Attitudinal factors and the intention to learn English in pre-vocational secondary bilingual and mainstream education*

Jenny Denman[1], Erik van Schooten[1], and Rick de Graaff[2]
[1] Rotterdam University of Applied Sciences | [2] Utrecht University

The effect of bilingual education (BE) on the attitude towards learning English of pupils in the first three years of pre-vocational secondary education in the Netherlands ($n = 488$) was investigated. Contrary to several other BE/CLIL studies, in the present study pupils choosing for a bilingual stream are not preselected based on their attitude or motivation. Attitude was measured using the Model of Planned Behavior (MPB). The best-fitting attitude model was one in which the perceived importance of learning English was a direct predictor of the Intention to make an effort, and not mediated by Affect. At the start of BE in year 1 (age 12), attitudinal differences between bilingual and mainstream pupils were non-significant, but after one or more years of a bilingual program, BE pupils scored significantly higher on four of the five MPB attitudinal constructs. BE appears to positively influence the attitude towards learning English of junior vocational students.

Keywords: Content and Language Integrated Learning, CLIL, junior vocational secondary education, vmbo, attitude, affective factors, Model of Planned Behavior, MPB, bilingual education

1. Introduction

Research from a wide variety of countries and contexts has shown that pupils in some form of bilingual educational programs exhibit more positive attitudes and stronger motivation towards learning English than those in monolingual streams (Doiz, Lasagabaster, & Sierra, 2014; Mearns, De Graaff, & Coyle, 2017; Merisuo-Storm, 2011; Sultan, Borland, & Eckersley, 2012; Sylvén & Thompson, 2015; Verspoor, De Bot, & Xu, 2015). These studies have focused almost exclusively

* Online appendix available from https://doi.org/10.1075/dujal.18005.den.appendix

https://doi.org/10.1075/dujal.18005.den
Dutch Journal of Applied Linguistics 7:2 (2018), pp. 203–226. issn 2211-7245 | e-issn 2211-7253

on selected, higher-achieving pupil populations (Bruton, 2015; Rumlich, 2014), and the results may not be generalizable to a non-selective, lower-achieving population. Furthermore, several studies have shown that pupils enter bilingual education with a more positive attitude towards learning English than pupils in regular educational programs (Mearns, 2015; Rumlich, 2014), though this has not yet been shown for the pre-vocational streams. The question is whether the more positive attitudes found in these studies can also be found – and maintained over time – in a non-selective, less academically-skilled population. Such a population can be found in the Netherlands because of the highly-streamed nature of secondary education and the existence of bilingual education (BE) programs with a Content and Language Integrated Learning (CLIL) pedagogical approach for all streams. The least academic stream in Dutch secondary education is called *vmbo,* or pre-vocational secondary, and is further substreamed into four cognitive ability levels. In total, approximately 50% of the secondary pupil population attends a four-year pre-vocational secondary stream. The four substreams thus represent 'average' to 'far below average' for academic ability. Once the pupil's stream/substream has been determined, pupils and their parents are free to choose any school in that stream, as well as choose for either a bilingual education program (hereafter called 'BE program' and 'BE pupils') or a regular Dutch-language program ('non-BE').

BE programs in the Netherlands date from 1989, but until recently these were limited to the university-track and general secondary schools. In 2009 the first pre-vocational school started with BE and at the time of this research there were 27 pre-vocational secondary schools with a BE program, none of which yet had national BE accreditation. Nearly all of these offer English-Dutch BE (De Graaff & Van Wilgenburg, 2015), which is the focus of the research reported here.

The Dutch pre-vocational BE approach is set out in the national Standard for bilingual pre-vocational education (Nuffic, "Standaard tweetalig vmbo"). It entails a less intensive bilingual program than the higher streams, with subjects taught using CLIL pedagogy and English comprising approximately 30% of the curriculum. It is pertinent that there is a higher percentage of English-medium provision in the first two years (age 12–14), and a smaller amount in the last two years, when the pupils receive regular Dutch-language instruction for most of the previously English-medium subjects (age 14–16) since all final national examinations are in Dutch. The pupil population is characterized by a higher percentage of first- and second-generation immigrants than in higher streams, particularly in the urban regions (Hartgers, 2007), and the lower the pre-vocational stream, the higher the percentage of non-Western immigrants and a related higher prevalence of home languages other than Dutch (Centraal Bureau voor de Statistiek, 2014). Clearly this population differs from the academically-inclined population com-

prising most CLIL/BE programs and research. If BE has a positive effect on pupils' attitude towards learning English, it could be even more beneficial for this lower-achieving group. Furthermore, since secondary school pupils' attitudes towards learning English (Cenoz, 2001) – or towards school in general (Lasagabaster & Sierra, 2009; Pels, Jonkman, & Drost, 2011; Van Nuland, 2011) – tend to become more negative over time, it is relevant to explore whether bilingual pre-vocational programs have a mitigating effect on this downward trend.

In a review of CLIL research in Europe, Pérez Cañado (2012) calls for studies which address "attitudinal and affective factors" and use "solid empirical evidence" (2012, p. 330). Attitudes, as defined by Keil in the International Encyclopedia of Education (1991), are positive or negative feelings of an individual towards objects, persons or ideas, and thought to be learned, changeable, and to influence behavior. Attitude theory has developed into an attitude model, the Model of Planned Behavior (MPB) (Ajzen, 2005; Ajzen & Fishbein, 1980), validated in hundreds of studies in fields such as behavioral psychology (Armitage & Connor, 2001; Conner & Sparks, 2005; Godin & Kok, 1996;) and education (Davis, Ajzen, Saunders, & Williams, 2002; Sideridis & Padeliadu, 2001; Van Schooten, De Glopper, & Stoel, 2004;) and has shown strong predictive validity (Ajzen, 1991). In the MPB, Intention (in this study the intention to make an effort to learn English, such as paying attention, actively participating in class, doing homework, asking questions, heeding corrective feedback, etc.), causes the actual behavior (not measured in this study), and is in turn determined by three constructs: Affect, Subjective Norm, and Perceived Behavioral Control (PBC). Affect represents the person's feelings about performing the behavior and is thought to be caused by Cognition, the evaluation of the expected results of performing the behavior. Subjective Norm stands for the person's own moral ideas about the desirability of the behavior, and PBC is the degree to which the person perceives impediments to performing the behavior (Ajzen, 2002). The MPB and its postulated causal structure are shown in Figure 1.

The literature also mentions two other models, so it is relevant to consider these. Some researchers (Ajzen, 2005; Bagozzi & Burnkrant, 1980; Dillon & Kumar, 1985) suggest that for some domains, Affect and Cognition might be one and the same construct. Other studies show that a model in which Affect and Cognition each have a direct path to Intention fits better (Bagozzi & Burnkrant, 1979, 1985; Brinberg, 1981; Fen & Sabaruddin, 2008; Triandis, 1977; Valois, Desharnais & Godin, 1988).

