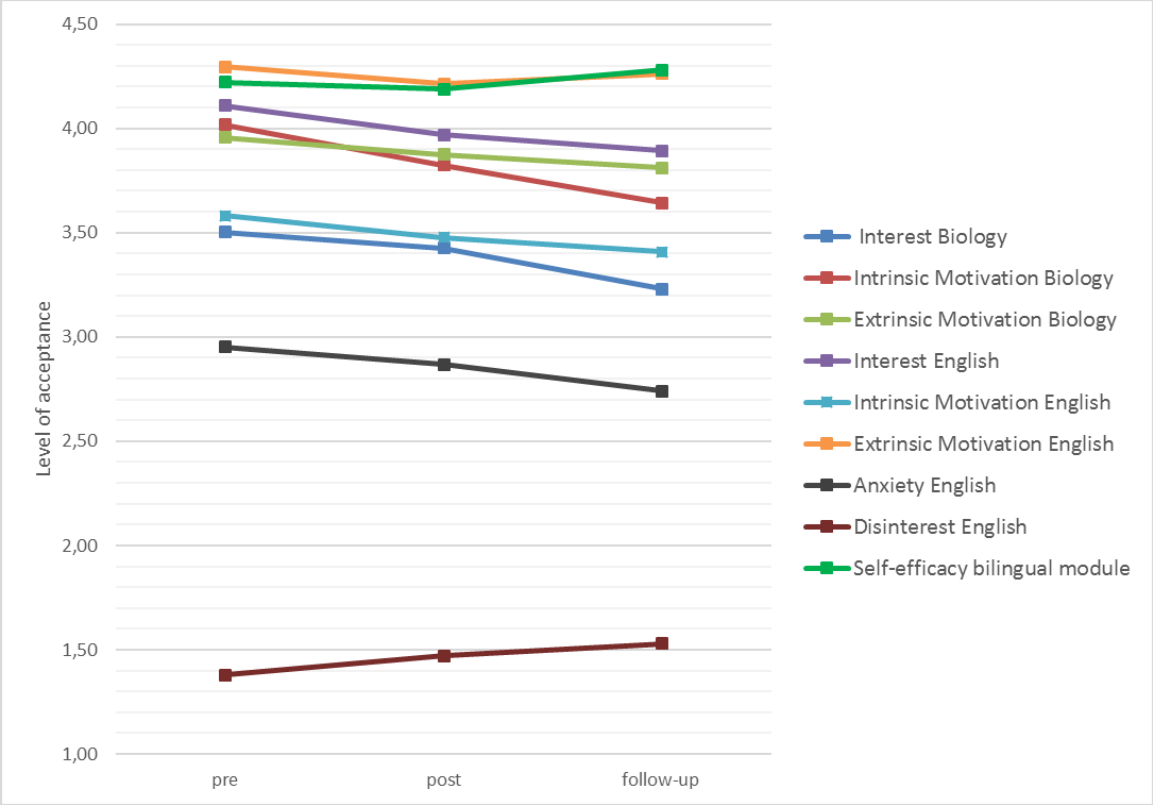

Figure 2. Illustration of construct means within the test group (n = 73) over the measurements of pre (t_0), post (t_1) and follow-up (t_2). In order to improve readability, the y-axis was cut at 4.5 (with the actual scale reaching from 1 to 6)

When comparing t_0 and t_1, the mean values for *extrinsic motivation Biology*, *intrinsic motivation English,* and *interest English* decreased slightly within the control group. For *interest Biology*, *intrinsic motivation Biology*, *anxiety* and *disinterest English*, *extrinsic motivation English* and *self-efficacy*, slight increases were recorded (see Figure 2). Just as could be observed in the test group, the affinity for English and Biology also decrease in the control group. As there are no significant changes regarding English motivation within the control group, hypothesis H2 can be verified.

Looking at the overall range of answers given at t_0, it stands out that all possibilities of the answer scale (strongly disagree (1) to strongly agree (6)) were used for the construct *anxiety for English*, that, however, only a quarter of answers is given in the range between 3.5 and 6. Answers given with in the *disinterest English* construct behave in a similar manner; even though the complete range of the answer scale was used, 75% of them are in between 1 and 1.25. Compared with *extrinsic motivation English*, a quarter of all answers were in the range of 5.33 and 6.

