



university of  
 groningen

# **Differentiation to improve the articulation between levels**

In the teaching of English in primary and secondary education  
 in the Netherlands

## **PhD thesis**

to obtain the degree of PhD at the  
 University of Groningen  
 on the authority of the  
 Rector Magnificus Prof. E. Sterken  
 and in accordance with  
 the decision by the College of Deans.

This thesis will be defended in public on  
 Monday 21 November 2016 at 16.15 hours

by

**Antonie Pleun de Kraay**

born on 2 August 1966  
 in Rotterdam

**Supervisor**

Prof. dr. C.L.J. de Bot

**Co-supervisor**

Dr. E.J. van Schooten

**Assessment Committee**

Prof. dr. K. van Veen

Prof. dr. T. Tokuhamma-Espinosa

Prof. dr. P. Burmeister

## **Acknowledgments**

Let me take this opportunity, before all is read and discussed, to acknowledge the people who inspired and supported me on this journey. First and foremost, I would like to thank Prof. Dr. Kees de Bot, who saw something in the research proposal I sent him almost six years ago. I could not have wished a better promotor; he was always quick to react and assist where possible, while he consistently allowed this thesis to be my own work. To say 'his door was always open' is an understatement, considering the international phone calls made outside of regular working hours. The second person I am eternally indebted to, who is in all others things second to none, is my co-supervisor, dr. Erik van Schooten. He is quick to counter reproach with friendly compliments; he lectures with genuine interest and in-depth discussions, interspersed with stories of his five wonderful daughters, which goes to show his mind is as razor-sharp as his wit. Erik initiated me into the amazing world of statistics, helped me to stay focused on the research as planned and kept me on track.

I would also like to thank the teachers who were involved in the experimental phase of the research:

Aksana van Steenbrugge, Elly Hanse and Joanne Kelly, as well as those involved in the preliminary rounds of discussions and material development: Hananae Hammouchi, Meyns Verhaar and Sandor van der Laan.

They are all passionate colleagues, willing to discuss the opportunities differentiation might offer and to open up their classrooms and teaching to an experiment. I would also like to thank my colleagues at the Rotterdam University School of Education, whose warmth and empathy provided me a workplace in which I could combine my duties as a teacher trainer with the hard work of being a researcher. They bravely withstood all my complaints and enjoyment at carrying out research during all the phases I underwent over the past five years. My parents, sadly my mother is no longer with us to see what I achieved, always firmly believed in every step I took. Their love and pride gave me the backbone to face and overcome all my challenges. Finally, I must express my deepest love and humblest gratitude to Ietske Roest, my wife, for her unfailing support and continuous encouragement throughout the years of researching and writing this thesis. This accomplishment would not have been possible without her. Thank you, my love.

**Table of contents:**

Acknowledgements.....	1
Annotated glossary.....	5
1. Introduction.....	6
Research question .....	8
1.1. Sub-questions.....	8
1.2. Sub-studies.....	10
2. Historical context of the teaching of English as a foreign language in primary and secondary education.	
2.1. Introduction.....	11
2.2. Method.....	11
2.3. Results	
Nineteenth century education and foreign language teaching.....	12
Grammar-translation.....	13
Reform movement (Europe).....	15
Progressive movement (USA).....	16
Behaviourism.....	17
The Oral Approach and Situational Language Teaching (Great Britain).....	18
The Audio-lingual Method (USA).....	18
Mastery Learning.....	19
The Communicative approach.....	20
The Common Europe Framework of Reference.....	22
English language teaching in Dutch primary education.....	23
Early English language teaching in The Netherlands.....	25
Early English language teaching in Rotterdam.....	26
Articulation between levels.....	27
2.4. Conclusion.....	29
3. Questionnaire Survey on the differences in English language teaching in primary and secondary education in The Netherlands.	
3.1. Introduction.....	32
3.2. Method used in the questionnaire survey on the differences in English language teaching in primary and secondary education in The Netherlands	
3.2.1. Design.....	33
3.2.2. Sample selection and characteristics.....	33
3.2.3. Instrument.....	35
3.2.4. Instrument distribution.....	36
3.3. Analyses.....	36
3.4. Results.....	37
3.4.1. Sample representativeness.....	37
Results of the representativeness analyses.....	40
3.4.2. Analysis Results Questionnaire.....	41
3.4.2.1 Amount of time spent on the four language skills (listening speaking, reading and writing), grammar and culture.....	42
3.4.2.2. Variety in language learning activities and assessments used in the classroom.....	44
Components found.....	45
Differences in language teaching methodology used between the primary and secondary education sample.....	48
3.5. Conclusion of the questionnaire survey on the differences in English language teaching in primary and secondary education in the Netherlands.....	52
4. Literature study on differentiated instruction: Historical context (A), Theoretical	

foundations (B), Building blocks (C) The role of Teachers in differentiated instruction (D) The effectiveness of differentiated instruction (E).	
4.1. Introduction.....	54
4.2. Method.....	55
4.3. Results.....	56
4.3.1. Results (A) - the historical context of differentiation	
4.3.1.1. Differentiation in nineteenth century education.....	57
4.3.1.2. The Progressive Movement.....	58
4.3.1.3. Aptitude Treatment Interaction.....	61
4.3.1.4. Differentiation in the final decades of the twentieth century.....	62
4.3.1.5. Differentiation in the first decades of the twenty-first century.....	66
4.3.2. Results (B) - the theoretical foundation of differentiation	
4.3.2.1. A workable definition of differentiated instruction.....	67
4.3.2.2. A research base for a new educational approach.....	70
4.3.2.2.1. Diversity of students.....	71
4.3.2.2.2. Effects of (not) responding to the diversity .....	73
4.3.2.2.3. Brain Research .....	75
4.3.2.2.4. Learning style.....	75
4.3.3. Results (C) - Building blocks of differentiated instructions.....	77
4.3.3.1. Students as partner in the educational process.....	78
4.3.3.2. Readiness, Interest and Learning Profile.....	81
4.3.3.3. Emotions.....	83
4.3.3.4. Motivation.....	84
4.3.3.5. Content, process and product.....	86
4.3.3.6. Grading and assessments.....	88
4.3.3.7. Feedback.....	91
4.3.3.8. Grouping students.....	92
4.3.4. Results (D) - The role of Teachers in differentiated instruction.....	93
4.3.5. Results (E) - The effectiveness of differentiated instruction.....	97
4.4. Conclusion of the literature study on differentiated instruction: theoretical foundations (B), building blocks (C), the role of Teachers in differentiated instruction (D) and prerequisites for a language teaching programme based on the principles of Differentiated Instruction.....	98
5. The development of a language teaching programme based on differentiated instruction.	
5.1. Introduction.....	102
5.2. The foundation of a language teaching programme based on differentiated instruction.....	102
5.3. The development of a language teaching programme based on differentiated instruction.....	105
5.4. The language teaching programme based on differentiated instruction.....	107
5.5. Conclusion: The implementation of the language teaching programme based on differentiated instruction as experimental treatment.....	112
6. Quasi experiment to assess the effect of a language teaching programme based on differentiated instruction on the attitude towards learning English and the English language skills development of students in the first year of secondary education	
6.1. Introduction.....	114
6.2. Method	
6.2.1. Design.....	118
6.2.2. Instruments.....	118
6.2.3. Instrument distribution.....	123
6.2.4. Sample.....	124

6.2.5. Fidelity check.....	125
6.3. Analyses.....	126
6.4. Results.....	130
6.5. Conclusion.....	140
7. Delphi Study	
7.1. Introduction.....	143
7.2. Set up of the Delphi study.....	145
7.3. Results.....	149
7.4. Conclusion.....	156
8. Conclusion and discussion	
8.1. Sub-questions answered.....	158
8.2. General conclusion.....	168
8.3. Discussion.....	170
9. English and Dutch summary.....	172
10. References.....	175
11. Appendix	
Table of contents .....	198
1. Questionnaire on the methodology used for English language teaching (Ch. 3)...	199
2. Structure matrix from the PCA (Chapter 3).....	201
3.1. The 37 specific types of implementation barriers .....	203
3.2. Rubrics as used during the experimental phase (Chapter 5):	
3.2.A. Rubrics used in the AYLITT project,.....	204
3.2.B. Rubrics as developed for the language teaching programme based on differentiated instruction.....	205
4. Attitude test used during the experimental phase (Chapter 6).....	207
5. Vocabulary MC test (Chapter 6).....	208
6. Dictation (Chapter 6).....	223
7. Editing test texts (Chapter 6).....	224
8. Writing test assignments (Chapter 6).....	226
9. Background information form for the students in the experiment (Chapter 6)....	227
10. Additional tables Chapter 6:	
10.a. Confirmatory factor analyses MPB pre- and post-test groups,.....	229
10.b. MI analyses to verify whether experimental and control group differ,.....	231
10.c. Correlations of sums of the MPB scores and Results of multi-level analyses of the effects of the treatment on language skills scores after controlling for covariates.....	234
10.d. Correlations of Reading skills scores. ....	273
11. Delphi Round 1 -Rephrased Complete (Chapter 7). ....	274
12. Delphi Round 1 -Rephrased and summarised for round 2 (Chapter 7). ....	280
13. Delphi Round 2-Rephrased and summarised for round 3 (Chapter 7). ....	285
14. Delphi Round 3 proposed text (Chapter 7).....	293
12. Curriculum Vitae.....	301

## **Annotated glossary**

This glossary consists of terms, with the definition or description, used in this dissertation. The terms will be elaborated upon in the chapters where they are used.

### **Differentiate (to)**

In this dissertation the general meaning is used to indicate someone or something is made different in some way or that the difference(s) between people or things are observed, recognized or used.

### **Differentiation in education**

*“To have differences develop between parts (e.g. schools, departments, year groups, subgroups, individual students) of an educational system (e.g. national teaching institutes, college, department, year group) with respect to one or multiple aspects (e.g. aims, contact hours, instructional approaches)” (De Koning, 1973, p. 3)*

### **Differentiated Instruction**

Differentiated instruction is embedded in the complexity of educational practice and as such does not allow a simple definition. Tomlinson (2014) says teachers that differentiate instruction because they: “... strive to do whatever it takes to ensure that struggling, advanced, and in-between learners, students with varied cultural heritages and children with a broad array of background experiences all grow as much as they possibly can each day, each week, and throughout the year.”