The first question is which of these three aforementioned models is most suitable for measuring BE and non-BE pupils' attitudes towards learning English. Using the best fitting model, we can answer the following questions: how strong the relationships are between the various constructs of the MPB, and which constructs of the MPB show the strongest relationships with the Intention to learn

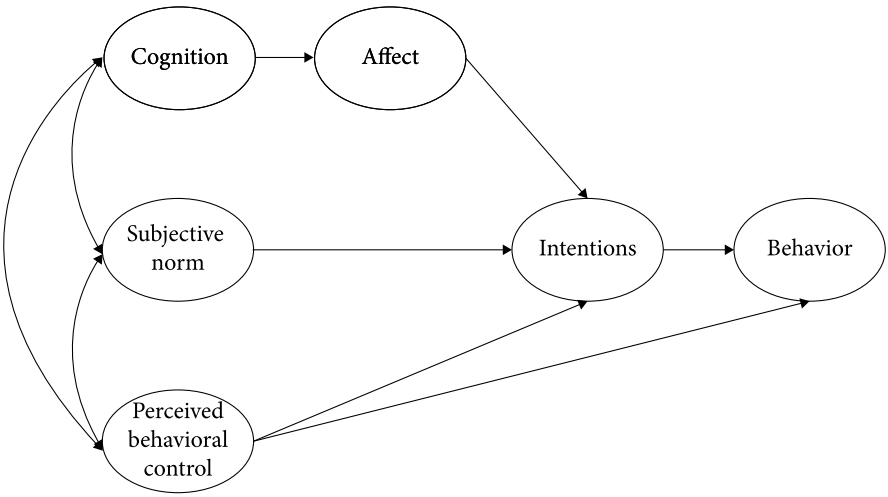

Figure 1. MPB as operationalized by Van Schooten et al., 2004

English. Subsequently, we also want to compare scores of BE and non-BE pupils on the MPB constructs in order to verify whether BE at this pre-vocational level is beneficial for the attitude towards making an effort to learn English. Before this comparison however, it is crucial to verify whether the designed instrument measures the same constructs on the same scale for the BE and the non-BE pupils. (Koomen, Verschueren, Van Schooten, Jak, & Pianta, 2012).

If BE changes the way in which pupils perceive the meaning of the MPB questionnaire items, the relationships between item-scores and latent traits might change, resulting in a difference in the constructs measured in both groups. In research the assumption is often made that a test measures the same constructs in different groups (metric invariance) and/or on the same scale in different groups (scalar invariance); in other words, is measurement invariant (Cheung & Rensvold, 2002; Meredith, 1993; Meredith & Teresi, 2006). The results of a comparison may be invalid if measurement invariance does not hold (Vandenberg & Lance, 2000); however, if both metric and scalar invariance for the BE and non-BE pupils hold, then we have 'strong factorial invariance' (Meredith & Teresi, 2006). If the measurement instrument is measurement invariant, then we will be able to explore differences in MPB scores between BE and non-BE pupils and determine whether the BE pupils in this study score higher on the MPB constructs than the non-BE pupils, even after controlling for other individual characteristics such as gender and language background.

Recalling the more positive attitude of more academically-inclined BE pupils in the previously-mentioned research results, we expect a causal effect of BE on the MPB constructs (higher for the BE pupils). Given that this cross-sectional

measurement includes three different class years, we predict that if there is a positive causal effect of BE on the different MPB constructs, we will find higher scores on these constructs for BE pupils. We also expect that the longer pupils follow a BE program, the larger the differences between BE and non-BE students will be. Prior research has shown that BE pupils in more academic streams have a more favorable attitude towards English from the outset than their non-BE peers (Mearns, 2015; Rumlich, 2014), likely caused by a twofold selection: the BE stream selecting pupils with a positive attitude, and by positive pupils also selecting BE. The present study examines if this is also the case for our non-selective, non-elite stream sample: whether pupils at the beginning of pre-vocational BE also have a more positive attitude towards learning English than their non-BE counterparts. If we do not find this difference at the start of BE in our sample (year 1), but do find it in year 2, a causal interpretation of the positive effects of BE on pupil attitude is more plausible.

The final research questions are, therefore, whether the hypothesis holds that BE pupils score higher on the MPB constructs, and whether the difference between BE and non-BE pupils is larger in year 2 than in year 1. For differences in year 3 we had no clear expectations, as the intensity of the BE program diminishes starting in the third year. Because of the cross-sectional design, differences in scores between BE and non-BE pupils could be caused by factors other than BE participation. Therefore differences between BE and non-BE pupils on MPB scores were also tested after correcting for some pupil characteristics.

The research questions for this article, then, are:

1. Which of the three theoretical MPB models fits better, and to what extent?
2. How strong are the relationships between the intention to learn English and the other constructs of the MPB in the Dutch pre-vocational context?
3. To what degree is the best fitting model, and thus the instrument based on the MPB, measurement invariant for the BE/non-BE groups? And when measurement invariance holds,
4. Do these BE pupils score higher on the MPB components than non-BE pupils?
5. How large are the differences in scores for BE and non-BE pupils on the MPB components in years 1, 2, and 3 and do these differences increase as pupils gain more BE experience?

2. Method

2.1 Participants

All 27 Dutch pre-vocational secondary schools with a BE program were invited to participate and all six responding schools were included. None of these schools select pupils for BE based on previous motivation; four of the schools had no selection criteria whatsoever and two administered an English placement test and a brief interview,[1] so the only true selection criteria is pupil self-selection There is only minimal attrition from the BE streams: schools report that once first-year pupils choose BE, they continue for the 4-year duration of pre-vocational secondary education and almost no pupils drop out of the BE program. Three schools were BE-only, with no non-BE streams; these schools provided BE classes for the experimental group. The other three schools had parallel BE and non-BE streams; these three schools provided both BE (experimental group) and non-BE (control group) classes. One additional non-BE school was included in the control group. Twenty-five entire class groups of pupils from the seven schools participated (242 boys and 278 girls, mean age 13 years and 3.2 months). Each school had between two and six classes participating, for a total of 15 BE classes ($n = 293$) and 10 non-BE classes ($n = 242$). BE/non-BE class size was approximately the same; some BE pupils were logistically unable to participate in the research because of a slightly different vocational profile. The participants were in year 1 ($n = 251$), 2 ($n = 168$) and 3 ($n = 116$).

Although our sample was non-randomly selected, it does not seem very biased. Firstly, the average CITO score (a standardized national primary education exit test), the main tool used to stream pupils into the pre-vocational secondary levels, was within the range of the national CITO ranges for 24 of the 25 classes. Secondly, the different BE organizational approaches found in the experimental group schools reflect the different BE approaches found nationally. Thirdly, the percentage of pre-vocational secondary pupils with a non-Dutch parental background is 30.7% in our study, and 33.6% nationally (Centraal Bureau voor de Statistiek, 2016). All of the schools in our sample are located in the urban west, the most densely populated area of the country, where also the majority of pre-vocational secondary BE programs can be found.