The means and standard deviation of the constructs *anxiety* and *disinterest English* are higher in the test group than in the control group (see Table 11); for anxiety, 25% of the test group answered between 4 and 6, while the last quartile ranges from 2 to 6 in the control group.

Table 11. Overview of means and standard deviations for *anxiety* and *disinterest English* in test and control group at t_0 and t_1

Construct	Group	N	Mean	SD
Anxiety English Pre	Test	82	2.8841	1.32026
	Control	31	1.7871	0.75354
Anxiety English Post	Test	82	2.7829	1.22291
	Control	31	1.8409	0.74131
Disinterest English Pre	Test	82	1.4756	0.99893
	Control	31	1.1048	0.34018
Disinterest English Post	Test	82	1.5528	0.99919
	Control	31	1.1694	0.42518

A further comparison of results shows that 76.83% (63 of 82) of students in the test group and 87.10% (27 of 31) of students in the control group evaluated themselves with the same or even higher self-efficiency values after having completed the module. Before the start of the bilingual module, no significant differences are found between the test and control group, which leads to the rejection of hypothesis H4 (see Table 12).

Table 12. t-Test for the comparison of *self-efficacy* between test and control group before the bilingual module

Group	N	Mean	SD	T	df	Sig.	Effect size d
Test	82	4.1738	0.86334	-1.124	111	0.264	-0.237
Control	31	4.3698	0.72166				

When just looking at the two subgroups of Biology- and English-oriented students, their affinity towards the opposite subject was investigated (Figure 3). It became apparent that for the English-focused students (only 20 who were present for all three testings), a decrease regarding their Biology-affinity from 3.56 (t_0) to 3.43 (t_1) to 3.40 (t_2) points occurred, which makes no significant difference. For the only 9 Biology-focused students a slight increase from 3.18 to 3.20 to 3.30 points could be detected for their English-affinity, which is still not statistically significant.

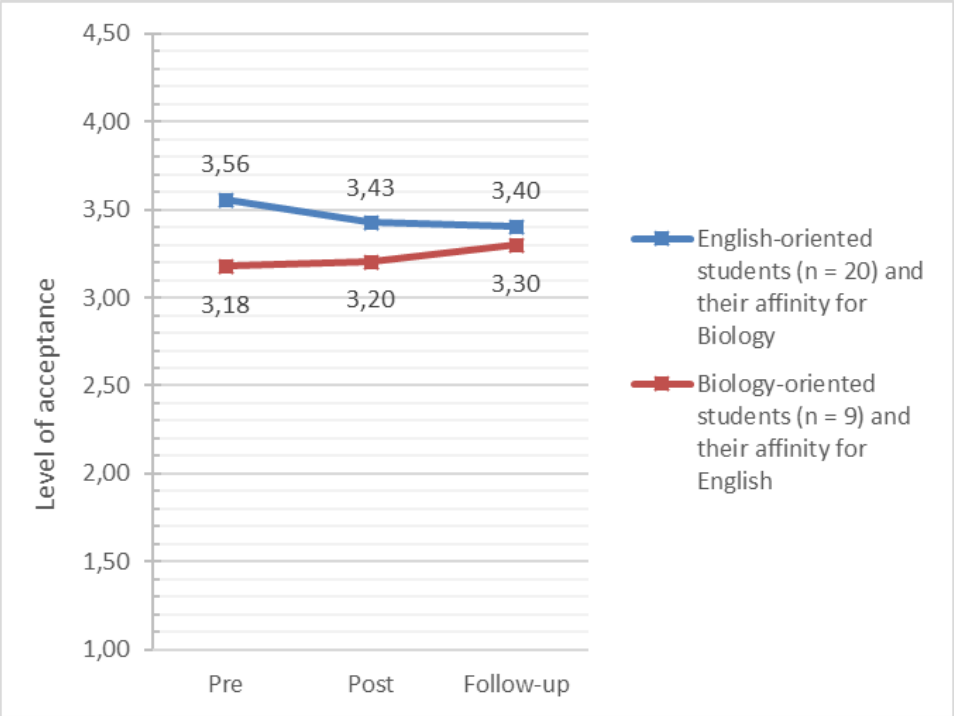

Figure 3. Development of the affinity for Biology in the case of the rather English-oriented students (n = 20) and of the affinity for English in case of the Biology-oriented students (n = 9) within the test group over the course of the bilingual module

