

Pre-print version of article to be published in the Journal of Immersion and Content-based Language Education in 2021, DOI: 10.1075/jicb.19010.mea. Please do not cite directly from this version.

Sketching a motivational landscape: Motivational variation within bilingual secondary education in the Netherlands

Tessa Mearns & Nivja de Jong
Leiden University

Abstract:

Studies of motivation in bilingual education settings often address questions of differences between learners in bilingual programmes and those in mainstream education. Problematic in this respect is our increasing awareness of the inherent differences between these two learner groups, as learners in bilingual programmes have often chosen or been selected for a bilingual route (Mearns et al., 2017). The study presented here therefore does not seek to compare learners in bilingual and non-bilingual programmes, but rather to explore the nature of language learning motivation within the context of bilingual secondary education (BSE) in the Netherlands. Using a purpose-designed tool reflecting the L2 Motivational Self System (Dörnyei, 2009), this study investigated trends in motivation across genders, academic tracks and year-groups of nearly 2000 learners. Findings suggest that, although these learners all have bilingual education in common, differences between the motivations of these groups should not be overlooked.

Keywords: bilingual education; CLIL; language motivation; Netherlands; L2 Motivational Self System; gender; secondary education

Introduction

Content and Language Integrated Learning (CLIL) and other forms of bilingual education are often hailed as being an educational ideal in terms of efficiency, learning outcomes and learner motivation (Dalton-Puffer & Smit, 2013). While bilingual approaches appear in theory to be potentially rich opportunities for the learning of both language and subject content (Coyle et al., 2010), however, studies have suggested that this may not always be the case (e.g. Sylvé, 2013). Furthermore, as the majority of bilingual education programmes involve a degree of selectivity or self-selection, it is not always clear whether there is a value-added difference between cohorts in bilingual and mainstream programmes (Bruton, 2011).

Likewise, as motivation to learn a new language (henceforth L2 motivation) is increasingly recognised as a complex and context-bound phenomenon (Dörnyei & Ushioda,

2011), it follows that L2 motivation in a bilingual education context will be equally complex. Somers and Llinares (2018) remark that repeated claims that CLIL is a motivating approach have little scientific basis. Furthermore, as argued by Mearns (2015) and Mearns et al. (2017), the inherent differences between the learners who opt in or out of bilingual secondary education in the Netherlands suggest that direct comparisons between the two groups of learners may be of little value in determining the impact of the approach on their motivation.

For this reason, this research paper seeks not to compare learners in bilingual and mainstream programmes, but to explore L2 motivation among pupils who have all chosen a bilingual education route. Motivation has been referred to as “an attractive point of intersection between theory and practice in the psychology of language learning” (Boo et al., 2015, p. 149), highlighting that increasing our understanding of the ways in which learners are motivated can help us to understand and develop educational practices. In deepening our understanding of motivational trends among learners in bilingual education, it is hoped that we can gain insight into the types of learners who choose this route in order to better respond to their needs.

1. The context for research: Dutch bilingual secondary education

Bilingual secondary education (henceforth BSE) in the Netherlands is a well-established educational model in which learners who are mostly native speakers of Dutch follow approximately 50% of their lower-secondary education (age 12-15) in English. In upper-secondary, the majority of subjects transition back to Dutch in order to prepare learners for their final central examinations, which they sit in Dutch. From this point, the curriculum for English focuses on the International Baccalaureate (IB) alongside the standard Dutch curriculum.

Schools that offer BSE are required to apply for accreditation by the National Network of Bilingual Schools, which has laid out a quality standard for BSE (henceforth BSE

Standard; Nuffic, 2012, originally published in Dutch in 2003).¹ One of the criteria stipulated in the BSE Standard is that teachers should be adequately trained in CLIL. Therefore, while BSE is in essence an organisational principle (Mearns & de Graaff, 2018a), it also has pedagogical implications.

Approximately 50% of secondary school pupils in the Netherlands follow the higher general (HG for the purpose of this paper) or pre-academic (Pre-A) track in secondary education and 50% a pre-vocational (Pre-V) track. Placement in a particular track is based largely on the advice of the child's primary school. It has been suggested that learner characteristics such as cognitive skills, motivation, study skills, learning preferences and academic performance differ across tracks (Michels, 2006). As Michels emphasises, however, lack of empirical research in this regard makes it difficult to establish an objective picture of whether these differences are inherent to learners, the perception of teachers or the result of a Pygmalion effect.

The majority of BSE schools in the Netherlands offer a bilingual Pre-A track, although an increasing number also offer a bilingual option at HG and/or Pre-V levels. In the current study, data were only collected from learners in Pre-A and HG BSE tracks as the BSE experience across these tracks was expected to be similar. Pre-A and HG classes often follow near-identical curricula, with differentiation only in the level of performance expected in assessments (Stichting Havoplatfom, 2019). Likewise, the BSE Standard (Nuffic, 2012, 2019) is aimed at both Pre-A and HG departments and draws no distinction between the two beyond the organisational factors and attainment levels outlined in Table 1.

¹ From 2019, a revised Standard for BSE has been introduced (Nuffic, 2019). While the 2019 Standard is relevant to the implications of the current study, the data explored here were collected in the context of BSE according to the original 2003 Standard.

Table 1. Differences between BSE in Pre-A and HG streams, according to the original BSE Standard (Nuffic, 2012)²

Track	English after 3 years (CEFR)	IB English	Contact hours in English during senior years
HG (5-years)	B1	English B Higher Level or English A Language & Literature Standard Level	850 out of 3,200 (27%) over 2 years
Pre-A (6-years)	B2	English A Language & Literature Higher or Standard Level	1,150 out of 4,800 (24%) over 3 years

2. Learner variables, (L2) motivation and CLIL

While it is impossible to draw definitive conclusions regarding the interaction between situational and individual variables and motivation, a number of trends have emerged from the literature. Relevant to the current study is the tendency for females to be more motivated language learners than males (Csizér & Dörnyei, 2005; Henry, 2009) and— in secondary school contexts – for L2 motivation to be lower among older learners (Lamb, 2017). In some cases, these trends have been found to hold up in CLIL or bilingual education contexts, while others have suggested that CLIL could serve to counteract them and thus level the playing field.

In the case of gender differences, Kissau (2006) suggests that the female advantage in L2 motivation may be in part due to features of the language classroom that are more appealing to girls, but also due to broader contextual factors such as societal expectations. Mills (2014) purports that these factors might also influence boys' self-efficacy in language learning, which can in turn affect motivation. Lasagabaster (2008) has hypothesised that CLIL could help to increase boys' motivation for language learning. This hypothesis appeared to hold up in studies by Merisuo-Storm (2007), and Mearns and de Graaff (2018b), who

² Please note that the requirements with regard to the senior years and the IB have changed in the 2019 Standard (Nuffic, 2019), although there is still no distinction drawn between HG and Pre-A tracks beyond the organizational level.

observed a narrower gap between the motivations of boys and girls in bilingual classes than in mainstream education. In the latter study, which was carried out in the context of Dutch BSE at HG level, BSE boys even reported stronger motivation for English than their female peers.