There was no exclusion of participants on the basis of previous scores, learning difficulties, or behavioral problems. Interviews with the BE coordinators, teachers, and focus groups of pupils confirmed the low threshold for participating

1. The BE coordinators at these schools reported that in general, this is a formality, and in practice almost no pupil is excluded from the BE program.

in the BE stream, thus further supporting the claim of non-selectivity of BE for these pre-vocational secondary groups. The BE segment of this population, therefore, cannot be called selective, unlike the BE streams at higher types of secondary education in the Netherlands and elsewhere (Bruton, 2011, 2015; Cenoz, Genesee, & Gorter, 2014; Rumlich, 2014; Verspoor et al., 2015).

2.2 Instruments

Data were gathered by means of two questionnaires, a biodata questionnaire for personal information (gender, age, parents' first language and pupils' language spoken with parents) and an MPB questionnaire with items representing all MPB constructs (except Behavior). As there was no extant MPB instrument for measuring attitude towards learning English, the MPB items were specially constructed based on MPB questionnaire guidelines (Ajzen, n.d.), written in statement form in Dutch and evaluated on a 5-point Likert scale (1 = completely disagree; 5 = completely agree). It was piloted with a group of first-year pre-vocational secondary pupils from a school not participating in the research project. In a 'think-aloud' session with the head researcher, the pilot pupils told what they thought each item meant and indicated if there was any confusion. Subsequently, several of the items were re-worded or dropped. The final questionnaire contained 40 MPB items (see Appendix). An example statement for each MPB construct is given below:

- Cognition: People who are good at English have a better future.
- Affect: I like learning English.
- Subjective Norm: I would be ashamed later if I didn't know English well.
- PBC: I don't have enough time to learn English (rescored).
- Intention: I intend to work on my English every day.

2.3 Procedure

The cross-sectional sample contains pupils in years 1–3 in class groups participating in either a BE program, or a regular non-BE program with Dutch as the only language of instruction. The questionnaires were administered at the beginning of the school year under the supervision of a teacher and a researcher. On several occasions either the teacher or the researcher had to read the statements aloud to a pupil hindered by reading difficulties.

3. Analyses

The first research question (To what extent does each MPB model fit the data, and which of the three fits best?) was answered through confirmative factor analyses (Muthén & Muthén, 1998–2004), using maximum likelihood. Exact fit is reflected by a non-significant χ^2 test. Since the power of χ^2 becomes very large with large sample sizes, and because exact fit is a very strict criterion for social science research (MacCallum, Browne, & Sugawara, 1996), the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA) and the Root Mean Square Residual (SRMR) were also used. Values of CFI and TLI higher than .90 (Bentler, 1992) indicate good fit. Values of RMSEA below .05 are indicative of close fit, .05–.08 fair fit, .08–.10 mediocre fit and values > .10 misfit (MacCallum et al., 1996). For SRMR, values below .08 indicate good fit (Hu & Bentler, 1999). Modification indices were used to identify items causing misfit. Nine of the 40 MPB items were removed, after which 31 items remained: 9 for Cognition, 5 for Affect, 5 for Subjective Norm, 6 for PBC, and 6 for Intention. To determine the best-fitting model, the fits of these three nested models were compared by subtracting the values of χ^2 and the numbers of degrees of freedom of both models to create a new χ^2 test to verify the significance of fit improvement.

To answer the second research question, concerning which components of the MPB best predict the intention to learn English, the structural model of the best-fitting model (see above) is interpreted. Because suppressor effects are possible in a structural model, the zero-order correlations between sums of the items per MPB construct are also examined. The power of the structural equation modeling is influenced among other things by the number of indicators per factor and the magnitude of the factor loadings (Wolf, Harrington, Clark, & Miller, 2013). Having five to nine indicators per factor we intended to sample at least 400 pupils, so our sample of 488 pupils used for the structural equation modeling is adequate.

To answer the question concerning measurement invariance, the resulting best-fitting model was refitted as a two-group model (BE-pupils vs non-BE pupils). Three nested two-group models were fitted: one implying strong factorial invariance (factor loadings and intercepts constrained to be equal in both groups), one implying metric, but no scalar invariance (factor loadings constrained to be equal in both groups, intercepts free to vary over groups) and one implying no measurement invariance (factor loadings and intercepts free to vary over groups). Fit improvement between nested models was verified through a χ^2 test based on the difference in χ^2 and in degrees of freedom between both models.

The fourth research question (Do BE pupils score higher on the MPB components than non-BE pupils?) is answered by means of regression analyses.

Dependent variables are the sums of the items for each construct of the MPB. The reliability (Cronbach's alpha) of each of the sums is also calculated. Regression analyses were conducted separately for each sum representing an MPB construct. These analyses are also used to verify whether first-year BE and non-BE pupils differ in their scores on the MPB sums at the outset.

Since we have a clustered sample (pupils within classes within schools), analyses were conducted multilevel. Significance of fit improvement by adding a variance component to the model was verified by means of the χ^2 distributed difference in deviance of both nested models. Since variances cannot be negative, the probability found should be divided by 2 (Hox, 2010). Multilevel regression analyses were conducted using the program ML-Win (Rasbash et al., 2000). The difference between scores on MPB constructs of BE and non-BE pupils was evaluated both with and without correcting for pupil characteristics (covariates). Covariates used were the dummies indicating the pupil's pre-vocational stream level, gender, the native tongue of the father and mother (Dutch or not), and the language spoken with the father and mother (Dutch or not) Covariates were added to eliminate alternative explanations and to increase power; only covariates that significantly predict the dependent variable were included.

The significance of independent variables was calculated in two ways: first, by the χ^2 test calculated as the difference between both deviances of a model with and a model without the predictor(s) with as degrees of freedom the difference in numbers of parameters estimated in both models. Secondly, significance of predictors was evaluated by dividing the regression coefficient of the predictor by its standard error, resulting in a t-score with a number of degrees of freedom equaling the sample size minus the number of predictors minus 1 (N-p-1). For pupil level variables, the sample size is the number of pupils; for class level variables, the number of classes; and for school level variables, the number of schools (Hox, 2010).

The last research question (How large are the differences in scores for BE and non-BE pupils on the MPB components in the three years, and do these differences increase as pupils gain more BE experience?) was answered by testing the effect of year and the interaction effect between BE and the pupil's year for each MPB construct. To estimate this interaction effect, two interaction terms for each construct were added to the model: the interaction between BE and year 2 and between BE and year 3, implying that year 1 is the reference group. These analyses were conducted both with and without correcting for significant covariates. For testing contrasts between years 2 and 3, the analyses were also repeated with year 3 as reference group.