Discussion

The results allow for a more in-depth characterisation of the two groups. The control group (average mark of 2.26) displays a significantly better performance in English when compared to the test group (average mark of 2.72), as it is often found in literature (Dallinger & Jonkmann, 2015, p. 75/76; Rumlich, 2014, p. 75; see Table 8). Indeed, it seems to be the case that admission criteria for students interested in bilingual classes presupposes a selection of highly motivated students and thus applies this concept to a selective group (Dallinger & Jonkmann, 2015, p. 75/76; Rumlich, 2014, p. 75). Using the test of correlation between self-evaluated affinity and calculated orientation by means of a Chi²-test confirms that calculating a mean value based on interest, intrinsic, and extrinsic motivation for each subject results in a high consistency, therefore legitimising the determination of the affinity-foci.

It is intriguing to observe that none of the students in the control group displayed a pure Biology-affinity, but that 60% prefer English (see Table 3). As a result, it may be interpreted that the reason for choosing a bilingual class could culminate from a student's heightened interest in English, whereas the content subject does not seem to play as big of a role. No particular affinity

for Biology may be attributed to the schedule, as bilingual subjects usually change periodically and students may not voluntarily choose Biology if other options are available. The distribution of subject foci in the test group exhibit that the year 10-class is a diverse group; fortunately, the majority display a double focus (n=23, 28%) and a similar number of students (n=22, 26.8%) feel rather attached to English (see Table 3). Only 12.20% reveal an affinity for Biology, which confirms the negative development of scientific interests with increasing age or academic year (Krapp, 1998, p. 187).

In general, the most "ideal" case did not occur between the pre- and post-test. Students that were not previously interested in any of the subjects did not develop an interest in either subject with the help of a bilingual module (see Table 3). However, the bilingual module may still be deemed as important, as four students of the test group with a previous English-focus could be considered having an affinity for both subjects after having completed the bilingual module. The treatment could have assisted the students by leading to the feeling that including the foreign language made Biology more endurable and likable. The shift from a double-focus to a pure English-affinity in six students can reinforce the idea that even temporary bilingual units give rise to preference development. However, one has to keep in mind that by introducing a bilingual module in Biology, Biology-oriented students' enjoyment of the subject can be spoiled, as only five of ten students remained with an interest in Biology after the post-test (t_1), while four did not display any particular preference anymore.

Developments in the control group presented themselves differently. The number of students with a double-focus stayed the same and even "won over" five more students by shifts from the English-affinity group. It can be assumed that the topic was decisive for the shift towards a two-fold affinity. Two of five students which did not prefer any of the subjects prior to the enzyme unit were oriented towards English and Biology after the unit, so it seems as if bilingual education can still affect profiling after a certain time.

Correlation calculations of the English marks and affinity point out that these two factors are not interrelated in the control group. This leads to the assumption that students are highly motivated and interested in the language without needing the recognition provided by marks. In addition, poorer marks will not dissuade them from mastering English as effectively as possible. On the other hand, a high correlation between the Biology marks and affinity and the English marks and affinity could be respectively noticed in the test group, which suggests that marks and affinity influence each other positively. The previously stated hypotheses will be further discussed individually; they are again listed in Table 13.

Table 13. Overview of stated hypotheses. A ✘ means the falsification of a hypothesis, while ✓ indicates a verification.

H1	After having completed the bilingual module, the test group's motivation towards the use of English will be significantly higher than before the module.	✘
H2	After having completed the bilingual module, the control group's motivation towards the use of English will not result in significant differences when compared to before the module.	✓
H3	The test group's self-efficacy with regard to the bilingual module will be significantly higher after having completed the module.	✘
H4	The control group's self-efficacy will be significantly higher than the test group's self-efficacy before the bilingual module.	✘
H5	Students in the test group who are English-oriented will have a significantly higher affinity for Biology after having completed the bilingual module when compared to before the module/t0.	✘
H6	Students in the test group who are Biology-oriented will have a significantly higher affinity for English after having completed the bilingual module when compared to before the module/t0.	✘
H7	After three months, the students in the test group will have a significantly higher motivation towards the use of English than before the bilingual module.	✘