Studies that have investigated differences in motivation across different year-groups or changes in motivation over time within CLIL or bilingual education contexts have echoed findings from elsewhere in L2 motivation research, providing little evidence that a bilingual approach counteracts the “attitudinal decline” referred to by Lasagabaster and Sierra (2009, p. 13) (e.g. Doiz et al., 2014; Mearns & de Graaff, 2018b; Mearns et al., 2017). A longitudinal study by Thompson and Sylvén (2019) highlights that, while CLIL learners decreased in anxiety over time, so too did their motivation. An exception to this trend is a study by Denman et al. (2018), in which attitudes towards learning English among learners in Dutch Pre-V BSE were more positive following one or more years in the programme, but then dropped once the majority of subjects had switched back to Dutch in the senior years.

In terms of L2 motivational differences among learners studying at different academic levels in mainstream Dutch secondary education, published research has not been found. In terms of CLIL and bilingual education, publications suggested that CLIL might be particularly beneficial for learners with a less academic profile (Genesee, 2007). Verspoor et al. (2015) found that Pre-A BSE learners scored higher than HG BSE learners in terms of both English language proficiency and L2 motivation, but also that HG and Pre-A BSE learners had different preferences in terms of approaches to teaching and learning.

3. A Framework of L2 Motivation: The L2 Motivational Self System

According to a review published in 2015 (Boo et al., 2015), the most commonly applied model of L2 motivation in the current age is Dörnyei’s (2009) L2 Motivational Self System (L2MSS). This framework also contributed to the design of the current study.

The L2MSS draws on a number of earlier models from both L2 motivation and cognitive psychology (for an overview, see Dörnyei & Ushioda, 2011 or Mearns, 2015). The model focuses on the image the learner holds of themselves as a speaker of the target language in the future, as well as on their present experience of learning the language. The L2MSS consists of three elements:

- The *Ideal L2 Self*, i.e., the image one has of the L2-speaker they would like to become and the ways in which they envisage themselves using the L2
- The *Ought-to L2 Self*, i.e., the image one has of the L2-speaker that others (e.g. parents, friends, teachers, society) expect them to become
- The *L2 Learning Experience*, i.e., the experience of learning the language (usually in the language classroom)

According to Dörnyei, the more vivid and realistic the image, and the better the alignment between the Ideal and the Ought-to L2 selves, the more motivating these components will be (Hadfield & Dörnyei, 2013). A positive learning experience will complement these and further strengthen motivation.

The L2MSS has formed the basis of multiple studies and a number of original research instruments (for examples, see contributions in Dörnyei & Ushioda, 2009). Early studies focused not only on investigating motivation but also on validating and critiquing Dörnyei's original model. The Ideal L2 Self has held its ground in various studies. With respect to gender-differences, it has generally appeared stronger among female learners (Azarnoosh & Birjandi, 2012; Henry, 2009; Henry & Cliffordson, 2013; Ryan, 2009).

In contrast to this, the Ought-to L2 self has been found to be stronger among boys (Azarnoosh & Birjandi, 2012; Heras & Lasagabaster, 2014). Questions have been raised, however, as to the reliability of the Ought-to L2 self, which has been found to be an unreliable

predictor of language learning success (Al-Hoorie, 2018; Dörnyei & Chan, 2013). An explanation that has been offered for this unreliability is that experienced expectations of others were not always internalised by learners.

Relatively little attention has been paid in research to the L2 Learning Experience (Al-Hoorie, 2018), although this element has been identified as having higher predictive value of motivated behaviour than the self-oriented elements of the L2MSS (Lamb, 2012). Likewise, Dörnyei (2001) highlights first and foremost the importance of a safe, secure and positive learning environment if learners are to be motivated to learn. In the current study, which focuses on learners in the specific BSE context, about which little is known in motivational terms, the learning experience is of particular interest but also requires specific attention in terms of survey design.

Two earlier research projects exploring L2 motivation in CLIL or bilingual education used the L2MSS in their theoretical framework (the Netherlands: Mearns & de Graaff, 2018b; Mearns et al., 2017; Sweden: Sylvén & Thompson, 2015; Thompson & Sylvén, 2019). While both sets of studies aimed to draw comparisons with learners in mainstream education rather than focusing explicitly on distinct learner groups within bilingual programmes, some conclusions regarding trends within CLIL/BSE groups could nevertheless be drawn. In earlier studies in the Dutch BSE context, Mearns and de Graaff (2018b) found that BSE boys appeared to have a stronger vision of their future L2 self than did BSE girls, while in Mearns et al. (2017), first-year BSE learners were more positive than second or third-year BSE learners regarding the experience of English lessons. Studies from the Swedish CLIL context did not reflect the same trend in either respect. In fact, they found no significant differences between boys and girls in the variables directly related to the L2MSS. In other variables not directly related to the L2MSS, Mearns and colleagues found no significant differences between the BSE sub-groups (by year and gender). In Sweden, CLIL boys appeared more

confident language learners than CLIL girls (Sylvén & Thompson, 2015). Also, within the CLIL group, participants appeared to become less anxious but also less motivated for English over time.

4. Aims and focus of the current study

The goal of the current study is to identify differences in L2 motivation between groups of learners within BSE in order to inform future research, policy and practice, rather than to compare BSE learners to those in mainstream education or evaluate relationships between learners' motivation and their attainment. This could contribute to existing – and somewhat contradictory – findings regarding differences between boys and girls, and between learners in different year-groups, and provide new insights into differences between more or less academically-oriented groups of learners. The research question that will be addressed in this study is therefore:

To what extent does L2 English motivation of learners in Dutch bilingual secondary education (BSE) vary between

- a. boys and girls?
- b. learners from different educational tracks?
- c. different year-groups?

5. Method

The survey explored here was part of a large-scale questionnaire study into L2 motivation in BSE, of which only the quantitative findings are reported here. The research design and process are described below.

5.1 Data collection

Data were collected from pupils following BSE in the HG or Pre-A tracks at 19 schools across different regions of the Netherlands in September-October 2017. Data collection occurred online during school time, under the supervision of a teacher who had

been provided with written instructions regarding how to facilitate data collection, and a short set of instructions to be read out to pupils. The questionnaire was written in Dutch.

By way of consent, schools were asked to distribute an information letter and opt-out form to pupils and their parents at least two weeks before the start of data collection. Pupils who returned the opt-out form did not participate in the research.

5.2 Recruitment and participants

All 132 schools in the National Network of Bilingual Schools were invited to participate in the study. Of the 30 schools that responded to the invitation, 19 actually participated. Schools were asked to invite all pupils in first, third and fifth year of Higher General (HG) and Pre-academic (Pre-A) bilingual education (BSE) to respond to the questionnaire. At the beginning of the school year, first-year secondary school pupils are usually aged between 11 and 13, third-year between 13 and 15, and fifth-year between 15 and 17.

2078 responses were collected, of which 90 were removed due to being incomplete, duplicate responses or invalid (e.g., the same answer selected for every scale). Responses were considered incomplete if any of the Likert-scale items were missing. Responses missing demographic data were retained, although it was not possible to include them in all of the analyses.

Of the 1988 valid responses (between 26 and 286 per school), 51.4% were from girls (N=1,021) and 44.7% from boys (N=888). 4% of respondents (N=79) preferred not to share this information or did not complete this section of the questionnaire. These cases were excluded from analyses that included gender as an independent variable.