## 1. Introduction

In 1986, English became a compulsory subject in primary education in The Netherlands<sup>1</sup>. While secondary schools have, by law, been required to offer English since 1863 (Wilhelm, 2005) the articulation between Dutch primary and secondary education levels of English language teaching has always been problematic to the point it has been referred to as a connection problem (e.g.: Oostdam & Van Toorenburg, 2002; Holdinga, 2007). The term 'articulation between levels' has been used frequently to refer to the connection between parts of the language education system, such as between primary and secondary education and between secondary and tertiary education. In particular, the group of Richard Lambert at the National Foreign language centre in Washington promoted research on articulation between levels (Rhodes & Marsh, 1992; Brecht, & Lambert, 1999; Ingold, 2002; Liddicoat & Baldauf Jr., 2008). This study focusses on the connection between primary and secondary education and the possibility to improve the articulation between these levels with regard to the English teaching in The Netherlands.

Related research focuses on different aspects of the articulation between levels dependent on the view taken by the researcher. Connection problems in the articulation between levels can be seen to stem from the different ways in which English language teaching has evolved in primary and secondary education in the Netherlands. The connection problem arises from the long standing tradition of English language teaching in secondary education on one hand, and the relatively new subject of English language teaching in primary education on the other. Researchers like Wilhelm (2005; 2009) and writers like Kwakernaak (2011) describe the changes English teaching in Dutch secondary schools has gone through in past centuries and is still going through today. English teaching in *primary* education in The Netherlands, on the other hand, has only been around for about 30 years. The majority of Dutch schools offer the required amount of English teaching to children between the ages of ten and twelve, albeit in a limited fashion (Herder & De Bot, 2005). In the past decade, however, the number of Dutch primary schools offering English to 4 to 12-year olds has grown from a mere 45 to approximately 1000, which is one-seventh of all

---

<sup>1</sup> Compulsory after the law on primary education of 1986: 'Overgangswet WBO, 6 maart 1986, nr 138a – 19054' ,came into effect.

Dutch primary schools (Europees Platform, 2016). The amount of teaching offered in primary schools in The Netherlands can differ by no less than between two and eight years.

Other aspects of the problematic articulation between levels are the aim with which English education is offered (Onderwijsraad, 2008), the way in which teachers are trained (Thijs, Tuin & Trimbos, 2011), the methodology used to teach English (Edelenbos, 1993), or the whole complex interplay of policy, syllabus development, material development, teacher training and school practice (Nekoda, 2003).

Given the distinct ways to approach English instruction in The Netherlands, the articulation between levels seems to be rather complex. Part of this research intends to explore how the teaching of English differs between primary and secondary education in The Netherlands, as well as how these differences have grown historically.

As might be evident from the information above, the differences in the amount of English teaching offered in primary education have grown in the past decade; this can be projected to lead to problems in the articulation between levels. For example, it is clear that when some first year secondary school students have been taught English from the age of four while the rest of the students in their class were taught English in a limited fashion during the final two years of primary education, the language skills of these children on entering secondary education will vary enormously (Bodde-Alderlieste & Schokkenbroek, 2013). These variations in levels of language skills are a challenge to the Dutch first year secondary school English teachers, as they require the teacher to either respond to student differences or remain unmoved. Unfortunately, the challenge of differentiating instruction is often met by disregarding the differences and having all students start at the same point and work at the same pace through the planned curriculum (Oostdam & Van Toorenburg, 2002). Heterogeneous student language levels are a challenge that teachers could possibly turn into additional learning gains for *all* students, if they were to differentiate their teaching (Tomlinson, 1999). In addition to better serving the current students in the Dutch system, the differentiation of English language teaching in secondary education could then take advantage of the enormous amount of time and resources invested in (early) English language teaching in primary education in the past decade without detracting from the attention needed by those who did not benefit from early

English teaching. Differentiation, in this case responding to the diversity in language skills levels, means that “teachers systematically plan for instruction that is responsive to a predictable range of human variance” (Tomlinson, 2004, p. 519). However, as of yet the varying English language skills levels of first year students in secondary education has not led to a differentiated teaching of English (Holdinga, 2007). Although quite a number of publications on differentiation in education are available and teachers are acquainted with the concept, only a limited number of teachers put it into practice (Latz, Speirs, Neumeister, Adams, & Pierce, 2008; Onderwijsraad, 2008), both in The Netherlands and internationally. Whether the lack of differentiated teaching in Dutch secondary education is due to the lack of teacher’s skills, the absence of differentiated programmes and good practice, unfamiliarity with credible research on the effectiveness of the approach, or something else, is unclear.

The situation described above, related to the growing diversity of English language skills levels of first year of secondary students in The Netherlands, combined with the apparent absence of differentiated education, leads to the following research question:

***How and to what extent do English teaching in the final year of primary school and first year of secondary education differ and what is the effect of differentiated English language teaching on the student’s attitude towards learning English and the development of English language skills in the first year of secondary education?***

### **1.1. Sub-questions**

To answer the main research question, the following sub-questions were formulated:

1. *How has English language teaching in primary and secondary education in The Netherlands been introduced and how has it developed historically?*

The historical developments in English language teaching in The Netherlands can shed a light on the apparent difference between English language teaching in Dutch primary and secondary education. Background information on how differences have arisen, or come into existence, is needed to clarify what has helped, or impeded, the articulation between levels.

2. *How does the teaching of English in the Netherlands differ between the final year of primary education and first year of secondary education?*

A closer look at the differences in approach to the teaching of English is needed to pinpoint the methodological activities that make the transition for students harder. These differences can help in the construction of a language teaching programme based on differentiated instruction plans, as well as help initiate discussions on the articulation between levels.

3. *Which aspects should a language teaching programme based on the principles of Differentiated Instruction for the first year of secondary school include, to ensure knowledge and skills acquired in primary school are used to their full potential and enable execution of activities on different levels?*

A language teaching programme based on differentiated instruction offered to secondary school teachers must be research-informed as well as grounded in practice. The aim of developing the language teaching programme based on differentiated instruction, is to find a sustainable solution for working with a wide variety in the level of language skills.

4. *To what extent does a language teaching programme based on differentiated instruction effectively enhance the attitude towards learning English and increase the learning gains of first year students in secondary education?*

The effects of a language teaching programme based on differentiated instruction must be measured to allow evaluation of the programme. The attitude as well as the language skills were measured twice to enable the effects to be reported.

5. *Which skills should (future) teachers of English in secondary schools acquire or possess to offer a language teaching programme based on differentiated instruction as mentioned above?*

The information on the effectiveness of the programme itself needs to be accompanied by information on anything that helped, or impeded the implementation of the programme. When the programme helps to deal with the variety of language skills levels, teachers need to know what to take into account when they start differentiating their language teaching.

## 1.2 Sub-studies

The research questions will be answered by five sub-studies; two literature reviews, a survey, a quasi-experiment, and a Delphi study. To answer sub-question 1, a literature review entitled: *'Historical context of the teaching of English as a foreign language in primary and secondary education.'* will be described in chapter 2. To answer sub-question 2, a questionnaire survey on *the differences in English language teaching in primary and secondary education in The Netherlands*, will be described in chapter 3. To answer sub-question 3, a literature study on *differentiated instruction: Historical context(A), Theoretical foundations (B), Building blocks (C), the role of teachers (D) and the effectiveness of differentiated instruction (E)*, will be described in chapter 4. The findings of this literature study will be used to construct a language teaching programme based on differentiated instruction, which will be described in chapter 5, to be used as treatment in the quasi experiment. To answer sub-question 4, a quasi-experiment will be performed, as will be described in chapter 6. The measurement of the effect of the programme leads to additional research questions, which will also be introduced in chapter 6. To answer sub-question 5, as well as to clarify the factors impeding differentiated language teaching and the implementation of the treatment offered, a Delphi study, as described in chapter 7, was carried out.

The following chapters will first clarify the differences in English language teaching between the final year of primary education and first year of secondary education by looking both at the history (chapter 2) and the current practice (chapter 3). Differentiated instruction is examined (chapter 4) and the findings used to develop a language teaching programme based on differentiated instruction (chapter 5). This differentiated programme is used by secondary school teachers and the effects of the programme on student learning gains and attitude are measured in a quasi-experiment (chapter 6). Schoolteachers' experiences with the programme are turned into advice and prerequisites for future implementation of differentiated instruction (chapter 7).

## 2. Historical context of the teaching of English as a foreign language in primary and secondary education.

### 2.1. Introduction

The literature study presented in this chapter presents a historical review of developments in the teaching of English as a foreign language in primary and secondary education. This historical review will help answer the first sub-question, *'How has English language teaching in primary and secondary education been introduced and how has it developed historically?'* The review looks at foreign language teaching in general in the western world and English language teaching in particular, with special attention to developments in The Netherlands. The findings are intended to clarify the roots of the differences in approach to English language teaching in primary and secondary education in The Netherlands.

### 2.2. Method used to search for literature on the historical context of the teaching of English as a foreign language in primary and secondary education.

A literature search of the journals, dissertations, articles and books in the 398 databases of CataloguePlus of Amsterdam University, was conducted. The use of this approach helped limit the hits to publications that are of methodological quality. The keywords used in the search initially related to publications on differentiation of education and English language teaching ('education' or 'curriculum' or 'teaching' and 'English' or 'language'), with the limitation of full text availability, which yielded 7010 hits. Synonyms for key words were also used (for example, 'education' in place of 'teaching'). Additional words that were used included; 'curriculum', 'teaching', 'English', 'English language instruction,' and 'language' which yielded 2294 hits.

The historical perspective on English language teaching in The Netherlands, turned out not to be a widely researched topic, with 'History of English language teaching' and 'Netherlands' or 'Dutch' yielding only 17 hits.

In addition to the digital library search, the so-called snowball method has been used, where key documents lead to citations or references of other sources on the same subject. The list of pivotal

moments and movements in history found in the initial search and the follow-up search for publications on these topics, are an example of this approach. To determine which articles were included for further analysis, the following criteria were used:

- The study needed to report on the developments and changes in English language teaching in general. Studies on historical developments and changes that did not have an international character, i.e. outside of the Dutch perspective, were excluded.
- The study needed to report on English language teaching in primary and secondary education, as this is the scope of this study.
- Both quasi-experimental and case studies and both quantitative and qualitative studies were included, as long as the method of differentiation was described in a sufficiently elaborate and transparent manner.