4. Results

Confirmatory factor analyses to answer the first question show that the model as proposed by Ajzen (1991) shows a good fit ($\chi^2 = 811.197$; $df = 427$; $p < .000$; RMSEA = .043; CFI = .90; TLI = .89; SRMR = .054). All items have significant factor loadings on the factor they are supposed to measure (all $p < .001$). The next model is that in which Cognition and Affect are combined into one construct (Dillon & Kumar, 1985). This model also has a good fit ($\chi^2 = 896.255$; df 428; $p < .000$; RMSEA = .047; CFI = .88; TLI = .87; SRMR = .057); again, all factor loadings of the items on their factor are significant (all $p < .000$). The difference in fit between both models is significant ($\chi^2 = 85.058$, $df = 1$, $p < .000$); therefore Cognition and Affect should not be combined into one construct. The third model fitted is the modified MPB in which Cognition has a path directly onto Intention, not mediated by Affect. Again the model shows a good fit ($\chi^2 = 755.175$; $df = 424$; $p < .000$; RMSEA = .040; CFI = .91; TLI = .91; SRMR = .050) and all item factor loadings are significant (all $p < .000$). The difference in fit between both models is significant ($\chi^2 = 56.022$; $df = 3$; $p < .000$). This last model, with Cognition loading directly onto Intention, is thus the preferred (less parsimonious) model.

To answer the second research question, the structural model of this best-fitting model was inspected. Since the structural model can be interpreted as a regression model (albeit after removing error) which means that highly correlated predictors (here: latent traits) might result in suppressor effects, we inspected not only the standardized coefficients of the structural model, but also the zero-order correlations between the item sums of all constructs measured in the instrument, taking into account that the zero-order correlations are lower because the sums contain error and the latent traits in the structural model do not. The structural model is presented in Figure 2 and the zero-order correlations are presented in Table 2.

In the structural model (Figure 2), it appears that only Cognition has a significant loading on Intention (standardized loading = .621). Cognition also correlates significantly and highly with Subjective Norm ($r = .815$) and Affect ($r = .786$). The correlation between Cognition and Perceived Behavioral Control (PBC) is rather low ($r = .198$), albeit significant. Significant correlations are also found between Affect and Subjective Norm ($r = .604$) and between Affect and PBC ($r = .450$). The result of the model fit therefore leads to the conclusion that for the pupils in the sample, Cognition (perceived importance) is the most important precursor of the Intention to work hard to learn English. Furthermore, those who score higher on Cognition also tend to score higher on Subjective Norm and Affect, and those who score higher on Affect also tend to score higher on PBC. In order to also

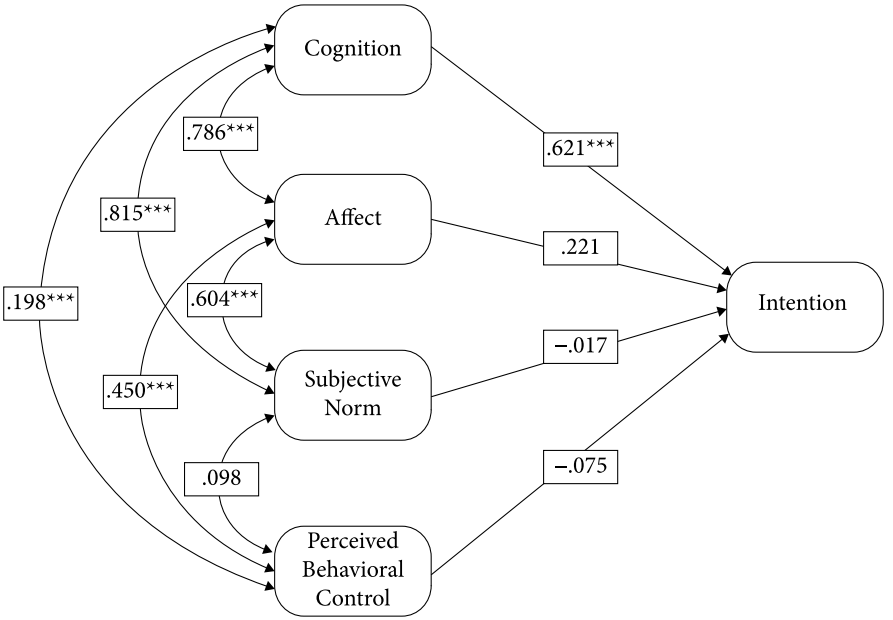

Figure 2. Structural part of the third MPB model tested (best fitting model) (* = $p < .05$; ** = $p < .01$; *** $p < .001$). Two-directional arrows indicate correlations; one-directional arrows indicate paths (loadings)

inspect the zero-order correlations, we first inspected the reliabilities of the sums (see Table 1).

Table 1. Reliabilities (Cronbach's alpha) of sum scores MPB

Variable	Number of items	Cronbach's alpha	N	Range item-test correlations
Cognition	9	.69	488	.236–.437
Affect	5	.85	488	.475–.751
Sub. norm	5	.72	486	.658–.716
PBC	6	.64	486	.239–.521
Intention	6	.76	487	.324–.589

The Cronbach's alphas for Cognition and PBC are lower than for the other constructs. It is likely that these relatively low alphas were found because they measure different, non-interchangeable aspects of 'usefulness' (Cognition) and 'practicability' (PBC) which do not necessarily have to show high intercorrelations in order to present a valid sum score indicative of their respective constructs (Davis et al., 2002).

Comparing the standardized regression coefficients and correlations from the structural model (Figure 2) with the correlations in Table 2, we can see that with the exception of PBC, all constructs show relatively high correlations with Intention. Only Cognition is a highly significant predictor of Intention in the structural model, but the strong zero-order correlations between the Cognition, Affect, and Subjective Norm on the one hand and Intention on the other mean that causal relationships between these constructs and the Intention to work hard to learn English cannot be excluded. PBC also correlates significantly with Affect (zero-order correlation of .287 and in the structural model .450) but not with the other constructs.

Table 2. Zero order correlations between sums scores MPB (higher scores = more positive attitude)

	Affect	**Sub. norm**	**PBC**	**Intention**
Cognition	.608 ** (487)	.576 ** (485)	.074 (486)	.564 ** (486)
Affect		.481 ** (485)	.287 ** (485)	.553 ** (486)
Subjective norm			−.004 (484)	.470 ** (485)
PBC				.035 (486)

** Correlation is significant at the 0.01 level (2-tailed).

Next, measurement invariance of the instrument for BE and non-BE pupils was verified (research question 3). To do so, first the best-fitting model was fitted separately in both groups (BE pupils and non-BE pupils). The model fit is reasonably good (the non-BE group $\chi^2 = 642.509$; $df = 424$; $p < .000$; RMSEA = .048; CFI = .87; TLI = .86; SRMR = .063: the BE group $\chi^2 = 669.066$; $df = 424$; $p < .000$; RMSEA = .046; CFI = .89; TLI = .87; SRMR = .063). Then three nested two group models were fitted (see Table 3, Appendix). Since none of the differences in fit between the models is significant, the most parsimonious model is accepted (c.f. Bagozzi & Yi, 1988). This implies that the measurement based on the third model, the modified MPB, shows strong factorial invariance over the BE- and the non-BE-group: the MPB scores of both groups of pupils can be validly compared.