80.5% of responses (N=1601) were from pupils in Pre-A, 18.3% (N=364) from HG and 1.2% (N=23) were missing this information. The HG group included pupils in first-year combination classes for HG/Pre-A and HG/Pre-V, designed to delay the choice of track until

second or third year. The heavy weighting towards Pre-A is a reflection of the larger number of Pre-A BSE programmes on offer.

45% of respondents (N=894) were in their first year of secondary school, 33.2% (N=661) in third year and 20.6% (N=410) in fifth year. This information was missing from 1.2% of responses (N=23). The lower number of responses from fifth-years is because not all schools offer bilingual education beyond third year.

5.3 Survey design

The survey reported here refers to the quantitative data collected through a series of items that formed the largest section of a purpose-designed online questionnaire, referred to henceforth as the English Motivation in Bilingual Education Questionnaire (EMBEQ).

The scales and items used in the EMBEQ were based on the Student Motivational State questionnaire (SMS) developed by Papi and Abdollahzadeh (2012) for the Iranian EFL context. The SMS was selected as a basis due to its length, being considerably shorter than other instruments such as Ryan's (2009) Motivational Factors Questionnaire (MFQ), and also because it included all three components of the L2MSS (Ideal L2 Self, Ought-to L2 Self, English Learning Experience), as well as two scales that related to other affective factors: Motivational Intensity and Linguistic Self-Confidence. Due to the differences between the research contexts, however, considerable adaptations were necessary (Dörnyei & Ushioda, 2011). These adaptations were carried out in two phases: firstly, manual adaptation for the Dutch BSE context, and piloting in that context; secondly, principal component analysis to confirm the validity of the survey scales. Each of these phases is explained below.

5.3.1 *Manual adaptation for Dutch BSE*

In the initial adaptation of the scales used by Papi and Abdollahzadeh (2012), cultural differences were addressed through changes in wording or omission or replacement of irrelevant items. For example, the original item, "I imagine myself living and making friends

in a modern community, using English” from the Ideal L2 Self scale in the SMS was changed to “I imagine myself using English to make friends in other countries,” as it was thought that the “modern community” would not be a relatable ambition for young people in the Netherlands.

Further adaptation was necessary in order to take into account that respondents from BSE did not only experience English in their English lessons, but across a range of curriculum areas. In the Motivational Intensity and Linguistic Self-Confidence scales, such adjustments largely involved rewording the items. For example, “When I am in English class, I volunteer answers as much as possible” (SMS Motivational Intensity scale) became “When I am in lessons where English is used, I volunteer answers as much as possible” for the purpose of the EMBEQ.

With regard to the items referring to the learning experience, it was necessary to distinguish between (a) the experience of English lessons, (b) the experience of learning a subject that happens to be taught in English and (c) the experience of learning English *through* another subject. The SMS English Learning Experience scale was therefore replaced with three separate scales: the English lesson experience scale, the subject learning experience scale, and the English-through-subject learning experience scale. For the final two scales, respondents were asked to reflect on the experience of learning their favourite English-medium subject to avoid interference from differences between subjects.

The resulting survey consisted of 47 Likert-style items to which participants were asked to select a response on an odd-numbered, 5-point scale (Preston & Colman, 2000) from “Strongly disagree” (1) to “Strongly agree” (5). This version of the instrument was piloted with 95 first and third year pupils from a BSE school that did not participate in the final study, and structure and wording were adapted slightly on the basis of their feedback.

5.3.2 *Principal component analysis and resulting alterations*

A series of factor analyses and tests of reliability were carried out using the data from the 1988 respondents in the current study in order to further refine the survey design and ensure its fitness for the Dutch BSE context. Principal component analysis (PCA) suggested that some small adjustments to the original tool would increase its validity in the current context. Seven of the original 47 items were removed, either because they did not load onto any factors or because they loaded onto factors that did not make theoretical sense. The resulting final version of the survey included 40 items across eight scales. A full account of the PCA and a complete list of survey items are available via the IRIS online repository (Marsden et al., 2016; IRIS, 2020³).

The resulting eight scales used in the analysis of responses are summarised in Table 2, along with data pertaining to the reliability (Cronbach's alpha, α) and internal consistency (summability: Goeman & de Jong, 2018) of the scales. Cronbach's alpha (α) scores for all eight scales were above .6. This is considered an acceptable level of reliability in SLA research (Dörnyei, 2003). Internal consistency (gauged by summability) was between .3 and .5, meaning that 30-50% of the variance between items of the subscale is captured in the mean score of that subscale.

Table 2. Scales 1-8 of the EMBEQ, including operationalisation, Cronbach's alpha (α) and Summability

Scale	Operationalisation	Items	N	α	Summability
1. Ideal L2 Self	<i>Images or expectations of what the L2 could mean for who you become in the future</i>	5	1988	.69	.32
2. Ought-to L2 Self	<i>Ideas about what other people expect of you with regard to the L2</i>	6	1988	.79	.39
3. Motivated Learning Behaviour	<i>Self-reported motivated behaviour when learning English in school</i>	4	1988	.78	.37
4. Motivation for English	<i>Desire to learn/use English outside of school/BSE</i>	3	1988	.64	.37
5. L2 Self-Confidence	<i>Confidence in own ability to learn English</i>	7	1988	.80	.37

³ Materials can be found by searching for the first author by name or for the term "EMBEQ."

Scale	Operationalisation	Items	N	α	Summability
6. English Lesson Experience	<i>Experience of learning English through English lessons</i>	6	1988	.89	.59
7. Subject Learning Experience	<i>Experience of learning favourite English-medium subject</i>	5	1517	.86	.56
8. English-Through-Subject Learning Experience	<i>Experience of learning English through favourite English-medium subject</i>	4	1517	.77	.45

5.3.3 Additional data

Respondents were offered the opportunity to elaborate on their responses. These optional, open responses are considered here as a supplement to the quantitative data, providing additional insight into the interpretation of items and the thinking behind the responses, although they are not handled as a data source in their own right.

5.4 Data analysis

Quantitative analyses were carried out using SPSS 25. For each scale, an analysis was carried out to ascertain the effects of Gender, Track and Year. Interaction effects were not included as these did not relate to the research questions. As the data were drawn from participants in different schools, our data were structured hierarchically and the assumption of independence of observation needed for carrying out (M)ANOVA was not met. We therefore opted to analyse the data using linear multilevel analyses (Snijders & Bosker, 1999). The multilevel procedure explicitly models dependencies in the data by including School as a random factor.

Indeed, analyses showed that School contributed significantly to variance for all eight scales. School was therefore taken as random intercept. Year, Track and Gender were added as random slopes where necessary to find the most fitting model. We checked the random structure of each model by comparing the model fit of two competing models (e.g. including or excluding the random slope) using chi-square. In order to examine the effects of Year for

Scales 1-6, which included three groups of respondents (Years 1, 3 and 5), EM Means posthoc tests were included in the analyses.

Regarding the effects of Gender, cases in which no gender was specified were excluded. Fifth-year responses were excluded for Scales 7 and 8 on the basis of participants' comments that they had difficulty identifying a favourite English-medium subject as so many of their subjects were now taught in Dutch. Cases in which no valid favourite subject was identified (N=79) were also excluded from the analysis of Scales 7 and 8. Missing responses were excluded pairwise.