Applying these criteria to the initial set of articles resulted in a selection of 68 titles. As this study is not a systematic review of literature at a specific moment in time, with an added extensive snowball method effect, the literature used has been extended during the five years of this study.

### **2.3. Results of the search for literature on the historical context of the teaching of English as a foreign language in primary and secondary education.**

#### **2.3.1. Historical Development of English language teaching in primary and secondary education in The Netherlands.**

##### **Nineteenth century education and foreign language teaching**

A significant historical review of English language teaching in The Netherlands begins in the 19th century.

This century was chosen because it was during the nineteenth century that education was organised in the form of schools and year groups, as is largely still current practice.

The 19th century brought a lot of changes to the educational system of The Netherlands as a whole as well as to foreign language teaching in particular (Boekholt & de Booy, 1987).

Publications on English language teaching methodology used in The Netherlands in the first half of the nineteenth century do not mention any prescribed or advised method for general education. This is in itself

not surprising since the Dutch government did not prescribe how education was supposed to be delivered and languages were taught in the same way that Latin and Greek had been offered in the past centuries. Primary education in The Netherlands has already been organised by law in 1857, before a law on the organisation of secondary education was passed in 1863. These laws deal primarily with the structure of education; the state leaves the organisation of schools and schooling to municipalities, religious organisations and non-governmental institutes (Boekholt & de Booy, 1987; van Els, 1992). The content of what was being taught and the methodology used was largely left to the teachers. Foreign language teaching was no exception to this rule.

Foreign language teaching in The Netherlands around 1880 did not have a consistent appearance (van Els, 1992; Wilhelm, 2009). Without universities that offered a degree in foreign languages, and given the non-existence of teacher training institutes and the absence of methodological direction from the government, ‘...teachers were free to choose their own methods and materials.’ (Wilhelm, 2009, p. 4).

Language teachers in general not only lacked external guidance or direction; they were also not internally organised. They did not share their views in journals and teacher associations only emerged in the final decades of the 19<sup>th</sup> century (Wilhelm, 2009).

### **Grammar-translation**

Despite the lack of guidance or shared platforms, practically all foreign language teachers used the grammar-translation method until the end of the nineteenth century. This method, also referred to as the ‘Classical method’, was based on the centuries-old practice of learning the classical Latin and Greek languages. The method consisted of memorisation and application of grammatical rules, syntactic structures, complemented with rote memorisation of vocabulary and translation of literary texts (Larsen-Freeman & Freeman, 2008; Smith, 2005; Wilhelm, 2009).

The world changed in the 19<sup>th</sup> century and The Netherlands changed with it. Dutch society changed from agricultural to industrial and this was reflected in mass public schooling that resembled a factory model in which children are taught in year groups and are all given the same treatment (Stoll Lillard, 2005).

Perhaps the industrial revolution and its growth in scientific research and progress instigated a search for the best method to teach languages. Or, as Brown (2002, p. 9) puts it, 'a single, ideal method, generalizable across widely varying audiences, that would successfully teach students a foreign language in the classroom.'

This scientifically based approach to language teaching concerned multiple areas. For instance, language teaching reformers like the Irishman Prendergast (1806-1886), and the Frenchmen Marcel (1793-1896) and Gouin (1831- 1896) were the first to refer to child language learning as a model for teaching languages and to suggest that a syllabus be developed based on their findings (Richards & Rodgers, 1986). British phonetician, Sweet (1845–1912) was instrumental in the development of the notion of the phoneme or distinctive sound. As a member of the International Phonetic Association (IPA), founded in 1886 by Passy (1859–1940), Sweet contributed to the International Phonetic Alphabet (IPA) designed to accurately transcribe the sounds of any language. This shift in focus from historical changes in languages over time, to the study of language as a self-contained and structured system, can be considered to be the beginning of Modern Linguistics. The subsequent suggestions for language teaching, generated by the developing linguistic research, were welcomed by reform-minded teachers (Richards & Rodgers, 1986).

Throughout this study the terms 'approach' and 'method' have been used in the same way they were found in the research referred to. A historical review of 'approach' and 'method' is beyond the scope of this study, but when they are used to discuss or describe without reference, the Richards and Rodgers (1986) model has been used. According to Richards and Rodgers (1986), approach and method are hierarchically related in that they inform as to the views on language teaching, albeit on different levels. In their view, an approach refers to theories about the nature of language and language learning that serve as the source of practices and principles in language teaching. A method consists of the practices and principles, organisationally determined by a design based on the theories of the approach, and is practically realized in procedure (Richards & Rodgers, 1986).

The above paragraphs clearly show that the major changes in language teaching were related to, if not a consequence of, the scientific progress made in the Western world during the second half of the nineteenth century. In the USA, for instance, the belief that psychology would revolutionize education led

to the formation of the American Psychological Association (APA) in 1892 (Llera & Sánchez, 2011). At the time people in general believed science would help better the world and in that sense viewed it differently than they had in previous centuries (Knoll, 2009). Science as we know it today stems from activities and work viewed at the time as being part of contributions to broad philosophical, theological and moral concerns. When William Whewell coined the term 'scientist' in Britain in 1833, most important British scientists of the time still preferred to refer to themselves as 'natural philosophers' (Yeo, 2003). Research concerning education and newly developed methods to teach (foreign) languages, slowly reached the language teachers in The Netherlands. For instance, in 1880 Gouin published his innovative method of acquiring foreign languages through reproduction of situations in which a specific use of language is demanded (Vonk, 1993). Fourteen years later proof of discussions on topics like this were to be found in the Netherlands, in the 1894 meeting of the Secondary Schoolmasters Association (Wilhelm, 2009).

### **Reform movement (Europe)**

At that time a more or less organised movement for methodological change influenced language teaching (Van Els, 1992) in the surrounding countries. This movement, also referred to as the 'Reform movement' (Wilhelm, 2009) is said to have begun in 1882 with a publication by Wilhelm Viëtor (Smith, 2005).

Viëtor was a teacher of English who, in his writing, attacked the excesses of the grammar-translation method used for modern language teaching in German schools. Although Viëtor's publication may have been the signal for the reform movement, dissatisfaction with the method had been around for quite some time (Howatt, 1982).

This dissatisfaction with the grammar-translation method gained momentum through publications by distinguished scholars like Sweet; support from the new science of psychology; one of the first attempts to develop a scientific theory of language learning by Franke and the foundation of an association of English teachers in France (the 'Phonetic Teachers' Association') in 1886. These leaders of the Reform movement agreed on a set of basic principles; teachers should:

- give priority to the spoken language;
- no longer have students translate whole texts in writing;

- use connected texts rather than isolated sentences and
- only teach grammar inductively, by presenting examples of language to generate rules and make generalizations. This inductively offered grammar teaching contrasted to the grammar-translation method, where students were offered grammatical rules and had to apply them by translating sentences between the target and source language.

These innovative ideas and publications (Thiergen, 1900) from around Europe triggered individual teachers to publish and debate new impulses such as the Berlitz approach, the Direct Method or the Gouin method, all exponents of the reform movement principles mentioned above. The lack of organisation of leading institutes, however, prevented these initiatives from growing into a sizable and concerted Dutch Reform movement in language teaching (Wilhelm, 2009).

In the Netherlands perhaps the most tangible result of this European movement for foreign language teaching was the addition of an oral exam to the exam programme of secondary schools (Higher Burger Schools or HBS), while the earlier mandatory test of grammar rule knowledge was made obsolete (Wilhelm, 2009). This HBS was one of the two possible secondary schools described in the law on secondary education in 1863 in preparation for a professional career. The other was the grammar school which, unlike the HBS, offered Latin and was the preparation for university (van Els, 1992).

### **Progressive movement (USA)**

The Reform Movement in Europe coincided with the Progressive Movement in the United States and both movements can be seen as part of the popular 19<sup>th</sup> century belief that scientific progress would help change the world. All the scientific advancements suggested science would help change language education, through the development of the best method, to successfully teach all students a foreign language (Brown, 2002).

Where the Reform Movement endeavoured to change (language) education, the Progressive Movement comprised a popular effort to 'insure the survival of democracy in the United States by the enlargement of governmental power to control and offset the power of private economic groups over the nation's institutions and life' (Link, 1959, p. 836). In reaction to corporate abuse, factory workers started a social

reform movement, while the middle class tried to eliminate corruption in government and improve efficiency through modernisation, emphasizing scientific, medical and engineering solutions. These movements together are called the Progressive Movement, although the individual reactions and movements mentioned never cooperated on a nation-wide scale (Filene, 1970).

Language teaching moved away from classical rote memorisation for the transfer of knowledge towards a more naturalistic approach. Educators and language teachers based their teaching on observations of how children learned skills and their mother tongue. Pedagogical and psychological ideas and theories were tried and tested beyond the laboratory (schools).

### **Behaviourism**

Thorndike (1874-1947) was the first modern psychologist who brought a scientific approach to the study of learning (Ostrowska, 2014). His approach to education was quite detached. He distrusted emotions and believed mankind could be perfected by a scientifically-discovered truth (Berliner, 1993). Thorndike took the idea of psychology as a purely objective experimental branch of natural science, with as its goal the prediction and control of behaviour, from Watson (1878-1958), also known as the father of American behaviourism (Berliner, 1993), and applied it to education. Or, as Thorndike said: "The nature of the pupil as well as the nature of the stimulus decides his response. To arouse, direct and select from his responses is the work of the teacher." (Thorndike, 1906, p. 41).

Behaviourism has had an enormous impact, not only on education. It was thought that applying the methods of the natural sciences to the control of human behaviour, would help create future citizens for a modern more efficient and ordered state. According to Tomlinson (1997), the legacy of this rather mechanical model of mind and society can still be found in the use of behavioural objectives, drills, intelligence testing, achievement scales, tracking, and vocational training (Tomlinson, 1997).

Apart from behaviourism, the educational innovations mentioned did not immediately change every classroom and teaching practice, not even in the Western world. They did, however, influence education in the sense that different approaches, methods and techniques, based on scientific progress, were applied to

teaching practice. The drive to change education had moved from theoretical publications to classroom practice.