It appears that all MPB sums contain significant proportions of class level variance, but no significant proportions of school level variance. Regression analyses are therefore performed multilevel with a pupil and a class level.

In Table 4 we can see that BE pupils score significantly higher than non-BE pupils on four of the five MPB constructs: Cognition, Affect, PBC and Intention. Being a BE- or a non-BE pupil explains about 40% of the differences between classes for Cognition, Affect and Intention; for PBC this is 25%. The differences in mean scores for BE- and non-BE-pupils for Subjective Norm are non-significant.

The regression models also show that there is a non-significant difference between the scores of first-year BE and non-BE pupils for all attitude constructs (Cognition $t = -1.908$, $df = 19$, $p =$ n.s.; Affect $t = -.868$, $df = 19$, $p =$ n.s.; Subjective norm $t = -1.341$, $df = 19$, $p =$ n.s.; PBC $t = .243$, $df = 19$, $p =$ n.s.; Intention $t = -1.673$, $df = 19$, $p =$ n.s.).

Table 4. Results (effect sizes) of regression analyses with MPB-sums as dependent variables and variable indicating BE/non-BE as predictor (N varies between 483 and 488)

Dependent variable	Percentage of class level variance	Regression coefficient variable BE/non-BE (se between brackets)	Percentage of explained class level variance by BE/ non-BE	Percentage explained total variance
Cognition	7.7% ***	.218 ** (.070)	44.0%	3.4%
Affect	11.9% ***	.500 ** (.156)	36.4%	6.5%
Subjective Norm	5.1% **	.100 (.105)	–	–
PBC	4.1% **	.145 # (.081)	25.0%	1.0%
Intention	7.1% ***	.282 ** (.092)	43.9%	3.1%

# $p < .10$ (=5% one sided). * $p < .05$. ** $p < .01$. *** $p < .001$.
Percentages of explained variance only included if significant.

Correction for the significant covariates (year, home language, gender; Table 5), results in roughly the same picture, except that the differences between class means for Intention are now explained for 66.7% by whether the class is a BE class or not. Likewise, after correcting for significant covariates (only for Cognition, Affect and Intention) the differences in sums between BE and non-BE pupils in year 1 are again non-significant (Cognition $t = -1.822$, $df = 16$, $p =$ n.s.; Affect $t = -.741$, $df = 17$, $p =$ n.s.; Intention $t = -1.673$, $df = 19$, $p =$ n.s.).

To verify whether the difference between BE-pupils and non-BE-pupils increases for year 2 compared to year 1 (the fifth research question), the interaction between year and the BE variable was inspected with separate analyses for each MPB-construct as dependent variable: once with and once without correcting for significant pupil characteristics.

Table 6 (see Appendix, online) presents two models for each MPB-construct: one containing the main effects of year and BE and one with both interaction terms (year2*BE and year3*BE) added, enabling comparison of the nested models. Adding both interaction terms significantly improves model fit only for Affect, explaining 37.0% of the class bound variance and 4.2% of total variance. No pupil variance is explained, unsurprisingly since both year and BE are class-related variables.

Table 5. Results (effect sizes) of regression analyses with MPB-sums as dependent variables and variable indicating BE/non-BE as predictor after correcting for significant covariates

Dependent variable	% class level variance with covariates in the model	Unstandardized regression coefficient variable BE/non-BE (se between brackets) with covariates in the model	% of explained class level variance by BE/non-BE after correcting for effect covariates	% explained total variance by BE/non-BE after correcting for effect covariates	Significant covariates added to the model
Cognition	8.5% ***	−.212 ** (.073)	37%	3.1%	First language mother, language spoken by pupil with father, gender
Affect	17.4% ***	−.460 ** (.158)	31.9%	5.6%	First language mother, first language father
Subjective Norm	2.9%	−.089 (.091)	–	–	Year 2; year 3 (year 1 = ref. group)
PBC	–	–	–	–	No covariates significant for PBC
Intention	3.8% *	−.252 ** (.076)	66.7%	2.9%	Year, first language mother, language spoken by pupil with father, language spoken by pupil with mother

* $p<.05$. ** $p<.01$. *** $p<.001$.
Percentages of explained variance only included if significant.

Inspection of the regression coefficients reveals that for Affect the interaction between BE and year 2 has a significant coefficient of -.870 (Affect (year2*BE) $t = 2.890$; $df = 20$; $p < .01$ (1-sided); non-BE = '1', BE = 'o'). The interaction term consisting of year 3 and BE is not significant (see Figure 3). An additional analysis with year 3 as reference group (dummy year 3 replaced by dummy year 1 in the regression model) shows that the difference for Affect between BE and non-BE pupils in year 3 does not differ significantly from this difference in year 2 (without covariates: $t = 1.6693$, $df = 19$; $p =$ n.s.: with covariates: $t = 1.910$; $df = 17$; $p =$ n.s.). This implies that the difference in Affect between BE and non-BE pupils is significantly larger in the second year than in the first, but that difference in year 3 does not differ significantly from the same differences in years 1 or 2.

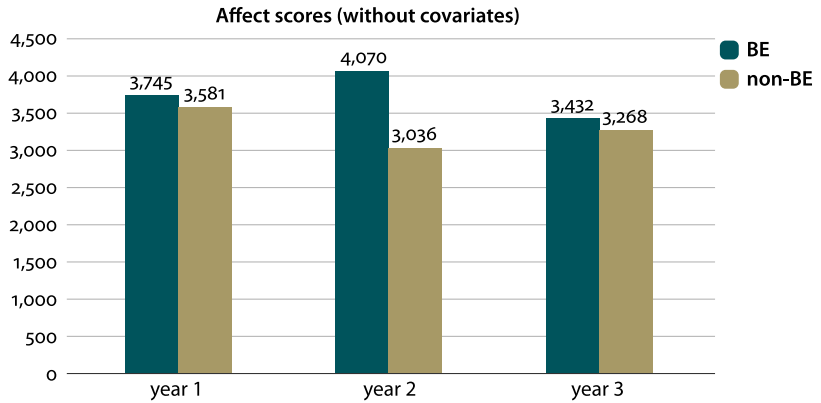

Figure 3. Affect scores for BE and non-BE pupils per year (without controlling for covariates)

In the analyses with PBC as dependent variable, both interaction contrasts in Table 6 (year 1 vs. year 2; year 1 vs. year 3) are positively significant for BE ($p < .05$), but the difference between models with and without the interaction terms is only significant at 10%. The proportion of explained variance at class level for PBC is larger than for Affect, 53.8%. The proportion of class level variance in PBC scores is very small, however, resulting in only 1.4% explained total variance. An additional analysis with year 3 as reference group (see Figure 4) shows that for PBC the contrast for the difference between year 2*BE vs. year 3*BE is not significant (no covariates: $t = -.026/.200 = .130$; $df = 19$; $p =$ n.s.).