Hypothesis H1 looks at the effectiveness of bilingual modules on the test group and assumes that the parallel inclusion of a foreign language into the subject content causes disguised motivational effects for English. This, however, cannot be proven as both extrinsic and intrinsic motivation for English decreased from t_0 to t_1; it still decreases until t_2, but on a rather small scale of 0.02 points for intrinsic motivation, while the extrinsic motivation increases by 0.03 points from t_0 to t_2 (see Appendix B). A possible explanation is that the students felt overwhelmed and surprised by the sudden intervention and did not always have the courage to get the German working material, so that they hold the foreign language accountable for their insecurity; this again is reflected in a constant or even decreasing motivation for the foreign language (Mewald, 2015, p. 111).

The second hypothesis, which expected that there are no significant motivational changes for English within the control group, can be verified on the basis of the results. A slight decrease of intrinsic motivation (0.22 points), but also a minor increase of extrinsic motivation (0.05 points) was noticeable (see Appendix A). Since these changes were insignificant, H2 is verified. This may show that even the use of different topics does not discourage the preference to use English in the control group.

Hypothesis H3 is based on the assumption that taking part in a bilingual module will eventually convey positive feelings in terms of a higher self-efficacy for students within the test

group. Once they realise that they are capable of understanding Biology in English, this would ideally result in an increased scholastic self-efficacy (see Table 9). A significant difference between pre and post could not be determined and self-efficacy slightly decreased, which leads to the rejection of the hypothesis. However, the value for self-efficacy rose again slightly in the follow-up test, which is discussed in relation to H7. The students may have felt more successful if they were able to prove themselves in an exam written in English. In this context, again, the factor of being caught off guard, meaning that they could not willingly decide for or against the intervention, might have played a role. Apart from that, it can be suspected that a longer or repetitive intervention would have yielded stronger positive effects.

Hypothesis H4 again relates to the self-efficacy, but now concerning a comparison of the test and control group for the pre-test (t_0). The previous characterisation confirmed that the control group consists of more highly motivated students than those in the test group. Students in the control group consciously chose the bilingual programme and thus feel presumably more confident about their abilities and self-efficacy if they have to face new and complicated tasks. However, the mean values of the test and control group differ by 0.2 points (see Appendix A) with the control group showing higher values (mean of 4.37) but no significant difference, thus discarding H4.

In the fifth hypothesis, it was examined whether the Biology-affinity students could be increased in a subgroup of the test group, namely those characterised with an English-focus at the beginning of the module. For the 22 English-oriented students, Biology as a bilingual subject seemed to be too demanding as the Biology-affinity decreases by 0.17 points (see Table 4). The desired indirect effects of bilingual modules do not seem to have been effective, therefore falsifying H5.

Hypothesis H6 focused on the other subgroup of the test group; it was presumed that the Biology-oriented students would experience an increase in their English-affinity. Indeed, there was a minor increase of 0.07 points (see Table 5), but this was insignificant, leading to the falsification of H6.

The seventh hypothesis examined the long-term changes of the test group regarding English motivation. Even though intrinsic motivation as well as interest for English decreased and extrinsic motivation increased slightly, no significant changes could be observed over time (see Figure 3 and Table 10). This unfortunately means that the bilingual module did not evoke the desired effects and did not show a positive long-term impact on the students' linguistic affinity. However, the value of self-efficacy increased to 4.28 when looking at the follow-up test, after a

recorded low of 4.15 directly after the bilingual module (see Appendix C). Equally favourable is the decrease in the anxiety to use English, which drops once again between the post and follow-up tests.

General observations of the results also allow for the following statements: *anxiety* and *disinterest English* are fortunately not very common in both test and control group. Still it is striking that the answers of the test group included all the values that were provided by the scale, while *anxiety* was answered to in the control group using 1 to 4 and *disinterest* using only 1 to 3. From this, one can infer that the test group is altogether more anxious about using English, which is supported by the observation that the last quartile of answers for anxiety ranges from 4 to 6 in the test group, while it covers an even bigger range from 2 to 6 in the control group. Thus, 75% of the students in consistent bilingual programmes are not frightened or only a little anxious to use English. To answer the research questions, unfortunately no meaningful statements can be made. Up to this point the sample size of students with specific interests is very limited and the changes for their affinities regarding the opposite subjects are not conclusive.