### **The Oral Approach and Situational Language Teaching (Great Britain)**

Applied linguists in Britain, like Palmer and Hornby, developed the basis for a principled approach to methodology in language teaching. They endeavoured to strengthen the scientific foundation, described for the Direct Method, for an oral approach to teaching English and did so through "...a systematic study of the principles and procedures that could be applied to the selection and organisation of the content of a language course." (Richards & Rodgers, 1986).

Although their developed approach to foreign language teaching methodology was based on the Direct Method, it differed in the sense that it involved systematic principles of:

- selection, of lexical and grammatical content;
- gradation, of content through organisation sequencing; and
- presentation and practice, of items in a course.

These principles were referred to as the *Oral Approach* to language teaching. This approach grew in Britain until it was the accepted British approach to language teaching in the 1950s. In the following decade it was extended and further developed. One of the principles added was that students needed to practice 'situationally', meaning that the language had to be taught linked to the real life situation in which it was used. This approach was aptly named *Situational Language Teaching*, due to its emphasis on situations, and from 1965 onward became an established name. In particular, because of language teaching materials known as the '*Situational English*' series, developed and published for worldwide use by the Commonwealth Office of Education.

### **The Audio-lingual Method (USA)**

Foreign language teaching in the US had not progressed far beyond the Direct Method during the first half of the twentieth century (Richards & Rodgers, 1986). That may be because the incoming waves of immigrants from Europe at the time already spoke their 'foreign' languages. The (American) English used in

the United States might have been foreign to the immigrants, but it was the language the immigrants usually acquired on the job. This changed briefly when the US entered the Second World War in 1942 and the government needed fluent speakers of languages used in the countries involved. Fifty-five American universities were involved in the development of the Army Specialized Training Program (ASTP), which consisted of short and intensive, oral-based courses. Although the ASTP methodology and successes impressed the linguists in America, this did not lead to widespread application of the methodology. The launch of the first Russian satellite in 1957 prompted the U.S. Government to take a number of measures to prevent Americans from becoming isolated from scientific advances made in other countries. One of the measures taken was funding for intensive research and development of foreign language teaching. Language teaching specialists drew on the ASTP experiences, the British approach and insights from behaviourist psychology, the combination of which led to the Audio-lingual Method. The Audio-lingual Method embodied the perspective that the content of language can be linguistically broken down and made part of the student's behaviour through drills and pattern practice (Richards & Rodgers, 1986). The period of the method's most widespread use in the US was in the 1960s. It was also exported around the world.

The developments, as described, did not immediately have an effect on (language) teaching in the Netherlands. Chronologically speaking, a milestone in language teaching in the Netherlands was reached in the early 1960s. From 1962 to 1965, the language project (Dutch: Talenproject) studied the effect of the use of achievement tests instead of literally translating during the final exam (Gras, 1969; Staatsen, Heebing & van Renselaar, 2009). These shorter exams, which were again a step away from the grammar-translation method, worked well for the new school types planned by the Dutch government in the new educational system. The achievement test contained items for vocabulary and idiom, grammar and reading comprehension and '...proved to be suitable substitutes for the translations.' (Gras, 1969, p. 168).

### **Mastery Learning**

In the 70s, the Audio-lingual method reached The Netherlands and it revived the discussion on the use of everyday speech and grammar in language teaching. A lot was expected of a new technique called 'Mastery

Learning'. Mastery Learning had been developed by Benjamin Samuel Bloom (1913 –1999). Although this strategy was based on behaviourist principles, it is of interest to this study because of Bloom's view on differentiating instruction (Guskey, 2005). Influenced by, amongst others, the Winnetka Plan of Washburne (cf. above), Bloom recommended using assessments at the end of each unit as part of the instructional process, to diagnose individual learning difficulties (feedback) and to prescribe individual remediation procedures (correctives). This strategy would give teachers the practical means to vary and differentiate their instruction in order to more effectively meet students' individual learning needs. With this strategy Bloom hoped to eliminate the achievement gap and drastically reduce the variation in students' achievement level, by increasing the variation in instructional approaches and learning time (Guskey, 2005; Reigeluth, 2012).

Mastery Learning and the Audio-lingual method seemed to work well together in The Netherlands. In effect, students were required to wear headphones in 'language labs' and mimic, drill and repeat language, often without comprehension (Brown, 2002). Methodologically speaking, the technological advancement into the 'language labs' did not really revolutionise the character of foreign language teaching. Although the assignments were (mostly) performed individually, the activities, vocabulary and grammatical structures students needed to work through remained the same for all. The advancement towards differentiation was the option to individually pace the work.

The result was the wide acceptance that the application of grammar rules is more effective than learning them by heart. On top of that, translation assignments were changed into cloze exercises, and listening skills, which were non-existent to some Dutch foreign language teachers prior to 1970, were here to stay (Wilhelm 2009). The different language skills that now had become part of language teaching, combined with the Audio-lingual method, provided Dutch language teachers with tools to offer scaffolding or challenge to both struggling and advanced students.

### **The Communicative approach**

The last approach to have affected language teaching was the communicative approach. This approach emerged in the English speaking world in the 70s but impacted Dutch education in the early 80s.

The behaviourist approach to language teaching had been prevalent<sup>2</sup> for more than half a century and had developed and progressed into neo- and radical behaviourism (cf. Chomsky, 1959; Celce-Murcia, 2007).

However, discontentment of language teachers and applied linguists in the US with this approach to language teaching grew at the end of the 60s, and the beginning of the 70s.

Discontentment with structurally competent students who remained communicatively incompetent; unable to perform a simple communicative task, led to the insight that the ability to manipulate the structures of the language correctly is only a part of what is involved in learning a language (Swan, 1985; 2007). According to Swan (1985) this discontentment made the language teachers and applied linguists receptive to the criticism towards behaviourist approaches to language teaching, as well as to the following developments towards a communicative approach.

The name *communicative approach* is based on the term *communicative competence* coined by Hymes (1972), in reaction to Chomsky's focus on linguistic competence (the rules for describing sound systems, combining sounds into morphemes and morphemes into sentences) in which the consideration of social factors was beyond the domain of linguistics (Celce-Murcia, 2007).

In Hymes's view, for the acquisition and competent use of languages people needed to gain *sociolinguistic competence* (the rules for using language appropriately in context) as well as linguistic competence (Celce-Murcia, 2007). Or, as Hymes put it: "rules of use without which the rules of grammar would be useless" (1971, p. 278).

In the following decades the 'strategic competence' (i.e. the ability to compensate for problems or deficits in communication) and discourse competence (the ability to produce and interpret language beyond the sentence level) have been added to the model by Canale and Swain (1980) and Canale (1983) respectively.

---

<sup>2</sup> Prevalence of the behaviourist approach should be seen in the light of the model of Richards and Rodgers (1986). Since the beginning of the twentieth century behaviourism had been the leading learning theory. On this general approach other language learning theories and methods had been developed and applied. For the Oral Approach or the Audio Lingual method, for instance, behaviourism had remained the foundation and through all these developments grammar and translation always played an important role.

This shift towards communicative competence in the United States and Canada coincided with developments in Europe.

### **The Common Europe Framework of Reference**

The Council of Europe, founded in 1949 for the defence of human rights, parliamentary democracy and the rule of law, worked with the member states towards agreements on standardised social and legal practices and awareness of the shared values that make up a European identity (Little, 2006).

This led to the European Convention on Human Rights in 1950 and the European Cultural Convention of 1954 in which the second article focussed on the need for all member states to encourage the study of each other's languages, history and civilisation. In their view, especially the languages and English as *lingua franca* were indispensable in the needed mutual understanding, effective educational and cultural exchange and the envisioned mobility of European citizens. Since the early 1970s, the work of the Council of Europe on language policy and language education had two main concerns: the analysis of students' communicative needs and the description of the language needed to fulfil those needs (Little, 2006). To work on those concerns, a group of experts started to investigate the possibility of developing language courses on a unit-credit system: a system in which learning tasks are broken down into "portions or units, each of which corresponds to a component of a student's needs and is systematically related to all the other portions" (Van Ek, 1975, p. 12). The group used the analysis of the European student needs by British linguist Wilkins (1972). Wilkins described the communicative definition of language by dividing the meanings that lay behind the communicative uses of language into functional and notional categories (Wilkins, 1972).

The analysis of communicative needs and the description of functional and notional categories of languages used, led to the definition of a 'threshold level' of communicative proficiency in a foreign language.

Described in 'Threshold Level English in a European Unit/Credit System for Modern Language Learning by Adults' (Van Ek, 1975) and the elaboration and promotion of the concept of autonomy in foreign language learning. The Council of Europe incorporated the threshold level specifications which have had a great impact on language programme design and textbooks published in Europe.

Work of the Council of Europe and publications on the theoretical basis for a communicative or functional approach to language teaching, helped to shape the Common European Framework of Reference (CEFR), published in 2001, as well as its companion piece, the European Language Portfolio (ELP). The CEFR is a framework that describes foreign language proficiency at six levels. These six levels are respectively higher and lower interpretations of the classic division into basic, intermediate and advanced. The Council of Europe adopted a branching principle, starting from the initial division into three broad levels: A (Basic User), B (Independent User) and C (Proficient User), subdivided in A1 (Breakthrough), A2 (Waystage), B1 (Threshold), B2 (Vantage), C1 (Effective Operational Proficiency) and C2 (Mastery).

Descriptions of the foreign language proficiency in these six levels are formulated as 'can-do' statements for the four language skills (listening, speaking, reading and writing, although speaking is described as two different skills: dialogues and monologues (Verhelst, Avermaet, Takala, Figueras, & North, 2009).

The Communicative approach gained national and international prominence through the swiftness with which textbook writers, curriculum development centres and governments applied the principles for communicative language teaching, the CEFR and ELP (Wilkins, 1972; Little, 2006).

In 1986, Dutch secondary school exams saw the introduction of speaking skills. During the final exam marks were to be obtained for speaking itself, where previously, the oral exam consisted of a teacher and student discussion on (obligatory) literature. At the time, secondary school foreign language teachers, methodologically-speaking, attempted to differentiate by dividing the subject matter in basic, deficiency and challenging materials based on Mastery Learning. This reorganisation of exercises and materials resulted in stronger emphasis on grammar and idiom ((Hulshof, Kwakernaak, & Wilhelm, 2015)).