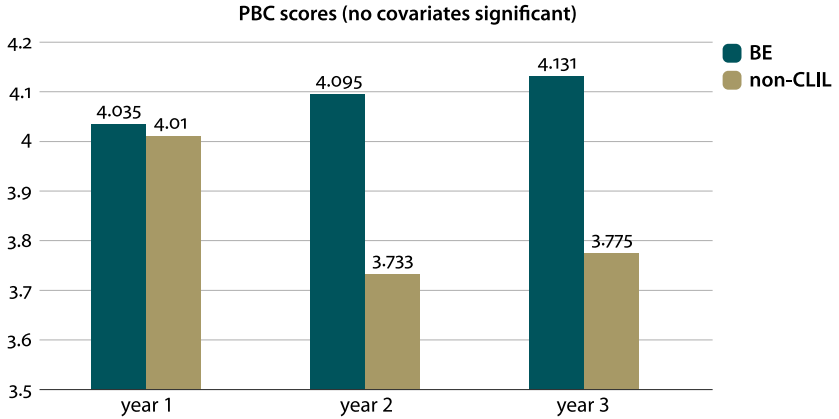

Figure 4. PBC scores for BE and non-BE pupils per year

5. Discussion and conclusions

The data collected from over 500 pre-vocational secondary pupils yielded revelatory results concerning the prediction of their intention to learn English and the difference between BE and non-BE pupils' attitudes in years 1, 2, and 3 towards learning English. The first research question concerned the choice of the best structural model. Analyses showed that the best-fitting model was one in which Cognition and Affect served independently as predictors of Intention. Although highly correlated (correlation in the structural model .786, $p < .001$), it appears that these two constructs represent distinct aspects of the attitude towards learning English (Breckler & Wiggins, 1989; Trafimow & Sheeran, 1998): broadly stated, utilitarian goals versus emotional factors. For instance, a pupil might intend to work hard because he or she enjoys learning English (Affect), but also because he or she believes that a good command of English will be useful at senior vocational school (Cognition). This is noteworthy, as many other studies (reported in Ajzen, 1991) have found that the effect of Cognition on Intention is mediated by Affect, implying that Cognition does not influence Intention directly.

The second research question asks which components of the MPB best predict Intention. Although the regression coefficient from Cognition to Intention shows the strongest relationship, the zero-order correlations between two of the other three predictors (Affect and Subjective Norm) are also high, implying they are suppressed to a certain extent by Cognition. Still, Cognition predicts best, and generally explains the same variance which the other two predictors share with Intention; only PBC shows a minimal relationship with Intention, both in the model and in the zero-order correlations. Apparently the perceived importance and usefulness of learning English are quite relevant for teenagers who are not particularly academically-inclined.

The best-fitting structural model results show that there are high correlations between believing learning English is important (Cognition), and liking it (Affect), and between believing learning English is important and perceiving social pressure to learn it (Subjective Norm). There is a medium-strength correlation between enjoying learning English and perceiving few impediments to learning it (Perceived Behavioral Control); this correlation might be lower because pupils like learning English for various non-academic reasons, such as gaming, television, films, music, and social media, and perceive few difficulties in learning it, which may be partly due to the prevalence and accessibility of English in daily life in the Netherlands. Finally, there is a weak but still significant correlation between believing English is important and perceiving few impediments to learning it, as was also found by Green, Miller, Crowson, Duke, and Akey (2004). Importantly, though, the structural equation modelling showed that in our sample PBC does not contribute to the prediction of Intention: a perception of impediments to learning English does not predict the intention to make more or less effort. PBC is in relation to the other constructs the 'odd one out'.

One strength of this kind of behavioral model result is its possible contribution to designing effective interventions for producing behavioral change (Fishbein, in Connor & Armitage, 1998), so the results concerning the strength of the predictors may be useful in designing interventions to elicit the desired behavior at school. Since the structural model results are for all pupils in the study (BE and non-BE together), these possible interventions are not limited to a BE population. For instance, BE programs in the Netherlands include a 'European and International Orientation' (EIO) aspect which may indirectly emphasize the importance of English globally; this could be strengthened and extended to non-BE pupils in order to tap into the strength of the Cognition construct.

The nested two-group models (BE and non-BE pupils) were fitted in order to determine measurement invariance (research question three). The resulting analysis showed strong factorial invariance, so that the results from the two groups could validly be compared in order to answer the fourth research question, concerning the comparison of attitudes of BE and non-BE pupils. The results here confirm the results of other studies (Doiz et al., 2014; Lasagabaster & Sierra, 2009; Mearns et al., 2017; Merisuo-Storm, 2011; Rumlich, 2014) regarding a more positive attitude among BE pupils: the BE pupils scored significantly higher on the model constructs Affect, Cognition, Intention and PBC than non-BE pupils. The results show that this more positive attitude was not yet present in first-year pupils at the start of BE, as none of the MPB constructs showed a significant difference between the BE and non-BE pupils at the outset; the two groups were comparable in terms of attitude towards making an effort to learn English at the start of prevocational secondary education. This is in contrast to results for more challenging

secondary streams, as was found in Mearns (2015) and Rumlich (2015), or highly selective programs (Doiz et al., 2014), and supports the schools' claim that there was no pre-selection of the BE pupils based on attitude.