Analysis of the optional additional comments on the scaled items was carried out in Atlas ti 8.2. Comments were categorized according to the item to which they referred, grouped together per scale and coded for demographic groups on the basis of gender, track and year. All comments were coded inductively for emergent themes that might support decisions regarding the retention of items following the PCA and assist interpretation of the quantitative data. The decisions taken regarding retention of items are explained in the account of the PCA published in the IRIS digital repository (Marsden et al., 2016; IRIS, 2020). As the spread of the comments offered was uneven due to this being an optional aspect of the questionnaire, these data were not considered representative of the group as a whole, nor of specific sub-groups.

6. Results

The results of the analyses are organized by effects of the three independent factors (Gender, Track and Year), rather than by dependent variable (scores on the eight scales). First, additional insights obtained through the optional comments are summarised.

6.1 Additional comments

From 1988 surveys, 2450 individual comments relating to the survey or survey items were filled out. Some comments were counted multiple times as they referred explicitly to

more than one item. The comments provided extra insight into respondents' interpretation of the survey items, which can aid our understanding of the results. To this end, they are presented per dependent variable. Where further relevant observations can be made regarding differences between genders, tracks and year-groups, these are also highlighted.

With regard to the Ideal L2 Self, additional responses generally offered further elaboration on the types of ambitions participants held with regard to their use of English in the future or emphasised their uncertainty in this regard. The ambitions expressed varied in terms of specificity. Girls tended to express more concrete ambitions than boys. Comments related to the Ought-to L2 Self were mostly used to reject the suggestion that learners were influenced in their choices by the views of others.

With regard to Motivated Learning Behaviour and Motivation for English, it appeared that keenness to learn and use English was offset for some learners by practical and social considerations, such as the time that would be required to do extra tasks and not wanting to raise their hand for every answer for fear of dominating the lesson. As reflected in the factor loadings, which divided the original Motivational Intensity scale into two, motivation to learn English did not appear to be directly related to school, some respondents (in particular boys) emphasizing that they had already been learning English before entering BSE: *"I already knew three-quarters of my English before I started BSE. That's because of playing computer games."* (3rd year Pre-A boy)

Most comments regarding L2 Self-Confidence referred to attributions of high rather than low confidence. The most prominent comments in this regard were attributed to aspects of learners' own character or behaviour, such as being good at languages, speaking English independently of school or working hard: *"I do well, but I'm not really proud of that or anything. It's just normal for me."* (3rd year Pre-A boy)

Some respondents attributed their confidence to the learning experience in school, for example expressing a belief that the BSE experience would help them to improve their English, or mentioning that they felt they were progressing in terms of specific aspects of language learning: *“We’ve never had to make any effort for English because we were thrown straight in at the deep end by being forced to speak English.”* (5th year Pre-A boy)

With regard to the learning experience, comments regarding English lessons were more negative than positive. These comments referred to a number of different reasons for liking or disliking lessons, of which the teacher featured most prominently.

Relatively few respondents added additional comments on the learning experience in lessons other than English. The majority of positive comments related to subject lessons referred either to the subject itself or to the teacher, rather than to teaching methods. Other comments were mostly explanations of why respondents would not like more lessons in their favourite subject, i.e. due to their timetable already being too full. A comment that stood out from one respondent referred to the difference he experienced between English lessons and English-medium lessons for other subjects: *“Lessons **in** English are fun. English lessons are a satanic ritual for which I’d rather not be alive.”* (3rd year Pre-A boy)

Most fifth-year comments regarding subjects taught through English provide factual information on not following subjects in English anymore, which was what led to the decision to exclude fifth years’ responses to these items from the quantitative analyses. Some fifth years took the opportunity to express dissatisfaction at this situation:

“The classes on this subject in 1st till 3rd grade were in english but now they’re in dutch concerning the finals are coming up. I do miss the english classes as most programs we work with are still in english” (5th year HG girl, original English text)

The above observations regarding the additional comments will be referred to in the discussion of results, where they support interpretation of the quantitative findings.

6.2 Quantitative analyses

6.2.1 Gender

Table 3 offers a summary of the mean responses from boys and girls. Table 4 displays the results of the seven mixed model analyses for effects of Gender. As shown in Tables 3 and 4, significant main effects of Gender were identified for the Ideal L2 Self ($F(7, 1883.20) = 11.182, p < .001$), the Ought-to L2 Self ($F(8, 1883.61) = 9.510, p = .002$), L2 Self-Confidence ($F(7, 1884.94) = 4.095, p = .043$) and the English Lesson Experience ($F(8, 1868.64) = 6.253, p = .012$). Girls scored significantly higher for the Ideal L2 Self and English Lesson Experience, while boys scored significantly higher for Ought-to L2 Self and L2 Self-Confidence.

Table 3. Descriptive statistics for boys and girls (significantly highest scores in bold)

	Scale [§]	1 <i>Ideal</i>	2 <i>Ought</i>	3 <i>Behaviour</i>	4 <i>Motivation</i>	5 <i>Confidence</i>	6 <i>English</i>	7 <i>Subject</i>	8 <i>E-t-S</i>
girl	M	4.01	1.95	3.43	3.52	3.76	3.28	4.01	3.81
	N	1021	1021	1021	1021	1021	1021	783	783
	SD	0.58	0.63	0.65	0.73	0.55	0.76	0.68	0.64
boy	M	3.93	2.03	3.39	3.45	3.82	3.21	4.04	3.82
	N	888	888	888	888	888	888	685	685
	SD	0.57	0.65	0.63	0.76	0.55	0.79	0.74	0.68
Total	M	3.97	1.99	3.41	3.49	3.79	3.24	4.02	3.82
	N	1909	1909	1909	1909	1909	1909	1468	1468
	SD	0.57	0.64	0.64	0.74	0.55	0.77	0.71	0.66

[§] Titles of scales shortened here and in subsequent tables from 1. Ideal L2 Self, 2. Ought-to L2 Self, 3. Motivated Learning Behaviour, 4. Motivation for English, 5. L2 Self-Confidence, 6. English Lesson Experience, 7. Subject Learning Experience, 8. English-Through-Subject Learning Experience

Table 4. Results of the 8 mixed model analyses for main effects of Gender, with School as random intercept

Scale	Random slope(s)	df	F	p
1. Ideal L2 Self	-	7	11.182	.001*
2. Ought-to L2 Self	year	8	9.510	.002*
3. Motivated Learning Behaviour	year	8	2.253	.134
4. Motivation for English	-	7	3.584	.058
5. L2 Self-Confidence	-	7	4.095	.043*
6. English Lesson Experience	track	8	6.253	.012*
7. Subject Learning Experience	-	6	.652	.419

Scale	Random slope(s)	df	F	p
8. English-Through-Subject Learning Experience	-	6	.073	.787

* = significant at $p < .05$. ** $p < .001$

6.2.2 Educational Track

Table 5 displays a summary of the mean responses for each scale per Educational Track, while Table 6 shows the results of the mixed-model analyses for effects of Track. Pre-A scored significantly higher for the Ideal L2 Self ($F(7, 1418.98) = 3.885, p = .049$) and the Subject Learning Experience ($F(6, 1378.28) = 8.085, p = .005$), while HG scored higher for Motivated Learning Behaviour ($F(8, 1167.77) = 5.212, p = .023$), Motivation for English ($F(7, 1433.82) = 10.923, p < .001$) and L2 Self-Confidence ($F(7, 1166.35) = 4.210, p = .040$).