### **English language teaching in Dutch primary education (ten to twelve-year-olds)**

In 1986, the same year speaking skills were introduced in the final exams of secondary education, English was introduced as a compulsory subject for the final two years in primary education in the Netherlands (ten to twelve-year-olds). The introduction of a modern foreign language in the primary school programme had been in preparation for some time.

In the sixties, the introduction of English in primary schools had been proposed and in 1968, the "Utrecht Eibo<sup>3</sup> Project" was started. The project, whose activities were hosted at the National Institute for Curriculum Development in the Netherlands (SLO) in 1978, developed a method for teaching English in primary schools. The project explicitly opted for a communicative approach, in which the communicative competence of the students was the ultimate goal (Engel, Trimbos, Drew, & Groot Wilken, 2007). This communicative approach meant children were taught to use English dialogues in everyday, age appropriate, situations, selected from the 'Threshold Level English in a European Unit/Credit System for Modern Language Learning by Adults' (Stoks & Voortman, 1982). Because of the emphasis on the use of dialogues as means to communicate in English, the focus was on listening and speaking skills.

Since its introduction in the Dutch primary school curriculum, English has also been a compulsory subject of the primary school teacher training (Pabo<sup>4</sup>) curriculum in The Netherlands. In the period round the introduction (1984-1987), teacher trainers received in-service training at a number of secondary school teacher training institutes (NLO) to enable them to teach English at the Dutch Pabos.

Subsequently, educators from Pabo, NLO and secondary education, trained approximately 16,000 primary school teachers in how to teach English. The course consists of lectures, television broadcasts by Teleac<sup>5</sup>, radio broadcasts and course materials for independent study (Koster, 1986; Van Toorenburg & Bodde-Alderlieste, 2003). However, the Dutch government did not give this training a mandatory character and in 1985 the government declared all in-service primary school teachers qualified to teach English in primary education (Boer, 2003).

---

<sup>3</sup> 'Eibo' is a Dutch acronym that stands for: 'Engels in het basis onderwijs' literally meaning 'English in primary education' indicating the mandatory English language teaching to ten to twelve-year-olds.

<sup>4</sup> 'Pabo' is a Dutch acronym that stands for 'Pedagogische academie voor het basisonderwijs', literally Pedagogical academy for primary education; the primary school teacher training (bachelor) programme of the school of education of a university of applied sciences.

<sup>5</sup> Teleac: short for Television Academy, the Dutch national educational broadcasting organization from 1963 to 2010.

Since 1987, the Dutch National Institute for Educational Measurement (Cito) has carried out the Periodic Survey of Educational Achievement in primary regular and special education (PPON<sup>6</sup>). The PPON reports on virtually all learning areas in primary education in The Netherlands. In 1991, the first survey of English in primary schools was conducted (Vinjé, 1993). This survey was repeated in 1996 (Edelenbos, Schoot & Verstralen, 2000), 2006 (Heesters, Feddema, Schoot, & Hemker, 2008) and 2012 (Geurts & Hemker, 2013). All four surveys find English language teaching is officially on the timetable of the final two years of all primary schools in The Netherlands. Most Dutch primary school teachers (90%) teach English weekly. The average time spent on English teaching has grown 20% in 2012 compared to 2006. This increase is noteworthy, because time spent on English teaching had remained practically the same since 1991. Schools would teach one English language lesson a week for about forty-five minutes, but in 2012 that time had risen to 54 minutes (Geurts & Hemker, 2013). The differences between schools are substantial. The lowest number of hours spent on English teaching is 7 and the highest 90.

As in 1991, 1996 and 2006, in more than 90% of the schools English is taught by the class teacher.

Approximately half of the teachers indicate in 2006 and 2012 that they received English education at the Pabo. This is considerably more than in 1991 (0%) and 1996 (8%). In the earlier polls, 57% of teachers said they had taken an Eibo in-service training course; in 2006 this dropped to 14% and in 2012 to 2%.

Furthermore, approximately one third of teachers say they have had no specific training in English teaching whatsoever, pre-service or in-service.

### **Early English language teaching in The Netherlands (four to twelve-year-olds)**

According to the English attainment targets of the ministry, two years of English language teaching (ten to twelve-year-olds) in primary education are mandatory, but some schools offer more English than is expected by the ministry. Until 2000, the number of schools (about 20) that offered Early Foreign Language

---

<sup>6</sup> PPON is a Dutch acronym for Periodieke Peiling Onderwijs Nederland; literally translated: Periodic Survey Education Netherlands.

Teaching (VVTO<sup>7</sup>) English remained fairly constant (Europees Platform, 2010). After the turn of the century the number of schools that offered VVTO English grew tremendously. In 2001 that was more than 20 schools, in 2004 more than 45 and in 2007 more than 100 schools. In 2011 the number of VVTO schools was 448 and in 2016 the number of schools known to the European Platform to offer VVTO English had risen to almost 1000. Currently, a small number of schools offer English from the age of eight, but almost all schools start their VVTO English programmes with the four-year-olds in the first year of Dutch primary education (European Platform, 2016.).

### **Early English language teaching in Rotterdam**

The largest centre of growth of VVTO English teaching in the Netherlands is in the city of Rotterdam and the surrounding region because of the EarlyBird centre. The EarlyBird centre came into existence in 2003 as an initiative of the 'Stichting Bestuur Openbaar Onderwijs Rotterdam' (BOOR). BOOR is the foundation responsible for public teaching in the municipality of Rotterdam; its EarlyBird centre, which was founded to help primary schools teach English to young children, was established as a follow-up of the success of BOOR's bilingual secondary education.

The EarlyBird centre began by guiding and supervising a pilot for the development of 'more, better and earlier English' in primary education in two primary schools. Over the years, the EarlyBird methodology has been further developed, guided both by experience and scientific research. The EarlyBird centre is now a nationally operating centre for expertise that works with various colleges, universities, school boards and school advisory services (EarlyBird, 2016). The EarlyBird centre helps schools at different levels with the introduction of the subject of English. The 'EarlyBird' primary schools preferably work with (near) native speakers or primary school teachers with additional training in the English language. English is part of a school's overall language teaching policy and schools are reviewed by the EarlyBird centre every four years using a set list of behavioural indicators to guarantee a constant quality. In Rotterdam, 28 of the 68 primary

---

<sup>7</sup> in the Dutch language *Early Foreign Language Teaching* is called 'VVTO', the acronym stands for 'vroeg vreemde talen onderwijs', which is used by schools that offer English language teaching from the age of four.

schools work with the EarlyBird methodology. In the rest of the Netherlands, this is the case for 109 of the approximately 7000 schools as well as 67 nurseries and day-care centres (EarlyBird, 2016).

Investments and activities in primary school English language teaching, like that of EarlyBird, have a positive effect on the regional teacher training institutes. After the success of the pilot described above, the primary school department of the School of Education of the Rotterdam University of Applied Sciences, for instance, sought an alliance with the EarlyBird centre in 2006 and started to cooperatively contribute to the development of VVTO English in Rotterdam and The Netherlands. In cooperation they developed additional English language teaching methodology courses, an 'Early English' minor and in-service training courses, which have been introduced into the regular curriculum (Corda, Kraay & Feuerstake, 2012).

Since 2010, both EarlyBird and the Rotterdam University collectively participate in the national VVTO platform (platform VVTO) together with three other teacher training institutes from across the country. This nationwide cooperation led to a national standard for VVTO – specialist primary school teacher, who is, for instance, expected to have a language skills level of at least C1 (Effective Operational Proficiency) of the Common European Framework.

### **Articulation between levels of primary and secondary school English language teaching**

From the outset, the articulation between the primary and secondary education level of English language teaching in the Netherlands has been problematic. Some problems were anticipated as is demonstrated by a publication from the SLO in 1984, two years before English teaching in primary education was formally introduced. The publication informed English teachers in the first year of secondary education about Eibo, including suggestions on how to fit subject matter to expected pre-knowledge and acquired skills. These suggestions were not very successful as they turned out to be difficult to attune to the course books in use in secondary education at the time (Van Toorenburg & Bodde-Alderlieste, 2003). In fact in 1991, only 5 years after its introduction, the Dutch schools' inspectorate advised the ministry of Education that unless specific measures were taken to improve the articulation between levels of English teaching as a subject, it

should seriously reconsider the continuation of Eibo<sup>8</sup> (Inspectie van het Onderwijs, 1991, p. 17). It was feared that Eibo would experience the same fate as 'Primary French' in England and Wales, which failed due to faulty articulation of primary and secondary school levels of French language teaching (Oostdam & Van Toorenburg, 2002).

The situation hardly improved despite the ministerial curriculum standards, or educational objectives (Dutch: 'kerndoelen'), formulated in 1993; revised in 1998 and 2005. Research from the period shows that no course books with a continued subject matter strand for English language teaching were available. Secondary school teachers of English hardly adapted their teaching repertoire to student pre-knowledge and mostly start from scratch (Edelenbos, 1993).

During the revision of the curriculum standards for primary education in 2002, a committee<sup>9</sup> advised the minister of education to keep English out of the compulsory part of the standards (Oostdam & Van Toorenburg, 2002). The ministerial choice to maintain the subject of English as part of the compulsory core of the curriculum standards has not helped the development of an uninterrupted continuation of English language teaching. In 2003, Van Toorenburg and Bodde-Alderlieste describe how articulation between primary and secondary levels of English language teaching is hard to find. Students entering secondary education find that progress for English is mainly measured by written tests in which correct spelling is required, although writing is not a part of the curriculum standards for English in primary education. So, students are confronted with a faulty articulation between levels of primary and secondary school and the implicit message is that acquired language skills are of no consequence to secondary school academic success.