Conclusion

The first run of the study has shown that motivational effects cannot be transferred quickly; i.e., that a CLIL-programme of any sort does not yield the results wished for as fast as would be expected. Even when looking at the follow-up results, motivation and interest for English and Biology decrease in the test group, therefore suggesting that the implication of only one module is not as effective. This can also be seen by the fact that the affinity foci remain with 50% consistency.

Thus, the findings confirm Helfrich's statement that, at points, "bilingual teaching seems to be very demanding for both learners and teachers, [and that] it may be a frustrating experience for average or below average children" (Helfrich, 2003, p. 32, as cited in Mewald, 2015, p. 111). This situation most certainly arises from the students' difficulties in wanting to express something that they do not have words for, meaning that their language competence restricts their output, which in turn leads to the feeling of excessive demand and failure. These feelings have negative effects on the affective-motivational variables (Rumlich, 2015, p. 315) and can possibly be overcome if the students get used to bilingual modules and the teacher, so that they are not afraid to ask questions or answer in the foreign language or just revert to their mother tongue, which is also accepted since content is the focus and not primarily language. However, such an assumption about the relation of CLIL and motivation can just be proven by further longitudinal studies (Doiz

et al., 2014, as cited in Sylvén, 2017, p. 57). It will still remain a very difficult issue to tackle since motivation is a multidimensional and dynamic construct and is deeply rooted in one's personality and learner biography (Riemer, 2016, p. 266/267).

Nonetheless, these results should not discourage bilingual education, as CLIL continues to be considered a "very important new paradigm in language teaching" (Wolff, 2003, p. 12) with its advantages being a high input, contentual focus as well as a bigger contextualisation of the language when compared to regular language classes (Kersten, 2012, p. 195).

References

Abendroth-Timmer, D. (2007). *Akzeptanz und Motivation: Empirische Ansätze zur Erforschung des unterrichtlichen Einsatzes von bilingualen und mehrsprachigen Modulen*. Frankfurt am Main: Peter Lang.

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. DOI: 10.1007/s11999-009-1164-4.

Brandstätter, V., Schüler, J., Puca, R. M., & Lozo, L. (2013). *Motivation und Emotion*. Heidelberg: Springer.

Breidbach, S. (2002). Bilingualer Sachfachunterricht als neues, interdisziplinäres Forschungsfeld. In S. Breidbach et al. (Eds.), *Bilingualer Sachfachunterricht: Didaktik, Lehrer-/Lernerforschung und Bildungspolitik im Spannungsfeld von Theorie und Empirie. Mehrsprachigkeit in Schule und Unterricht* (pp. 11-27). Frankfurt am Main: Peter Lang.

Dahiru, T. (2008). P – Value, a true test of statistical significance? A cautionary note. *Annals of Ibadan Postgraduate Medicine*, 6 (1), 21–26.

Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics. 4th edition. Los Angeles, CA: Sage.

Hannover, B. (1998). The development of self-concept and interests. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Eds.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (pp. 105-125). Kiel: Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel.

Holstermann, N., & Bögeholz, S. (2007). Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften, 13,* 71-86.

Kersten, S. (2012). Wortschatzarbeit im bilingualen Unterricht. In F. Lenz (Ed.), *Bilinguales Lernen. Unterrichtskonzepte zur Förderung sachfachbezogener und interkultureller Kompetenz* (pp. 195-205). Frankfurt am Main: Peter Lang.

KMK (2013). *Bericht „Konzepte für den bilingualen Unterricht – Erfahrungsbericht und Vorschläge zur Weiterentwicklung". Beschluss vom 17.10.2013.* Retrieved from http://www.kmk.org/fileadmin/ Dateien/veroeffentlichungen/beschluesse/2013/2011017-Konzepte-bilingualer-Unterricht.pdf

Kobayashi, Y. (2002). The Role of Gender in Foreign Language Learning Attitudes: Japanese female students' attitudes towards English learning. *Gender and Education*, *14(2),* 181-197.

Kondring, B., & Ewig, M. (2005). Aspekte der Leistungsmessung im bilingualen Biologieunterricht. *IDB – Berichte des Institutes für Didaktik der Biologie, 14 (1),* 49-62.