The final research question concerned the difference in MPB construct scores between BE and non-BE pupils in years 1, 2, and 3. The interaction effects for Affect and PBC between BE/non-BE and the pupil's year show that the differences between BE and non-BE pupils are significantly larger in second-year than in first-year pupils. This suggests that BE may have contributed to a growth in positive attitude, which is a salient finding, as Mearns et al. (2017) found that attitude differences in a more academic secondary stream were lower in third-year than first-year pupils. In year 3 in the current study – when BE is substantially reduced – the difference in Affect between BE and non-BE pupils disappears and is about the same as in year 1. However, for PBC the BE advantage seems to be maintained in year 3. As noted in the Introduction, the BE programs at pre-vocational secondary level in the Netherlands form a much more substantial part of the curriculum in years 1 and 2 than in years 3 and 4 due to Dutch-language vocational profiling and Dutch final examination preparation in the latter two years. This greater focus on BE in the first two years may explain some of the results in this study: as the programs shrink and third-year BE pupils change to regular Dutch-language schooling, the positive effects of BE on Affect seem to diminish. Another explanation could be the general trend towards less enthusiasm for learning as pupils age and progress to higher years, as reported by Mearns (2015) and Mearns et al. (2017). Regarding the increase in Perceived Behavioral Control, which is related to autonomy and self-efficacy, it may be that the extra exposure to English in the BE group, the heterogeneous input from different teachers, and a wider variety of language strategies give pupils the tools and skills to exert more control over their own learning, and that these tools and skills do not diminish when the BE program ends. For PBC, the benefits of BE for pre-vocational secondary pupils seem to support the purported "sustainable advantage" of pre-university bilingual education in the Netherlands that was found in Verspoor and Edelenbos (2011) and Verspoor et al. (2015). For Affect, however, this "sustainable advantage" seems to dwindle after the initial period of more intensive BE. Put simply, learning English may not be as enjoyable once the amount of BE decreases. Affect is, after all, the second best predictor of Intention, and sustained enjoyment of learning English indicates a sustained intention to make an effort to learn it. Clearly, longitudinal data analyses are needed to further support these claims. For this study, at least, BE and its related CLIL pedagogy seem to contribute to an increase in enjoyment and motivation of a wide variety of lower-achieving language learners, appearing to confirm some of the side effects of CLIL stated by Coyle (2006, 2013) – learner engagement and involvement, enjoyment, confidence, self-esteem, motivation, and autonomy.

Limitations and recommendations

Since the data are cross-sectional, definitive assumptions cannot be made about pupils' attitude development over time. Because it was not possible to obtain a measurement of actual behavior, we can only predict behavior from the Intention construct; a measurement of behavior could be included in order to verify or disprove the prediction of behavior from intention. Further research could replicate this study using a longitudinal design to establish whether the expected positive causal effect of BE on the constructs of the attitude model can be verified. Such a study might also explore the correlation between the attitudes of BE/non-BE pupils and their language proficiency development over time. Considering suggestions in the literature regarding potential benefits of CLIL for boys, a similar study with emphasis on gender differences might also be relevant (see Mearns & De Graaff, 2018). Finally, a similar study could be carried out in another national context, although outside the Netherlands this is difficult, as BE is seldom offered specifically for negatively-selected 'average to below average' secondary pupils. It is hoped that the present study will encourage more research into bilingual education and CLIL for a non-selective population.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, 32(4), 665–683. https://doi.org/10.1111/j.1559-1816.2002.tb00236.x

Ajzen, I. (2005). *Attitudes, personality, and behavior* (2nd ed.). Milton-Keynes: Open University Press/McGraw- Hill.

Ajzen, I. (n.d.) Constructing a TPB Questionnaire. In *Theory of Planned Behavior*. Retrieved from <http://people.umass.edu/aizen/pdf/tpb.measurement.pdf>

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*, 40(4), 471–499. https://doi.org/10.1348/014466601164939

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74–94. https://doi.org/10.1007/BF02723327

Bagozzi, R. P., & Burnkrant, R. E. (1979). Attitude organization and the attitude-behavior relationship. *Journal of Personality and Social Psychology*, 37(6), 913. https://doi.org/10.1037/0022-3514.37.6.913

Bagozzi, R. P., & Burnkrant, R. E. (1980). Single component versus multicomponent models of attitude: Some cautions and contingencies for their use. In J. C. Olson (Ed.), *NA – Advances in Consumer Research* (Vol. 7, pp. 339–344). Ann Abor, MI: Association for Consumer Research.

Bagozzi, R. P., & Burnkrant, R. E. (1985). Attitude organization and the attitude-behavior relation: A reply to Dillon and Kumar. *Journal of Personality and Social Psychology*, 49(1), 47–57. https://doi.org/10.1037/0022-3514.49.1.47

Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, 112, 400–404. https://doi.org/10.1037//0033-2909.112.3.400

Breckler, S. J., & Wiggins, E. C. (1989). Affect versus evaluation in the structure of attitudes. *Journal of Experimental Social Psychology*, 25(3), 253–271. https://doi.org/10.1016/0022-1031(89)90022-X

Brinberg, D. (1981). A comparison of two behavioral intention models. In K. B. Monroe (Ed.), *NA – Advances in Consumer Research* (Vol. 8, pp. 48–52). Ann Abor, MI: Association for Consumer Research.

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

Bruton, A. (2015). CLIL: Detail matters in the whole picture. More than a reply to J. Hüttner & U. Smit. (2014). *System*, 53, 119–128. https://doi.org/10.1016/j.system.2015.07.005

Cenoz, J. (2001). Three languages in contact: Language attitudes in the Basque Country. In D. Lasagabaster & J. M. Sierra (Eds.), *Language awareness in the foreign language classroom*. Zarauz: University of the Basque Country.

Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics*, 35(3), 243–262. https://doi.org/10.1093/applin/amt011

Centraal Bureau voor de Statistiek. (2014). *Jaarrapport Integratie 2014*. Retrieved from <http://www.cbs.nl/NR/rdonlyres/E6878ED8-0347-4ED0-8A8D-360AB79022B2/0/jaarrapportintegratie2014pub.pdf>

Centraal Bureau voor de Statistiek. (2016). *VO; leerlingen, onderwijssoort, leerjaar, herkomstgroepering, generatie* [data file]. Retrieved from <http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=80043NED&D1=2-11,18&D2=0&D3=0&D4=a&D5=0&D6=0&D7=0&D8=9&HDR=G2,G5,G6,G7,G4,T&STB=G3,G1&VW=T>

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Conner, M., & Armitage, C. J. (1998). Extending the theory of planned behaviour: A review and avenues for further research. *Journal of Applied Social Psychology*, 28, 1430–1464. https://doi.org/10.1111/j.1559-1816.1998.tb01685.x

Conner, M., & Sparks, P. (2005). The theory of planned behaviour. In M. Conner, & P. Norman (Eds.), *Predicting health behaviour: Research and practice with social cognition models* (2nd ed., pp. 121–162). Maidenhead: Open University Press.

Coyle, D. (2006). Content and Language Integrated Learning: Motivating learners and teachers. *Scottish Languages Review*, 13, 1–18.

Coyle, D. (2013). Listening to learners: An investigation into 'successful learning' across CLIL contexts. *International Journal of Bilingual Education and Bilingualism*, 16(3), 244–266. https://doi.org/10.1080/13670050.2013.7773

Davis, L. E., Ajzen, I., Saunders, J., & Williams, T. (2002). The decision of African American students to complete high school: An application of the theory of planned behavior. *Journal of Educational Psychology*, 94(4), 810–819. https://doi.org/10.1037//0022-0663.94.4.810

De Graaff, R., & Van Wilgenburg, O. (2015). The Netherlands: Quality control as a driving force in bilingual education. In P. Mehisto & F. Genesee (Eds.), *Building bilingual education systems: Forces, mechanisms and counterweights* (pp. 167–179). Cambridge: Cambridge University Press.