Oostdam and Van Toorenburg (2002) investigated the issue through questionnaire research among 147 teachers in primary and (the first year of) secondary education as well as 295 students in the first year of secondary education. They found English teachers in the first year of secondary education usually ignore

---

<sup>8</sup> 'indien er geen gerichte stappen worden gezet om de aansluiting voor het vak Engels te verbeteren, de wetgever zich ernstig dient te beraden of voortzetting [...] zinvol is' (Inspectie van het Onderwijs, 1991:17)

<sup>9</sup> Curriculum standards primary education committee (Dutch: 'Commissie Kerndoelen Basisonderwijs')

the rather varied levels of proficiency in English of their students. According to Oostdam and Van Toorenburg (2002), this frustrates students who find their education does not match their proficiency levels and gives primary school teachers the idea that their efforts are not being taken seriously. Oostdam and Van Toorenburg (2002) argue this is a vicious circle that needs to be broken. Curriculum standards for English in Dutch primary education have been, and still are, formulated as guidelines, which are difficult to measure and are not used to hold primary schools accountable for the English language proficiency of their students. According to Oostdam and Van Toorenburg (2002), this leads to a lack of professionalization among primary school teachers, a low priority given to English teaching in primary schools and a questionable status of the subject of English language teaching in teacher training programmes for primary education.

Between 2003 and the present, English teaching in primary education has become a dynamic field. The curriculum standards have been reformulated, more than 14% of all primary schools have started to offer (early) English teaching to all students from the age of four onwards, and all secondary final exams have been aligned with the CEFR. In 2014 the ministry of education even started a pilot project for bilingual primary education with 18 primary schools offering English 30% to 50% of the instruction time Bodde-Alderlieste (2015). However, all these developments do not mean the articulation between primary and secondary levels of English language teaching has improved. According to Oostdam (2009), the need for clear and measurable standards combined with the option to hold primary schools accountable for the English proficiency of their students has only grown with the advancement of early English teaching. In that sense Bodde-Alderlieste (2015) suggests primary schools need to make clear choices about their levels of ambition to possibly stratify the levels of English language proficiency with which students leave primary education. As far the articulation between primary and secondary levels of English language teaching is concerned, there is still a long way to go.

## **2.4 Conclusion of the literature study on the historical context of the teaching of English as a foreign language in primary and secondary education.**

The above pages briefly describe the developments in education, with a focus on English language teaching. This history clearly shows that ideas and approaches to language teaching have not developed along clear lines.

It is noticeable to find that current research on approaches in education are essentially recurrences of publications between approximately forty to almost a century old. While society has changed dramatically over the past century and changes are occurring ever more rapidly, education seems to make slow and repetitive moves towards ideal teaching practice.

Although it has been more than a century since Viëtor challenged the Grammar-Translation method, it seems foreign language teaching in The Netherlands has not yet moved consistently towards the communicative approach (Inspectie van het Onderwijs, 1999; Hermans-Nymark, 2006). Despite this apparent rigidity in language teaching methodology (Inspectie van het Onderwijs, 1999; Bonnet, 2004), neither the position of the subject or the methodology in use are static. The position of English as a subject has been reinforced by new legislation regarding the minimal requirements to pass the final secondary school exams (Ministerie van OCW<sup>10</sup>, 2011).

English teaching in primary education in The Netherlands has been around for almost 30 years and has been growing steadily, especially in the past decade. While mandatory as a subject for ten to twelve-year-olds, the number of Dutch primary schools offering English to four to twelve-year olds has grown from 45 to more than 1000, out of a total of around 7000 in the past ten years (Europees Platform, 2016).

The last twenty years has shown that new insights and ideas in English language teaching have emerged in the growing awareness that there is no 'best method' waiting to be found (Adamson, 2004; Gabrielatos, 2002; Kumaravadivelu, 2001; Larsen-Freeman & Freeman, 2008; Tosun, 2009).

---

<sup>10</sup> In Dutch: Ministerie van Onderwijs, Cultuur en Wetenschap (OC&W), literally meaning: Ministry of Education, Culture and Science.

Chronologically sorted methods, approaches, changes and innovations appear to be anything but a clear roadmap from ignorance to blessed applied science. Some authors even find it remarkable that publications found from innovators, sometimes more than a century old (Viëtor 1882), still sound like something that might have been published today (Gabrielatos, 2002; Grittner, 1975).

Despite it being an unclear roadmap, which also serves a *déjà vu* every now and then, the history described above does answer sub-question 1:

*“How has English language teaching in primary and secondary education in The Netherlands been introduced and how has it developed historically?”*

English teaching in Dutch secondary schools has been tried and tested by all educational reforms, changes and approaches described above. Rooted in the tradition of grammar-translation, it has endured the influence from individual reformers, cooperating teachers and theorists in ‘movements’ and even supranational bodies like the Council of Europe. These changes in approach, like the communicative approach in combination with the CEFR, are visible in textbooks and national exams (Little, 2006), although, as has been reported above, the teaching of grammar is still a fundamental part of teaching English as a foreign language (Hulshof, Kwakernaak, & Wilhelm, 2015).

English teaching in Dutch primary schools, on the other hand, was introduced with a project that explicitly opted for the communicative approach. Teacher training and textbooks have been fashioned accordingly. On top of that, English language teaching as a subject is only a small portion of the curriculum, which is offered completely by the classroom teacher. This means the English language teaching methodology is part of the wider pedagogical and methodological vision of primary education.

These historically explicable differences in current English teaching approaches should clearly show in the time spent on the language skills during teaching. In primary education the focus will be on speaking as product, with listening, reading and writing to support this. Primary education will probably use more songs, (language) games and role playing to support language acquisition and practice speaking, because of the communicative approach applied. Secondary education has a more academic focus and will probably spend more time on grammar and straightforward vocabulary acquisition, the tests of which will be seen as product of English language learning.

### **3. Questionnaire Survey on the differences in English language teaching in primary and secondary education in The Netherlands.**

#### **3.1. Introduction**

The literature study of chapter 2 above describes a divide between a relatively recent, more communicative, English language teaching approach in primary education and a more traditional English language teaching approach in secondary education in The Netherlands. The apparent distance between the English language teaching approaches used in these two school types is maintained, or even increased, by developments in both types of education.

The educational system of The Netherlands comprises eight years of primary education and four, five or six years of secondary education dependent on the stream students are in. Children in the Netherlands are allowed to attend primary school from the age of four; attendance becomes mandatory at the age of five. During primary education, pupils are in mixed-ability classes. Primary education lasts eight years, after which children, at the age of twelve, go on to secondary education. The secondary schools in the Netherlands are stratified into four mainstream school types:

1. schools for students with special education needs (LWOO – 4 years);
2. vocational education on four sub-levels (VMBO/MAVO – 4 years);
3. higher general education (HAVO – 5 years); and
4. pre-university education (VWO/Gymnasium – 6 years).

Students are selected for one of these mainstream school types on advice from the primary school. This advice is based on a pedagogical educational report and a standardized test. As Boer, Minnaert, and Kamphof (2013) describe, the selection at the age of twelve and the stratification is what distinguishes the Dutch educational system from other systems internationally.

In primary education the recent development of Early English language teaching has enhanced earlier established inconsistencies in the English language teaching articulation between primary and secondary education levels. English language teaching offered in primary education differs between schools in quality and quantity, staff, materials, approaches and attainment targets, as schools are relatively free in what is

offered as English language teaching as well as how. In secondary education the apparent rigidity in the development of applied language teaching methodology prevents a consistent move towards the communicative approach propagated by publishers, teacher organisations and the ministry. This description of a discontinuous learning-teaching trajectory for English language teaching in the consecutive steps of general public education in The Netherlands leads to sub question (2): ***“How does the teaching of English in the Netherlands differ between the final year of primary education and first year of secondary education?”*** Differences in English language teaching should show in the time spent on activities as well as the scope and frequency of activities used to teach English. This research endeavours to find differences on item level as well as discern possible clusters of activities that covary and which might indicate distinguishable types of language teaching methodology used.

### **3.2. Method used in the questionnaire survey on the differences in English language teaching in primary and secondary education in The Netherlands.**

#### **3.2.1. Design**

A cross-sectional survey design was used. In this survey teachers from the final year of primary schools and first year English teachers of secondary schools completed questionnaires on teaching time spent on the four skills and the scope of the applied language teaching activities in the classroom.

#### **3.2.2. Sample selection and characteristics**

##### **Sample selection – sample 1 – Primary Education**

From the ‘central financial institutions’ (CFI) department of the ministry of education the list of all known and active primary and secondary schools was used. The CFI list contained 6850 active primary schools in The Netherlands with a unique BRIN<sup>11</sup>-number. As a precaution for a possible low response rate, three groups of 200 primary schools were randomly selected, each numbered 1 to 200, although only one of

---

<sup>11</sup> The acronym BRIN stands for ‘Basic Registration Institutional Number’. These are unique numbers the Dutch ministry of education uses to identify schools in The Netherlands.

these groups was used as the intended sample. For each school in this intended sample that did not react, the school with the same number from the second sample was asked to participate. If this school also failed to react, the school with the same number in the third sample was asked to participate in the research. This procedure was used to maintain the representativeness of the sample as much as possible.

### **Sample characteristics – sample 1 – Primary Education**

The sample of primary school teachers that responded by completing the questionnaire (N=77) consisted of 23 males and 54 females. Their mean age was 40.4 (sd=11.9) years and all but four were born in The Netherlands. The whole sample had received primary school teacher qualifications in The Netherlands (74-Pabo<sup>12</sup>, 3-KLOS<sup>13</sup>) and 12 of them had additional teaching qualifications. The mean number of years that the teachers in the sample had been teaching in primary education was 16.2 (sd=11.5).

### **Sample selection – sample 2 – Secondary Education**

The same approach was used with the CFI list of all 546 active secondary schools. Due to the limited number of active secondary schools, all secondary schools were randomly divided in three groups of 182. Because BRIN numbers have been used in the sample selection, the data attained on applied methodology in Dutch primary and secondary schools reflects the English language teaching on school level. Each BRIN number stands for a primary school with one or more final grades. For secondary schools the BRIN number indicates schools with possibly several buildings and streams. This implies that for secondary schools the BRIN numbers might not be uniquely related to a school stream.

### **Sample characteristics – sample 2 – Secondary Education**

The sample of secondary school teachers that responded by completing the questionnaire (N=112) consisted of 24 males and 88 females. Their mean age was 43.3 (sd=11.4) years and all but twelve were born in The Netherlands (3 missing). The whole sample had received secondary school teacher

---

<sup>12</sup> 'Pabo' is a Dutch acronym that stands for: 'Pedagogische academie voor het basisonderwijs', literally: Pedagogical academy for primary education; the primary school teacher training (bachelor) programme of the school of education of a university of applied sciences.