Krapp, A.s (1998). Entwicklung und Förderung von Interessen im Unterricht. *Psychologie, Erziehung, Unterricht, 44,* 185-201.

Krechel, H.-L. (2003). Bilingual Modules: Flexible Formen bilingualen Lehrens und Lernens. In M. Wildhage (Ed.), *Praxis des bilingualen Unterrichts* (pp. 194-216). Berlin: Cornelsen Scriptor.

Mewald, C. (2015). Lexical Range and Communicative Competence of Learners in Bilingual Schools in Lower Austria. *Global Education Review, 2 (2),* 98–113.

MSW NRW (2011). *Bilingualer Unterricht in Nordrhein-Westfalen*. Retrieved from https://www.schulministerium.nrw.de/docs/Schulsystem/Unterricht/Lernbereiche-und-Faecher/Fremdsprachen/Bilingualer-Unterricht/Kontext/Broschuere_Bilinguale-Unterricht-in-NRW.pdf

MSW NRW (2013). Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen: Biologie. Retrieved from http://www.schulentwicklung.nrw.de/ lehrplaene/ upload/klp_SII/ bi/GOSt_Biologie_Endfassung.pdf

MSW NRW (2014). Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen: Englisch. Retrieved from http://www.schulentwicklung.nrw.de/ lehrplaene/upload/ klp_SII/e/KLP_GOSt_Englisch.pdf

Müller-Schneck, E. (2005). *Bilingualer Geschichtsunterricht: Theorie, Praxis, Perspektiven*. Frankfurt am Main: Peter Lang.

Palesch, Y. Y. (2014). Some common misperceptions about p-values. *Stroke; a journal of cerebral circulation,* 45 (12). DOI: 10.1161/STROKEAHA.114.006138.

Piesche, N., Jonkmann, K., Fiege, C., & Keßler, J.-U. (2016). CLIL for all? A randomised controlled field experiment with sixth-grade students on the effects of content and language integrated science learning. *Learning and Instruction, 44*, 108-116.

Powell, R., & Batters, J. (1985). Students' perceptions of foreign language learning at 12+: some gender differences. *Educational Studies, 1*, 11–23.

Preisfeld, A. (2016). Die Bedeutung bilingualen Experimentalunterrichts in Biologie für die fachliche und sprachliche Kompetenz. In B. Diehr, A. Preisfeld, & L. Schmelter (Eds.), *Bilingualen Unterricht weiterentwickeln und erforschen* (pp. 103-123). Frankfurt am Main: Peter Lang.

Prüfer, K. (2012). Bilinguale (englisch-deutsche) Module im Mathematikunterricht und ihre Auswirkung auf die Lernbereitschaft der Schüler/innen für das Sachfach: Theoretische Grundlagen und ausgewählte Aspekte eines Pilotmoduls. In F. Lenz (Ed.), *Bilinguales Lernen: Unterrichtskonzepte zur Förderung sachfachbezogener und interkultureller Kompetenz* (pp. 149-168). Frankfurt am Main: Peter Lang.

Riemer, C. (2016). Affektive Faktoren. In: E. Burwitz-Melzer, G. Mehlhorn, C. Riemer, K.-R. Bausch, & H.-J. Krumm (Eds.), *Handbuch Fremdsprachenunterricht* (pp. 266-271). 6th completely revised edition. Tübingen: UTB.

Rumlich, D. (2012). The effects of CLIL: Students' linguistic accuracy in relation to internal and external learner variables. In S. Kersten, C. Ludwig, D. Meer, & B. Rüschoff (Eds.), *Language learning and language use - applied linguistics approaches* (pp 115-127). Duisburg: Univ.-Verl. Rhein-Ruhr.

Rumlich, D. (2015). Zur affektiv-motivationalen Entwicklung von Lernenden im CLIL Unterricht. In B. Rüschoff, J. Sudhoff, & D. Wolff (Eds.), *CLIL Revisited: Eine kritische Analyse des gegenwärtigen Standes des bilingualen Sachfachunterrichts* (pp. 309-330). Frankfurt am Main: Peter Lang.

Satow, L. (1999). Zur Bedeutung des Unterrichtsklimas für die Entwicklung schulbezogener Selbstwirksamkeitserwartungen. Eine Mehrebenenanalyse mit latenten Variablen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 31 (4),* 171-179.