Table 5. Descriptive statistics for Pre-A and HG tracks (significantly highest scores in bold)

Scale	1 <i>Ideal</i>	2 <i>Ought</i>	3 <i>Behaviour</i>	4 <i>Motivation</i>	5 <i>Confidence</i>	6 <i>English</i>	7 <i>Subject</i>	8 <i>E-t-S</i>
Pre-A	M	4.00	1.99	3.38	3.47	3.78	3.20	4.05
	N	1551	1551	1551	1551	1551	1207	1207
	SD	0.57	0.64	0.63	0.74	0.55	0.77	0.71
HG	M	3.86	1.99	3.53	3.58	3.83	3.43	3.92
	N	344	344	344	344	344	261	261
	SD	0.59	0.65	0.67	0.76	0.54	0.77	0.71
Total	M	3.97	1.99	3.41	3.49	3.79	3.24	4.02
	N	1895	1895	1895	1895	1895	1468	1468
	SD	0.58	0.64	0.64	0.74	0.55	0.78	0.71

Table 6. Results of 8 mixed model analyses for main effects of Track, with School as random intercept

Scale	Random slope(s)	df	F	p
1. Ideal L2 Self	-	7	3.885	.049*
2. Ought-to L2 Self	year	8	.057	.812
3. Motivated Learning Behaviour	year	8	5.212	.023*
4. Motivation for English	-	7	10.923	.001*
5. L2 Self-Confidence	-	7	4.210	.040*
6. English Lesson Experience	track	8	.637	.449
7. Subject Learning Experience	-	6	8.085	.005*
8. English-Through-Subject Learning Experience	-	6	.026	.873

* = significant at $p < .05$.

6.2.3 Year-group

Table 7 displays a summary of the mean responses for each scale per Year-group, while Table 8 summarises the mixed model analyses for effects of Year. A significant main effect of Year-group was found in relation to all scales apart from L2 Self-Confidence.

Table 7. Descriptive statistics for Years 1, 3 and 5 (significantly highest scores in bold and lowest shaded)

Scale		1	2	3	4	5	6	7	8
		<i>Ideal</i>	<i>Ought</i>	<i>Behaviour</i>	<i>Motivation</i>	<i>Confidence</i>	<i>English</i>	<i>Subject</i>	<i>E-t-S</i>
1 st year	M	3.86	1.88	3.65	3.57	3.79	3.55	4.07	3.89
	N	870	870	870	870	870	870	848	848
	SD	0.57	0.60	0.58	0.72	0.53	0.68	0.67	0.63
3 rd year	M	4.06	2.03	3.30	3.45	3.79	2.92	3.95	3.72
	N	631	631	631	631	631	631	620	620
	SD	0.55	0.67	0.61	0.78	0.56	0.76	0.76	0.68
5 th year	M	4.08	2.16	3.07	3.37	3.78	3.09	-	-
	N	394	394	394	394	394	394	-	-
	SD	0.58	0.63	0.63	0.72	0.57	0.73	-	-
Total	M	3.97	1.99	3.41	3.49	3.79	3.24	4.02	3.82
	N	1895	1895	1895	1895	1895	1895	1468	1468
	SD	0.58	0.64	0.64	0.74	0.55	0.78	0.71	0.66

Table 8. Results of 8 mixed model analyses for main effects of Year, with School as random intercept, including results of EM means posthoc tests for Scales 1-6

Scale	Random slope(s)	df	F	p	EM means	
					years	p
1. Ideal L2 Self	-	7	36.361	<.001**	1-3	<.001**
					1-5	<.001**
					3-5	.308
2. Ought-to L2 Self	year	8	11.822	<.001**	1-3	.006*
					1-5	<.001**
					3-5	.074
3. Motivated Learning Behaviour	year	8	52.325	<.001**	1-3	<.001**
					1-5	<.001**
					3-5	.006*
4. Motivation for English	-	7	9.483	<.001**	1-3	.002*
					1-5	<.001**
					3-5	.212
5. L2 Self-Confidence	-	7	.034	.967	-	-
6. English Lesson Experience	track	8	111.943	<.001**	1-3	<.001**
					1-5	<.001**
					3-5	.051
7. Subject Learning Experience	-	6	6.845	.009*	-	-

Scale	Random slope(s)	df	F	p	EM means	
					years	p
8. English-Through-Subject Learning Experience	year	7	20.077	<.001**	-	

* = significant at $p < .05$. ** $p < .001$

For the Ideal ($F(7, 1573.98) = 36.361, p < .001$) and Ought-to L2 Self ($F(8, 21,28) = 11.822, p < .001$) (Scales 1-2), fifth-year respondents produced the highest score and first-year the lowest. The EM means posthoc tests show that the difference between first and third year (Ideal: $p < .001$; Ought: $p = .006$), and between first and fifth year (both: $p < .001$) were significant for both of these scales. The difference between third year and fifth year was not significant for either of these scales.

For Motivated Learning Behaviour ($F(8, 18.33) = 52.325, p < .001$) and Motivation for English ($F(7, 1583.40) = 9.483, p < .001$) (Scales 3-4), first-year respondents scored highest and fifth-year respondents scored lowest. Posthoc analyses highlight significant differences between first and third year (Behaviour: $p < .001$; Motivation: $p = .002$), and between first and fifth year for both of these scales (both $p < .001$). For Motivated Learning Behaviour, the difference between third and fifth year was also significant ($p = .006$), while for Motivation for English it was not.

With regard to the learning experience (Scales 6-8), first-year respondents gave the highest and third years the lowest scores for all three scales. For the English Lesson Experience (Scale 6) ($F(8, 1791.44) = 111.943, p < .001$), third-year participants gave the lowest score. The differences between first and third year ($p < .001$), and between first and fifth year were significant ($p < .001$), while the difference between third and fifth year was not. Posthoc tests were not performed for Scales 7 ($F(6, 1306.29) = 6.845, p = .009$) and 8 ($F(7, 1201.82) = 20.077, p < .001$) as they only included responses from first and third-years.

7. Discussion

Analysis of the quantitative Likert scale data on the basis of Gender, Educational Track and Year-group, and supporting respondent comments, suggest that trends in L2 motivation in bilingual secondary education (BSE) are not clear-cut. For each of these independent variables, there was no group that consistently expressed stronger motivation, suggesting that the most pertinent question is that of motivational *difference* rather than motivational *advantage*. That said, a number of new insights into the nature of these differences do emerge, awareness of which might help us to better understand and serve the needs of different learner groups within BSE or similar approaches. Here, we discuss and interpret the findings regarding differences between boys and girls, between educational tracks and between year-groups, before drawing conclusions regarding implications for research, policy and practice.

7.1 Gender difference, not gender gap

Boys and girls in this study self-reported similar levels of Motivation for English and of Motivated Learning Behaviour, although differences were observed in terms of the factors that might contribute to that motivation, namely aspects of the L2 self, L2 self-confidence and the learning experience.