<sup>13</sup> 'KLOS' is a Dutch acronym that stands for: 'Kleuter Leidster Opleiding School' literally 'Toddler Leader Training School'; teacher training specializing in reception years (four to six-year-olds), which has been made part of the Pabo in 1985.

qualifications in The Netherlands, on top of which 11 had Pabo and 1 KLOS qualifications, while 14 had other additional teaching qualifications. The mean number years the teachers of the sample had been teaching in secondary education was 14.9 (sd=10.6).

### 3.2.3. Instrument

The questionnaire used is an adapted version of the questionnaire used by Bonnet (2004, p. 203), see appendix 1. The final version of the questionnaire as used in this study, consisted of three sections. The first section was on demographic and background variables, with 7 questions on personal particulars like gender, age, country of birth, the level of teacher qualifications and the years respondents had been employed in education. The second section concerned the quantity of English language teaching offered in primary education, with 16 questions about the details of the organisation and nature of English teaching in the respondent's primary school, such as age groups taught, use of course books, presence of Native speakers and the amount of teaching time per year, week and meeting. The third and final section dealt with 41 questions on the teaching methodology used and the repertoire of language teaching activities. The first 7 questions dealt with the time spent on the four skills, listening, speaking, reading and writing, as well as the organisation of the classroom, culture<sup>14</sup> and grammar, the methodology used and repertoire of pedagogical activities. All teachers answered section 1 and 3, section 2 has only been answered by the primary school teachers.

- (1) Respondents were asked to indicate on a 13-point scale (1 = 'no time', 2 = 'a few minutes', 3 = '5 minutes', 4 = '10 minutes', 5 = '15 minutes', etc. until 13 = '55 minutes or more') how much time on average during lessons would be spent on these 7 different parts of language teaching throughout the year.

After one question about target language use, in the following 20 questions the respondents were given an extensive list of possible activities through which the four skills and grammar might be offered. These 20

---

<sup>14</sup> 'culture' in this study refers to all language teaching activities dealing with the history, culture, current affairs and news from the English speaking world.

possible activities were based on the activities offered in the teaching methodology part of the Rotterdam University of applied sciences teacher training curriculum. Activities included consisted of general language acquisition activities like 'reading' as well as specific activities like TPR<sup>15</sup> and drama. Respondents were asked to indicate on a seven point Likert scale whether they used these activities (1=almost never; 4= half of the meetings; 7= almost every meeting). The final 13 questions focussed on cooperative learning activities, possible differentiation approaches and testing characteristics. The list of activities offered in the questionnaire is an enumeration of possible language teaching activities, without clustering. Possible clusters of activities have been analysed statistically, as described in 4.3.2.2.

#### **3.2.4. Instrument distribution**

As stated in the above chapter on sample selection, both the randomly selected 600 of the 6850 active primary schools and 546 active secondary schools were divided into three portions and numbered A1 to C200. The approach was to send out a first portion of 200 and 182 respectively of the A portion in May 2012 and after two weeks send out a corresponding number from the B, or even C portion if a school failed to respond.

#### **Analyses**

To verify the representativeness of the sample from primary education, some school characteristics of the school in the sample were compared to the same characteristics of all other primary schools in the Netherlands by means of analyses of variance and cross tabulation with a chi-square test. Three groups were compared: schools in the sample, non-responding schools and all other primary schools in the Netherlands. The dependent variables used for the analyses of variance concern the proportion of students of low, medium and high social economic status in school respectively. For each of the analyses of variance, the assumption of homogeneity of variances was tested by means of Levene's test. The cross tabulation

---

<sup>15</sup> TPR (Total Physical Response) activities enforce the meaning of language used by physically acting out what is being said, or the other way around.

with chi-square was used to verify differences in geographical spread of the schools. For the chi-square test it was checked whether no cells had an expected value below 1 and not more than 20% of cells had an expected frequency below 5. Both of these assumptions must hold to conduct a chi-square test (Field, 2009). The additional data used was attained from the ministry of education (CFI).

The representativeness of the sample of secondary schools could not be verified the same way, since comparable data for secondary education was not available.

To answer sub-question (2) ***“How does the teaching of English in the Netherlands differ between the final year of primary education and first year of secondary education?”***, first independent t-tests were carried out in order to find whether time spent on the four language skills, grammar and culture differ significantly between primary and secondary education. Levene’s test was used to check, for each t-test, whether the assumption of homogeneity of variances was violated. If so, the results of the t-tests were corrected for the lack of homogeneity.

Second, it was checked whether a significant difference could be found in the frequency of the teaching activities used in the classroom as distinguished in the questionnaire. On top of that, an attempt was made to ascertain clusters of English teaching activities that teachers use with the same frequency or in the same combinations. This was done in the hope of finding clusters of activities that would indicate a certain approach to English language teaching as well as for data reduction. Comparison of components is easier than comparison on item level. For that reason, a Principal Component Analysis (PCA) on the frequency scores of the activities was conducted. Also, a second PCA on the sums of items that were based on the first PCA on item scores was conducted to search for second order components.

Leading up to the PCAs, Kaiser–Meyer–Olkin (KMO) and measures of sampling adequacy per item (diagonal of the anti-image correlation matrix) were used to verify whether the sample was large enough to conduct a PCA. Also Bartlett’s test of sphericity was used to verify whether the covariances between items were sufficient to conduct a PCA. Given the expected correlation between components, Direct Oblimin rotation was chosen. After analysis of items determining a component (first PCA), these items were summed to form a new variable. Again t-tests were used on these sums to verify whether the mean sums differed

significantly between the samples of primary and secondary education. Before conducting t-tests, the reliability of the sums was checked by means of Cronbach's alpha. The minimum value of Cronbach's alpha for conducting research on group level is .60 (Field, 2009).

### **3.4. Results of the questionnaire survey on the differences in English language teaching in primary and secondary education in The Netherlands.**

The results of the questionnaire research will be presented below. First the results of the non-response analysis will be offered. These results will be used to determine to what extent the sample of primary schools can be seen as representative for the national population of teachers in the final grade of primary school. Subsequently both the primary and secondary school sample will be described by means of personal details acquired through the questionnaire. The results of the t-tests and the PCA analyses aimed at answering the research question 'How does the teaching of English in the Netherlands differ between primary and secondary education?' will follow after that.

#### **3.4.1. Sample representativeness**

Two sample surveys were carried out as described above; one for primary education and one for secondary education. The rate of 74 teachers responding, while a questionnaire was sent to the random sample of 600 schools for primary education, was quite low with a response rate of 12%. With 111 secondary school teachers responding, from the 546 schools who received a questionnaire, the response rate of 20% was also rather less than satisfactory.

The response rate was so low, no generalisation can be built upon the findings just like that, so in order to allow discussion of the results in relationship to general educational practice in The Netherlands, comparison between the sample and the rest of the country was required.

To ascertain the degree of representativeness of the primary school sample represented by the responding teachers, the sample was analysed with the help of data regarding the complete population. The only data available on all the primary schools in The Netherlands consisted of information on social-economic student status and the geographical spread of the primary schools across the country. Of course, the information on school's student population and its physical location does not provide information on the

teaching of English. The assumption is, however, that the comparison will allow the results to be discussed as insight into what English language teaching looks like in practice.

**Additional data (1): average social-economic student status of students per school ('weighted student funding')**

Additional data on primary and secondary schools has been acquired from the Ministry of Education<sup>16</sup>. The added data on all students in primary education in The Netherlands consisted of the 'weight' (Dutch: leerlinggewicht) of students, which indicates the educational level of the parents and the spread of primary schools per province<sup>17</sup>.

Since 1985, the Netherlands has had a system of weighted student funding for all primary schools (Ladd & Fiske, 2010). This 'weight' is a good indicator for the socioeconomic status (SES) of the parents as it reflects their educational level.

Primary school students are divided into three categories on enrolment in the educational system. These categories are indicated with a 1.2, 0.3 or 0 'weight', which relate to the additional funding the school will receive for children in the first two of these three 'weights'.

Children whose parents have been in a special educational needs programme, or received primary education only, have a 1.2 weight. Children whose parents only received lower vocational training have a 0.3 weight and children whose parents received at least two years of general secondary education have a 0 weight. Schools receive additional funding for the 1.2 and 0.3 children to help them organise expected educational needs for these pupils.

---

<sup>16</sup> 23-06-2012: <http://www.cfi.nl/public/websitecfi/Default.aspx>

<sup>17</sup> <http://www.rijksoverheid.nl/onderwerpen/leerachterstand/vraag-en-antwoord/wat-is-de-gewichtenregeling-in-het-basisonderwijs.html>

### **Additional data (2): Geographical spread of primary schools**

The second dataset acquired from the Ministry of Education concerns the geographical spread of primary schools per province. Every BRIN-number is linked to a postal code and organised per province. The spread of all primary schools over the twelve provinces of The Netherlands is given, which enables a comparison with the spread of the primary schools in the response group. The geographical spread is important given the provincial and regional (rural and urban) differences in SES of the general population.

Like the data on student weight and geographical spread of schools, the randomization of the school sample and distribution of the questionnaire has been organised based on BRIN numbers. This means that in the school sample each respondent represents a BRIN-number. To allow comparison with unique BRIN-numbers of all primary schools only one questionnaire per school was used, which left 97.7% of the original data set for analysis.

### **Results of the representativeness analyses**

The first analysis was based on the school proportion of registered 'weight' of primary school students. To verify the representativeness of the school sample, three analyses of variance were conducted concerning school proportions of students of certain weights. As between factor, a variable is used that indicates whether a school belongs to the non-response (530 schools), the response (73 schools) or neither (6205 schools). As dependent variables, the proportions of pupils with each of the three weights were used.

Therefore, the analysis of variance was conducted three times: first with the proportions of 0-weight pupils in school as dependent variable, second with the proportions of 0.3 pupils in school as dependent variable and third with the proportions of 1.2 pupils in school as dependent variable. For each of the three analyses, first the assumption of homogeneous variances in each of the three groups was checked by means of Levene's tests.

For each of the three analyses, first the assumption of homogeneous variances in each of the three groups was checked by means of Levene's tests.

For the first analysis (% 0-weight students) the assumption of homogeneity of variances holds ( $F_{(2; 6085)}=1.323$ ;  $p=.266$ ). The analysis of variance shows the percentage of children with a 0-weight in the

response group, the non-response group and the schools that had not been approached did not differ significantly ( $F_{(2; 6805)}=1.129$ ;  $p=.323$ ).