Schwarzer, R., & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In Jerusalem, M., & Hopf, D. (Eds.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (pp. 28-53). Zeitschrift für Pädagogik, Beiheft 44. Weinheim: Beltz.

Sung, H., & Padilla, A. (1998). Student motivation, parental motivation, and involvement in the learning of Asian languages in elementary and secondary schools. *Modern Language Journal, 2*, 205 –216.

Sylvén, L. K., & Thompson, A. S. (2015). Language learning motivation and CLIL - Is there a connection? *Journal of Immersion and Content-Based Language Education, 3(1),* 28–50.

Sylvén, L. K. (2017). Motivation, second language learning and CLIL. In A. Llinares, & T. Morton (Eds.), *Applied Linguistics Perspectives on CLIL* (pp. 51-65). Amsterdam: John Benjamins.

Wolff, D. (2003). Integrating language and content in the language classroom: Are transfer of knowledge and language ensured? *ASp, 41-42,* 35-46. DOI: 10.4000/asp.1154.

Appendix A. t-Tests for the comparison of all constructs between t_0 (pre) and t_1 (post), differentiated into test and control group

Group	Construct	N	Mean	SD	T	df	Sig.
Test	Interest Biology Pre	82	3.43	1.01	0.996	81	0.322
	Interest Biology Post	82	3.35	1.16			
Control	Interest Biology Pre	30	3.41	1.15	-0.252	29	0.803
	Interest Biology Post	30	3.44	1.39			
Test	Intrinsic Motivation Biology Pre	82	3.94	1.13	2.628	81	0.010
	Intrinsic Motivation Biology Post	82	3.69	1.21			
Control	Intrinsic Motivation Biology Pre	30	3.64	1.30	-1.032	29	0.311
	Intrinsic Motivation Biology Post	30	3.81	1.36			
Test	Extrinsic Motivation Biology Pre	81	3.89	0.94	0.948	81	0.346
	Extrinsic Motivation Biology Post	81	3.82	0.98			
Control	Extrinsic Motivation Biology Pre	30	3.71	1.19	2.248	29	0.032
	Extrinsic Motivation Biology Post	30	3.38	1.20			
Test	Interest English Pre	82	4.07	1.19	2.119	81	0.037
	Interest English Post	82	3.92	1.26			
Control	Interest English Pre	31	5.05	0.92	1.368	30	0.181
	Interest English Post	31	4.89	1.00			
Test	Intrinsic Motivation English Pre	82	3.54	1.07	1.745	81	0.085
	Intrinsic Motivation English Post	82	3.43	1.09			
Control	Intrinsic Motivation English Pre	31	4.43	1.05	1.657	30	0.108
	Intrinsic Motivation English Post	31	4.21	1.14			
Test	Extrinsic Motivation English Pre	82	4.25	1.21	0.426	80	0.671
	Extrinsic Motivation English Post	82	4.20	1.19			
Control	Extrinsic Motivation English Pre	31	4.71	1.12	-0.318	30	0.753
	Extrinsic Motivation English Post	31	4.76	1.13			
Test	Anxiety English Pre	82	2.88	1.32	1.319	81	0.191
	Anxiety English Post	82	2.78	1.22			
Control	Anxiety English Pre	31	1.79	0.75	-0.347	30	0.731
	Anxiety English Post	31	1.84	0.74			
Test	Disinterest English Pre	82	1.48	1.00	-1.558	81	0.123
	Disinterest English Post	82	1.55	1.00			
Control	Disinterest English Pre	31	1.10	0.34	-1.609	30	0.118
	Disinterest English Post	31	1.17	0.43			
Test	Self-efficacy bilingual education Pre	82	4.17	0.86	0.403	81	0.688
	Self-efficacy bilingual education Post	82	4.15	0.93			
Control	Self-efficacy bilingual education Pre	31	4.37	0.72	-0.920	30	0.365
	Self-efficacy bilingual education Post	31	4.49	0.89			

Appendix B. t-Tests for the comparison of all constructs, at both t_0 (pre) and t_1 (post) between test and control group