It appeared that girls were more conscious of their Ideal L2 self, the vividness of which, according to Dörnyei (2009), could have a powerful motivating effect. In contrast, boys in the current study expressed more awareness of the Ought-to L2 self. The latter has been echoed elsewhere (Azarnoosh & Birjandi, 2012; Heras & Lasagabaster, 2014), where it was suggested that the difference between boys and girls in this respect may be due to boys being placed under more societal pressure to perform, although previous research in BSE did not highlight the same gender difference in this respect (Mearns & de Graaff, 2018b). It must be noted, however, that both the relatively low scores for this scale across all groups and the

dismissive comments from respondents raise questions as to the applicability of the Ought-to L2 self to survey research in the Dutch BSE setting. This may reflect the suspicion voiced elsewhere that awareness of external expectations does not necessarily lead to internalisation of those expectations (Al-Hoorie, 2018).

The fact that there were no significant differences in boys' and girls' Motivated Learning Behaviour or Motivation for English might support the suggestion in the CLIL literature that content-focused approaches might help to narrow the gender gap in terms of L2 motivation (Lasagabaster, 2008). An alternative interpretation in light of earlier findings that suggest inherent differences between boys in BSE and mainstream Dutch education (Mearns & de Graaff, 2018b), however, might be that boys who opt for BSE are naturally motivated learners of English, independently of their learning experiences. Noteworthy in this regard is that the only significant gender difference regarding the learning experience was in the higher rating assigned to English lessons by girls. This could suggest that approaches to the English curriculum in BSE schools favours girls, or at least that it does little to negate societal influences on boys' motivation to learn languages as a distinct discipline (Kissau, 2006), in spite of the stronger Ought-to L2 Self among boys with regard to language proficiency.

Furthermore, boys scored higher than girls with regard to L2 Self-Confidence, while it has been suggested elsewhere that traditionally female-friendly language classrooms have been posited to adversely affect boys' linguistic self-concept (Mills, 2014). This could imply that boys are indeed more attracted by the applied nature of language learning in BSE, as was suggested by Lasagabaster (2008) in a Spanish CLIL context, or that the inclusion of appealing subject content makes the language learning experience more enjoyable and more comfortable for boys than traditional language lessons (Coyle et al., 2010). The quotation cited in the Results section, in which a boy expresses strong negative feelings regarding "English lessons" but enjoyment of "lessons in English," could be an indication of this trend.

Somers and Llinares (2018) distinguish between L2 motivation and “CLIL motivation” (p. 1), whereby the latter refers to learners’ motivation specifically for learning language and content in an integrated way. The current results seem to indicate that CLIL motivation and even L2 motivation may exist among boys, where motivation for formal English lessons may not.

7.2 Differentiation beyond organisation

We drew attention already to the fact that the HG track in Dutch secondary education – including BSE – lacks a distinct identity in spite of evidence to suggest differences between these groups (Verspoor et al., 2015). Differences between responses from HG and Pre-A in the current study indeed suggest that each group may have its own L2 motivational profile. HG respondents self-reported higher levels of Motivation for English, Motivated Learning Behaviour and L2 Self-Confidence than Pre-A respondents, while Pre-A displayed a stronger sense of the Ideal L2 Self and were more positive regarding the Subject Learning Experience. In previous research focusing on HG pupils, Mearns and colleagues (Mearns, 2015; Mearns & de Graaff, 2018b; Mearns et al., 2017) found significant differences between the motivations and orientations of learners in BSE and mainstream groups. Those in BSE were likely to have chosen this route due to a desire to be challenged, enjoyment of the English language or an international outlook, and they also appeared to be more oriented towards learning in general. This, combined with the current findings, might suggest that HG BSE is an opportunity for pupils who are not academically-inclined to nevertheless take on a higher level of challenge. They may not be suited to the academically-oriented Pre-A, but can be enthusiastic and confident learners and users of English.

The more positive response from Pre-A regarding the Subject Learning Experience, however, could imply that teaching approaches in BSE are less well-suited to the HG learner. The BSE Standard differentiates between HG and Pre-A with regard to organisational rather than pedagogical features of their programme. That Pre-A learners were more positive about

an aspect of the learning experience in spite of the fact that HG learners expressed stronger motivation could be an indication that the way in which bilingual subjects are taught at HG level is less fitting to the type of pupil in the class. As there were no significant differences between the HG and Pre-A groups regarding the other two aspects of the learning experience and no additional insights were provided by their comments, however, this finding is tentative.

7.3 Shifting priorities?

In terms of the differences across year-groups, the oldest and youngest pupils displayed motivation in different areas. While the youngest respondents expressed stronger Motivation for English and Motivated Learning Behaviour, and were more positive regarding the learning experience, fifth-year participants responded more positively with regard to the L2 Self. In other words, while younger learners appeared to focus on their enjoyment of English and their positive learning experience, older learners were more aware of self-oriented motivators. This could be an indication that future self-guides (Dörnyei, 2009) gained significance as learners grew older. This was reflected in the additional comments, in which older respondents described more vivid ideal L2 selves than did younger learners. While cross-sectional data cannot confirm this motivational shift from focus on enjoyment to focus on future prospects, the evidence in its favour begs further research and could indicate that younger and older learners' motivation can be nurtured and stimulated in different ways.

The higher level of motivation and enjoyment expressed by younger learners reflects common trends in L2 motivation, which is known to decline among older learners, even in bilingual education or CLIL (Lasagabaster & Sierra, 2009). Dalton-Puffer and Smit (2013) posit that one of the main motivators of CLIL-type contexts may be the novelty of the teaching approaches, so it is not surprising that the BSE learning experience may be experienced as less motivating once that novelty has worn off. While this does not necessarily

imply a demotivating effect of BSE, it also does not provide evidence of a motivating effect. Lamb (2017) highlights that one of the irrefutable lessons research into L2 motivation has taught us is that “teachers are able to influence their learners’ motivation, both for better and worse” (p. 330). This is echoed in the emphasis on the role of teachers in the additional comments in the current study. As CLIL is dependent on effective and safe classroom interaction (Coyle et al., 2010), and CLIL is a central feature of BSE, it follows that pupil-teacher relationships might have a particularly strong impact on L2 motivation in BSE.

An alternative explanation for the less positive response from older learners might be the curriculum itself. Worth noting are the comments by some fifth-year respondents regarding the transition from English to Dutch in the final years of BSE. While the small number of comments that refer to this explicitly means that no firm conclusions can be drawn, they do remind us of the risk that the return to Dutch for the majority of the senior BSE curriculum might cause learners to lose enthusiasm. This effect was confirmed by Denman et al. (2018) in the context of Pre-V BSE (see also Rumlich, 2018).

Also with regard to curriculum, it is interesting to note that the only scale that does not show a continuous incline or decline from first to fifth year is the English Lesson Experience. For this scale, a significant decline was observed between first-year and third-year responses, followed by an incline (albeit non-significant) between third year and fifth year. As fifth-year data pertaining to other aspects of the learning experience are not included in the analysis, it is not possible to know whether this is a particular response to English lessons or to all English-medium lessons. That said, it is possible that this could be a reflection of the curriculum for English. While some BSE English departments introduce a heavier focus on literature, cultural and media studies in the junior years as preparation for the International Baccalaureate, many other schools continue to implement a traditional form-focused language curriculum in the first three years before switching to an almost entirely content-focused

programme in the senior years. Variation between schools in this regard is substantial (Dale et al., 2018). The dip in positivity regarding English lessons in third year, and indeed the tendency for comments regarding English lessons to be negative, might therefore be a reflection of learners' readiness to move on from traditional language lessons and embrace a more content-focused approach, as they are accustomed to in their other English-medium lessons.