Dillon, W. R. & Kumar, A. (1985). Attitude organization and the attitude-behavior relation: A critique of Bagozzi and Burnkrant's reanalysis of Fishbein and Ajzen. *Journal of Personality and Social Psychology*, 49(1), 33–46. https://doi.org/10.1037/0022-3514.49.1.33

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

Fen, Y. S., & Sabaruddin, N. A. (2008). An extended model of theory of planned behaviour in predicting exercise intention. *International Business Research*, 1(4), p108. https://doi.org/10.5539/ibr.v1n4p108

Godin, G., & Kok, G. (1996). The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion*, 11(2), 87–98. https://doi.org/10.4278/0890-1171-11.2.87

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology*, 29(4), 462–482. https://doi.org/10.1016/j.cedpsych.2004.01.006

Hartgers, M. (2007). Scholen in de Randstad sterk gekleurd. *Sociaaleconomische trends, 3e kwartaal 2007* (pp. 12–18). Rijswijk: Centraal Bureau voor de Statistiek.

Hox, J. J. (2010). *Multilevel analysis. Techniques and applications. Quantitative methodology series* (2nd ed.). New York, NY: Routledge.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. https://doi.org/10.1080/10705519909540118

Keil, L. J. (1991). Attitude development. In T. Husen, & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (Vol. 1, pp. 346–352). New York, NY: Pergamon Press.

Koomen, H. M., Verschueren, K., Van Schooten, E., Jak, S., & Pianta, R. C. (2012). Validating the student-teacher relationship scale: Testing factor structure and measurement invariance across child gender and age in a Dutch sample. *Journal of School Psychology*, 50(2), 215–234. https://doi.org/10.1016/j.jsp.2011.09.001

Lasagabaster, D., & Sierra, J. M. (2009). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal*, 1(2), 4–17.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Mearns, T. L. (2015). Chicken, egg, or a bit of both? Motivation in bilingual education (TTO) in The Netherlands (Unpublished PhD dissertation). University of Aberdeen, Scotland. Retrieved from <http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.646095>

Mearns, T., & De Graaff, R. (2018). 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. L., De Graaff, R., & Coyle, D. (2017). Motivation for or from bilingual education? A comparative study of learner views in the Netherlands. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2017.1405906

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525–543. https://doi.org/10.1007/BF02294825

Meredith, W. M., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care*, 44, S69–S77. https://doi.org/10.1097/01.mlr.0000245438.73837.89

Merisuo-Storm, T. (2011). *Literacy skills and learning attitudes in bilingual education. Education in Finland*. Taipei: Taiwan National Institute of Educational Resource and Research.

Muthén, L. K., & Muthén, B. O. (1998–2004). *Mplus user's guide* (4th ed.). Los Angeles, CA: Muthén & Muthén.

Nuffic, n.d. *Standaard tweetalig vmbo*. Retrieved from <https://www.nuffic.nl/bestanden/documenten/standaard-tweetalig-vmbo/view?searchterm=t-vmbo>

Pels, T., Jonkman, H., & Drost, L. (2011). *Socialiseren, leren, presteren: Over motivatie van leerlingen in het voortgezet onderwijs*. Verwij-Jonker Instituut. Retrieved from <https://www.verwey-jonker.nl/doc/jeugd/Socialiseren_Leren_Presteren_2845_web.pdf>

Pérez-Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism*, 15(3), 315–341. https://doi.org/10.1080/13670050.2011.630064

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, H., Draper, D., Langford, I., & Lewis, T. (2000). *A user's guide to MlwiN*. London. Institute of Education, University of London.

Rumlich, D. (2014). Prospective CLIL and non-CLIL students' interest in English (classes): A quasi-experimental study on German sixth-graders. In R. Breeze, C. Llamas Saíz, C. Martínez Pasamar, & C. Tabernero Sala (Eds.), *The integration of theory and practice in CLIL* (pp. 75–95). Amsterdam: Rodopi.

Rumlich, D. (2015). Zur affektiv-motivationalen Entwicklung von Lernenden im CLIL Unterricht. In B. Rüschoff, D. Wolff, & J. Sudhoff (Hg.), *CLIL revisited. Eine kiritische Analyse zum gegenwärtigen Stand des bilingualen Sachfachunterrichts* (pp. 309–330). Frankfurt: Peter Lang.

Sideridis, G. D., & Padeliadu, S. (2001). The motivational determinants of elementary school students at risk of having reading difficulties: An application of planned behavior theory with goal importance. *Remedial and Special Education*, 22, 268–273. https://doi.org/10.1177/074193250102200502

Sultan, S., Borland, H., & Eckersley, B. (2012). English as medium of instruction (EMI) in Indonesia public junior secondary school: Students' language use, attitudes/motivation and foreign language outcomes. *ACTA International TESOL Conference*. Cairn Australia, Victoria University.

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.3.1.02syl

Trafimow, D., & Sheeran, P. (1998). Some tests of the distinction between cognitive and affective beliefs. *Journal of Experimental Social Psychology*, 34(4), 378–397. https://doi.org/10.1006/jesp.1998.1356

Triandis, H. C. (1977). *Interpersonal Behavior*. Monterey, CA: Brooks/Cole.

Valois, P., Desharnais, R., & Godin, G. (1988). A comparison of the Fishbein and Ajzen and the Triandis attitudinal models for the prediction of exercise intention and behavior. *Journal of Behavioral Medicine*, 11(5), 459–472. https://doi.org/10.1007/BF00844839

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70. https://doi.org/10.1177/109442810031002

Van Nuland, H. J. C. (2011). Eliciting classroom motivation: Not a piece of cake (Unpublished PhD dissertation). University of Leiden. Retrieved from <http://hdl.handle.net/1887/16693>

Van Schooten, E., De Glopper, K., & Stoel, R. (2004). Development of attitude toward reading adolescent literature and literary reading behavior. *Poetics*, 32(5), 343–386. https://doi.org/10.1016/j.poetic.2004.07.001

Verspoor, M., & Edelenbos, P. (2011). Tweetalig onderwijs zorgt voor een duurzame voorsprong. *Levende Talen Tijdschrift*, 12(4), 5–13.

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

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73(6), 913–934. https://doi.org/10.1177/0013164413495237

Address for correspondence

Jenny Denman
Rotterdam University of Applied Sciences
Research Centre Urban Talent
Museumpark 40
3015 CX Rotterdam
The Netherlands
j.l.denman@hr.nl

Co-author information

Erik van Schooten
Rotterdam University of Applied Sciences
Research Centre Urban Talent
e.j.van.schooten@hr.nl

Rick de Graaff
Utrecht University
Utrecht Institute of Linguistics OTS
r.degraaff@uu.nl