Construct	Group	N	Mean	SD	T	df	Sig.
Interest Biology Pre	Test	82	3.43	1.01	-0,016	111	0,987
	Control	31	3.44	1.14			
Interest Biology Post	Test	82	3.35	1.16	-0,358	110	0,721
	Control	30	3.44	1.39			
Intrinsic Motivation Biology Pre	Test	82	3.94	1.13	1,176	111	0,242
	Control	31	3.65	1.28			
Intrinsic Motivation Biology Post	Test	82	3.69	1.21	-0,465	110	0,643
	Control	30	3.81	1.36			
Extrinsic Motivation Biology Pre	Test	82	3.89	0.94	0,877	111	0,382
	Control	31	3.70	1.17			
Extrinsic Motivation Biology Post	Test	82	3.82	0.98	1,983	110	0,050
	Control	30	3.38	1.20			
Interest English Pre	Test	82	4.07	1.19	-4,142	111	0,000
	Control	31	5.05	0.92			
Interest English Post	Test	82	3.92	1.26	-3,869	111	0,000
	Control	31	4.89	1.00			
Intrinsic Motivation English Pre	Test	82	3.54	1.07	-3,963	111	0,000
	Control	31	4.43	1.05			
Intrinsic Motivation English Post	Test	82	3.43	1.09	-3,359	111	0,001
	Control	31	4,.21	1.14			
Extrinsic Motivation English Pre	Test	82	4.23	1.21	-1,906	111	0,059
	Control	31	4.71	1.12			
Extrinsic Motivation English Post	Test	81	4.20	1.19	-2,261	110	0,026
	Control	31	4.76	1.13			
Anxiety English Pre	Test	82	2.88	1.32	5,515	93,430	0,000
	Control	31	1.79	0.75			
Anxiety English Post	Test	82	2.78	1.22	4,968	88,708	0,000
	Control	31	1.84	0.74			
Disinterest English Pre	Test	82	1.48	1.00	2,940	110,295	0,004
	Control	31	1.10	0.34			
Disinterest English Post	Test	82	1.55	1.00	2,858	109,407	0,005
	Control	31	1.17	0.43			
Self-efficacy bilingual education Pre	Test	82	4.17	0.86	-1,124	111	0,264
	Control	31	4.37	0.72			
Self-efficacy bilingual education Post	Test	82	4.15	0.93	-1,767	111	0,080
	Control	31	4.49	0.89			

Appendix C.

Depiction of means of all constructs at all points of measurements (pre, post, and follow-up) for both test and control group. Means are based on the number of students who were present at each of the testings

Construct	Group	Pre	Post	Follow-up
Interest Biology	Test	3.43	3.35	3,23
	Control	3.44	3.44	3,60
Intrinsic Motivation Biology	Test	3.94	3.69	3,64
	Control	3.65	3.81	4,01
Extrinsic Motivation Biology	Test	3.89	3.82	3,81
	Control	3.70	3.38	3,82
Interest English	Test	4.07	3.92	3,89
	Control	5.05	4.89	4,93
Intrinsic Motivation English	Test	3.54	3.43	3,41
	Control	4.43	4.21	4,23
Extrinsic Motivation English	Test	4.23	4.20	4,26
	Control	4.71	4.76	4,78
Anxiety English	Test	2.88	2.78	2,74
	Control	1.79	1.84	1,69
Disinterest English	Test	1.48	1.55	1,53
	Control	1.10	1.17	1,22
Self-efficacy bilingual education	Test	4.17	4.15	4,28
	Control	4.37	4.49	4,51

GLM for the inspection of differences for all constructs over time within the control group. Based on data of students who were present at all three testings (n = 28).

Construct	Sum of squares (Type III)	df	Mean squares	F	Sig.	Partial Eta²
Interest Biology	1.456	2	0.728	2.465	0.095	0.087
Intrinsic Motivation Biology	2.359	2	1.179	2.738	0.074	0.095
Extrinsic Motivation Biology	3.181	2	1.590	4.199	0.020	0.139
Interest English	0.412	2	0.206	1.115	0.335	0.040
Intrinsic Motivation English	0.536	2	0.268	1.298	0.281	0.046
Extrinsic Motivation English	0.203	2	0.102	0.261	0.772	0.010
Anxiety English	0.397	2	0.198	0.869	0.425	0.031
Disinterest English	0.167	2	0.083	2.000	0.145	0.069
Self-efficacy bilingual education	0.288	2	0.144	0.532	0.590	0.019