8. Limitations of this study

Being cross-sectional in nature, the current study only allows for tentative conclusions regarding motivational changes over time. As a quantitative study, it also does not take into account differences between individuals or the dynamic nature of motivation (de Bot et al., 2007). Longitudinal or in-depth qualitative studies of learner motivations and experiences of BSE would be a welcome complement to larger-scale quantitative studies such as this one.

Furthermore, this study does not take into account the more recent concept of CLIL motivation (Somers & Llinares, 2018), focusing instead on L2 motivation in BSE. It may be valuable to incorporate CLIL motivation in future studies of motivation in BSE, in order to investigate this concept in a different bilingual education setting, and to add further depth to the findings.

9. Conclusions and implications for practice and research

In light of the complexity and context-dependency of L2 motivation, the current study sought to investigate the L2 motivations of learners in the specific educational setting of Dutch bilingual secondary education (BSE). As indicated by previous studies (Mearns et al., 2017; Sieben & van Ginderen, 2014), learners in BSE are likely to be inherently different to learners who do not opt for a bilingual programme. The aim was, therefore, not to evaluate the motivational outcomes of BSE compared to mainstream Dutch-medium programmes, but to examine the motivational makeup of the learners within BSE by identifying trends in L2

motivation among learners of different genders, educational tracks and year-groups. It was hoped that greater understanding of these groups' motivations would help to inform developments and differentiation in policy and practice, in the Netherlands as well as in other CLIL or bilingual education settings.

The results of the study draw attention to differences between particular learner groups within BSE, highlighting that the L2 motivations of girls and boys, and of the general and pre-academic tracks may be qualitatively different and therefore influenced by different aspects of the learning experience. Noteworthy in these respects were the higher rating given to English lessons by girls, and the lower rating for subject learning by the less academic group, both of which highlight the need to accommodate different types of learners in BSE just as in mainstream programmes, not only on organisational levels but also in terms of curriculum and pedagogy. In addition, the seemingly higher level of L2 motivation reported by the learners in the less academic track serves as a reminder that learners who choose BSE often do so due to high levels of motivation, regardless of their academic inclinations.

Differences between year-groups suggest too that there may be opportunities to motivate learners in different stages of BSE in different ways, for example by making use of the more prominent L2-related ambitions of older learners and the initial enthusiasm of younger learners, but also serve as a warning that BSE does not necessarily appear to maintain a motivating effect over time. The relative negativity of third-year respondents with regard to English lessons is an area that warrants further investigation, for example to establish whether there is a connection with the curriculum.

Dörnyei highlighted over two decades ago that “the exact nature of the social and pragmatic dimensions of L2 motivation is always dependent on who learns what languages where” (1998, p. 275). In the light of the current study, it would appear that this statement is no less applicable in the context of bilingual education. In choosing BSE, learners such as

those in the current study have set themselves apart as motivated language learners, but this is not to say that they all have the same needs. Mearns (2015) proposed that education – bilingual or otherwise – might adopt a concept of “Motivational differentiation” (p. 305), taking into account varying learner needs not only in terms of interests, academic readiness and learning preferences (Tomlinson et al., 2003), but also with regard to what drives them to learn and keeps them learning in the longer term (Dörnyei, 2001). Future research could shed light on what forms these different needs might take and how they can best be served in practice. In the meantime, the results of the current study highlight that the BSE and broader CLIL communities may benefit from acknowledging learners’ motivational differences more overtly in order to ensure that they live up to their potential to enthuse and motivate learners.

References

- Al-Hoorie, A. H. (2018). The L2 Motivational Self System: A meta-analysis. *Studies in Second Language Learning and Teaching*, 8(4), 721-754.
<https://doi.org/10.14746/ssllt.2018.8.4.2>
- Azarnoosh, M., & Birjandi, P. (2012). Junior High School Students’ L2 Motivational Self System: Any Gender Differences? *World Applied Sciences Journal*, 20(4), 577-584.
- Boo, Z., Dörnyei, Z., & Ryan, S. (2015). L2 motivation research 2005e2014: Understanding a publication surge and a changing landscape. *System*, 55, 145-157.
<https://doi.org/10.1016/j.system.2015.10.006>
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 7-21.
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System*, 39(4), 523-532. <https://doi.org/10.1016/j.system.2011.08.002>

- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge University Press.
- Csizér, K., & Dörnyei, Z. (2005). The Internal Structure of Language Learning Motivation and Its Relationship with Language Choice and Learning Effort. *The Modern Language Journal*, 89(1), 19-36. <https://doi.org/10.1111/j.0026-7902.2005.00263.x>
- Dale, L., Oostdam, R., & Verspoor, M. (2018). Juggling ideals and constraints. The position of English teachers in CLIL contexts. *Dutch Journal of Applied Linguistics*, 7(2), 177-202. <https://doi.org/10.1075/dujal.18002.dal>
- Dalton-Puffer, C., & Smit, U. (2013). Content and Language Integrated Learning: A research agenda. *Language Teaching*, 46(04), 545-559. <https://doi.org/10.2167/beb459.0>
- Denman, J., van Schooten, E., & de Graaff, R. (2018). Attitudinal factors and the intention to learn English in pre-vocational secondary bilingual and mainstream education. *Dutch Journal of Applied Linguistics*, 7(2), 203-226. <https://doi.org/10.1075/dujal.18005.den>
- 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. <https://doi.org/10.1080/09571736.2014.889508>
- Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, 31(03), 117-135. <https://doi.org/10.1017/S026144480001315X>
- Dörnyei, Z. (2001). *Motivational Strategies in the Language Classroom*. Cambridge University Press.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing* (1st ed.). Lawrence Erlbaum.
- Dörnyei, Z. (2009). The L2 Motivational Self System. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (pp. 9-42). Multilingual Matters.

- Dörnyei, Z., & Chan, L. (2013). Motivation and Vision: An Analysis of Future L2 Self Images, Sensory Styles, and Imagery Capacity Across Two Target Languages. *Language Learning*, 63(3), 437-462. <https://doi.org/10.1111/lang.12005>
- Dörnyei, Z., & Ushioda, E. (2009). *Motivation, Language Identity and the L2 Self*. Multilingual Matters.
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and Researching: Motivation* (Vol. 2). Pearson Education.
- Genesee, F. (2007). Top Ten Most Consistent Findings from Research on Foreign Language Immersion. *The ACIE Newsletter*, 10(3).
https://carla.umnn.edu/immersion/acie/vol10/may2007_researchfindings.html
- Goeman, J. J., & de Jong, N. H. (2018). How well does the sum score summarize the test? Summability as a measure of internal consistency. *Educational Measurement: Issues and Practice*, 37(2), 54-63. <https://doi.org/10.1111/emip.12181>
- Hadfield, J., & Dörnyei, Z. (2013). *Motivating learning*. Longman.
- Henry, A. (2009). Gender differences in compulsory school pupils' L2 self-concepts: A longitudinal study. *System*, 37, 177-193. <https://doi.org/10.1016/j.system.2008.11.003>
- Henry, A., & Cliffordson, C. (2013). Motivation, Gender, and Possible Selves. *Language Learning*, 63(2), 271-295. <https://doi.org/10.1111/lang.12009>
- Heras, A., & Lasagabaster, D. (2014). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19(1), 70-88.
<https://doi.org/10.1177/1362168814541736>
- IRIS (2020, 15 September) *IRIS: A digital repository of instruments and materials for research into second languages*. IRIS. <https://www.iris-database.org/iris/app/home/index>

- Kissau, S. (2006). Gender differences in second language motivation: An investigation of micro- and macro-level influences. *Canadian Journal of Applied Linguistics*, 9(1), 73-96. <https://journals.lib.unb.ca/index.php/CJAL/article/view/19755>
- Lamb, M. (2012). A Self System Perspective on Young Adolescents? Motivation to Learn English in Urban and Rural Settings. *Language Learning*.
<https://doi.org/10.1111/j.1467-9922.2012.00719.x>
- Lamb, M. (2017). The motivational dimension of language teaching. *Language Teaching*, 50(3), 301-346. <https://doi.org/10.1017/S0261444817000088>
- Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Courses. *The Open Applied Linguistics Journal*, 1, 30-41.
<https://doi.org/10.2174/1874913500801010030>
- Lasagabaster, D., & Sierra, J. M. (2009). Language Attitudes in CLIL and Traditional EFL Classes. *International CLIL Research Journal*, 1(2), 4-17.
<http://www.icrj.eu/index.php?vol=12&page=73>.
- Marsden, E., Mackey A., & Plonsky, L. (2016). The IRIS Repository: Advancing research practice and methodology. In A. Mackey & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS Repository of Instruments for Research into Second Languages* (pp. 1-21). Routledge.
- Mearns, T. (2015). *Chicken, Egg or a Bit of Both? Motivation in bilingual education (TTO) in the Netherlands*. (Unpublished doctoral dissertation) University of Aberdeen, Aberdeen. <http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.646095?>
- Mearns, T., & de Graaff, R. (2018a). Bilingual education and CLIL in the Netherlands. The paradigm and the pedagogy. *Dutch Journal of Applied Linguistics*, 7(2), 122–128.
<https://doi.org/10.1075/dujal.00002.int>

- Mearns, T., & de Graaff, R. (2018b). Bucking the trend? Motivational differences between boys and girls who opt in or out of bilingual education. *Journal of Immersion and Content-Based Language Education*, 6(1), 1-26.
<https://doi.org/10.1075/jicb.17003.mea>
- Mearns, T., de Graaff, R., & Coyle, D. (2017). Motivation for or from bilingual education? A comparative study of learner views in the Netherlands.
<https://doi.org/10.1080/13670050.2017.1405906>
- Merisuo-Storm, T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. *Teaching and Teacher Education*, 23(2), 226-235. <https://doi.org/10.1016/j.tate.2006.04.024>
- Michels, B. (2006). *Verschil moet er wezen. Een werkdocument over verschillen tussen havo en vwo-leerlingen in de tweede fase en handreikingen om daarmee om te gaan.*
https://www.slo.nl/downloads/archief/verschil_20moet_20er_20wezen.pdf/
- Mills, N. (2014). Self-Efficacy in Second Language Acquisition. In S. Mercer & M. Williams (Eds.), *Multiple Perspectives on the Self in SLA* (pp. 6-22). Multilingual Matters.
- Nuffic. (2012). *Standard for bilingual education*. Nuffic.
- Nuffic. (2019). *Kwaliteitsstandaard tweetalig onderwijs 2.0*. Nuffic.
<https://www.nuffic.nl/publicaties/kwaliteitsstandaard-tweetalig-onderwijs-20/>
- Papi, M., & Abdollahzadeh, E. (2012). Teacher Motivational Practice, Student Motivation, and Possible L2 Selves: An Examination in the Iranian EFL Context. *Language Learning*, 62(2), 571-594. <https://doi.org/10.1111/j.1467-9922.2011.00632.x>
- Preston, C.C., & Colman, A.M. (2000) Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1-15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5)

- Rumlich, D. (2018). Current research on CLIL and bilingual education in the Netherlands. A discussion. *Dutch Journal of Applied Linguistics*, 7(2), 264-273.
<https://doi.org/doi.org/10.1075/dujal.00003.rum>
- Ryan, S. (2009). Self and Identity in L2 Motivation in Japan: The Ideal L2 Self and Japanese Learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (pp. 120-143). Multilingual Matters.
- Sieben, I., & van Ginderen, N. (2014). De keuze voor tweetalig onderwijs. *Mens en maatschappij*, 89(3), 233-255. <https://doi.org/10.5117/MEM2014.3.SIEB>
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.
- Somers, T., & Llinares, A. (2018). Students' motivation for content and language integrated learning and the role of programme intensity. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2018.1517722>
- Stichting Havoplatfom. (2019). *Visie*. <https://havoplatfom.nl/visie/>.
- Sylvén, L. K. (2013). CLIL in Sweden – why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism*, 16(3), 301-320. <https://doi.org/10.1080/13670050.2013.777387>
- 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. <https://doi.org/10.1075/jicb.3.1.02sylv>
- Thompson, A. S., & Sylvén, L. K. (2019). CLIL and Motivation Revisited: A Longitudinal Perspective. In L. K. Sylvén (Ed.), *Investigating Content and Language Integrated Learning. Insights from Swedish High Schools* (pp. 76-97). Multilingual Matters.
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., Reynolds, T. (2003). Differentiating Instruction in Response to

Student Readiness, Interest, and Learning Profile in Academically Diverse Classrooms: A Review of Literature. *Journal for the Education of the Gifted*, 27(2/3), 119-145. <https://doi.org/10.1177/016235320302700203>

Verspoor, M., de Bot, K., & Xu, X. (2015). The effects of English bilingual education in the Netherlands. *Journal of Immersion and Content-Based Language Education*, 3(1), 4-27. <https://doi.org/10.1075/jicb.3.1.01ver>

Abstract (Nederlands)

In veel onderzoek naar motivatie in tweetalige onderwijscontexten staan verschillen tussen leerlingen in het tweetalige of het reguliere onderwijs centraal. We worden echter steeds bewuster van het feit dat er tussen deze twee groepen leerlingen inherente verschillen zijn, bijvoorbeeld omdat tweetalige programma's vaak facultatief of selectief zijn (Mearns et al., 2017). Het doel van dit onderzoek was daarom niet om leerlingen in het tweetalige en het reguliere onderwijs met elkaar te vergelijken, maar om de motivatie van leerlingen te verkennen binnen het tweetalig voortgezet onderwijs (tto) in Nederland. Met een vragenlijst die speciaal ontworpen is om aspecten van het L2 Motivational Self System (Dörnyei, 2009) te peilen, onderzoeken we verschillen in motivatie tussen leerlingen op basis van gender, onderwijsniveau (havo of vwo) en leerjaar. De bevindingen suggereren dat – ondanks de gezamenlijke keuze voor het tto – de verschillen tussen deze groepen leerlingen met betrekking tot hun tto-ervaring niet te onderschatten zijn.

Author addresses

Tessa Mearns
ICLON Graduate School of Teaching
Leiden University
Kolffpad 1
2333 BN Leiden
The Netherlands
t.l.mearns@iclon.leidenuniv.nl

Nivja de Jong
ICLON Graduate School of Teaching
Leiden University
Kolffpad 1
2333 BN Leiden
The Netherlands