For the second analysis (% 0.3-weight students) the assumption of homogeneity of variances holds ( $F_{(2; 6085)}=.072$ ;  $p=.931$ ). The analysis of variance shows the percentage of children with a 0.3-weight in the response group, the non-response group and the schools that had not been approached did not differ significantly ( $F_{(2; 6805)}=.522$ ;  $p=.593$ ).

For the third analysis (% 1.2-weight students) the assumption of homogeneity of variances holds ( $F_{(2; 6085)}=2.935$ ;  $p=.053$ ). The analysis of variance shows the percentage of children with a 1.2-weight in the response group, the non-response group and the schools that had not been approached did not differ significantly ( $F_{(2; 6805)}=1.265$ ;  $p=.282$ ). In order to further ascertain representativeness Pearson's chi-squared test has been used. The Pearson chi-square test showed no significant deviation (.240) in the results of the cross tabulation of the spread of primary schools across the 12 provinces of The Netherlands compared with the response and non-response group.

### **Non-response-analyses Secondary education**

As stated earlier in 'Sample selection – sample 2 – Secondary Education' analysis of sample representativeness of data gathered from secondary schools was impeded by the fact that secondary schools tend to have a variety of branches under one BRIN number and data such as student 'weight' is unavailable.

The above paragraphs show that, despite the low response rate, the data won in primary education appear to be representative for the population of Dutch primary schools. The schools in the sample do not differ significantly from the rest of the Dutch population of schools when compared on distribution of weighted students or geographical spread. As far as secondary education is concerned such a distinct claim is unwarranted.

### 3.4.2. Analysis Results Questionnaire

The question: *How does the teaching of English differ between primary and secondary education?* is answered in two parts as described in the introduction to this chapter above. The first part concerns the amount of time spent on the four language skills (listening, speaking, reading and writing), grammar and culture. The second part concerns the variety in language learning activities and assessments used in the classroom.

#### 3.4.2.1 Amount of time spent on the four language skills (listening, speaking, reading and writing), grammar and culture.

The first analysis deals with the time spent on listening, reading, speaking, writing, grammatical tasks and culture activities. As described above in the introduction and chapter 2, primary education is expected to spend more time on speaking and less on grammar than secondary education, while secondary education is expected to spend more time on writing, reading and grammar. In table 3.1 the mean scores, standard deviations and standard errors can be found for both samples of primary and secondary education in The Netherlands. As can be seen in table 3.1, the SDs indicate that in both levels of education rather substantial differences in time are spent on language skills, grammar and culture. For an indication of the range (95% of the scores when normally distributed) the SD needs to be added twice and subtracted twice. The first mean of 4.29 (Listening activities in primary education) indicates teachers spend approximately 15 minutes each lesson on listening activities. The range of scores however was for 95% the 4.29 score plus and minus 2 times 1.31 (SD 1.315), so some of the teachers score around 1.67 (0 to a few minutes), while others score around 6.91 (25 minutes). Secondary school teachers in the sample have a mean score of 3.73, but the SD (times two 1.788) shows some of the teachers score around 1.942 (a few minutes), while others score around 5.518 (between 15 and 20 minutes). The high SDs in this table indicate that merely comparing the means with the scales do not represent consistent classroom behaviour throughout the sample. In this example the range of time spent on listening activities in the sample can run up to 25 minutes in primary education and approximately 20 minutes in secondary education.

**Table 3.1: Average time spent per lesson on the four skills, Grammar and Culture in Primary vs Secondary education (N = number of respondents; Mean = mean score, in which 1 = ‘no time’, 2 = ‘a few minutes’, 3 = ‘5 minutes’, 4 = ‘10 minutes’, 5 = ‘15 minutes’, etc. until 13 = ‘55 minutes or more’; SD=standard deviation= ; SE= standard error)**

	education	N	Mean	SD	SE
Listening activities	Primary	76	4.29	1.315	.151
	Secondary	111	3.73	.894	.085
Reading activities	Primary	77	3.83	1.229	.140
	Secondary	111	4.57	1.332	.126
Speaking activities	Primary	77	4.62	1.377	.157
	Secondary	111	3.58	1.379	.131
Writing activities	Primary	77	3.94	1.291	.147
	Secondary	111	4.38	1.695	.161
Grammar activities	Primary	77	2.90	1.071	.122
	Secondary	111	4.51	1.205	.114
Culture activities	Primary	77	1.95	1.307	.149
	Secondary	110	2.90	1.141	.109

To ascertain which means differ significantly between primary and secondary education, independent t-tests have been used. For each t-test first it has been checked whether the assumption of homogeneity of variances was violated. When this was the case the output has been corrected for unequal variances.

For ‘Writing activities’ the assumption of homogeneity of variances was violated ( $F_{(1, 185)}=4.262$ ;  $p=.040$ ).

The assumption of homogeneity of variances was not violated for the other five variables (Listening activities  $F_{(1, 186)}=2.506$ ;  $p=.115$ ; Reading activities  $F_{(1, 186)}=1.097$ ;  $p=.296$ ; Speaking activities  $F_{(1, 186)}=.354$ ;  $p=0.552$ ; grammar  $F_{(1, 186)}=.158$ ;  $p=.691$ ; culture  $F_{(1, 185)}=.038$ ;  $p=.845$ ).

For each of the tested variables the difference between primary and secondary education turned out to be significant (Listening activities  $t=3.467$ ;  $df=185$ ;  $p=.001$ ;  $r= -.247$ ; Reading activities  $t=-3.846$ ;  $df=186$ ;  $p=.000$ ;  $r=.271$ ; Speaking activities  $t=5.122$ ;  $df=186$ ;  $p=.000$ ;  $r= -.352$ ; Writing activities  $t=-2.033$ ,  $df=184,331$ ;  $p=.043$ ;  $r=.141$ ; grammar  $t=-9.466$ ;  $df=186$ ;  $p=.000$ ;  $r=.570$ ; culture  $t=-5.287$ ;  $df=185$ ;  $p=.000$ ;  $r=.362^{18}$ ).

The results reflect the situation as described in the introduction (p. 2) and also in the conclusion of chapter 2 (p. 32). In primary education more time is spent on listening and speaking to teach English and less time on reading, writing, grammar and culture, than in the secondary education sample. The Independent t-tests show how the averages for primary education differ significantly in the expected directions. The effect sizes

<sup>18</sup> In this paragraph ‘r’ is given as effect size: < .3 is small, .3-.5 middle; > .5 is big (Field, 2009)

for listening, reading and writing are small and for speaking, grammar and culture they are medium. These differences in the means are as expected and consistent with the described difference in English language teaching methodology between primary and secondary education. It seems clear, despite the response rate and unavailability of data for the comparison with the nation-wide population of secondary schools, teachers in the primary education sample indicate they spend considerably more time on speaking and less on grammar than the sample of teachers from secondary education.

#### **3.4.2.2. Variety in language learning activities and assessments used in the classroom**

Apart from the questions about time spent on the four language skills, grammar, history, culture and news of the English speaking world, 33 questions have been offered to obtain information about specific language teaching methodology used in the classroom. Respondents were asked to indicate on a Likert scale how often they used certain activities in their teaching. The Likert scale ran from 1 to 7 in which 1 stood for: 'I almost never use this activity', 4 stood for: 'I use this activity in half of all lessons' and 7 stood for: 'I use this activity in almost every lesson'.

Differences in language teaching methodology are checked per item, but also for clusters of items that appear to form one latent trait. Therefore, the clustering of items was checked by means of a Principal Components Analyses (PCA).

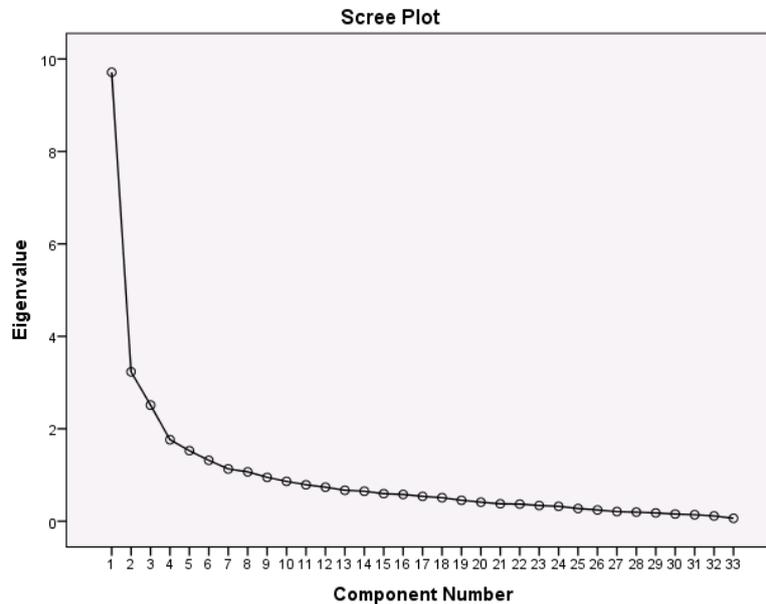
To verify which latent traits are measured by the items used, exploratory factor analyses or Principal Components Analyses (PCA) was conducted. The PCA shows the clustering of items and facilitates the interpretation of the content of the clusters found.

The sample size was sufficient for a PCA according to the KMO (.828) and all items show measures of sampling adequacy over .5 (range .687-.930). The items showed enough covariation for an exploratory factor analysis (EFA) (Bartlett's test of sphericity:  $\chi^2=2969.553$ ;  $df=528$ ;  $p=.000$ ). Direct Oblimin was chosen as rotation because of the expected correlation between components.

The analysis resulted 8 components with an eigenvalue over 1. The scree-plot, as shown in graph 3.2, does not help to ascertain the number of distinguishable components as it does not show a clear bend. A clear bend would indicate which factors explain most of the variability, as these factors would ideally be in the

steep curve of the pattern before the bend. Factors in the flat, or horizontal line after the bend, explain a very small proportion of the variability and are likely to be unimportant.

**Graph 3.2: Scree Plot of the PCA on the variety in language learning activities and assessments used in the classroom.**



On the other hand, the 8 components found explain the variances of items reasonably well, as the communalities of the items vary between .446-.864. and the 8 components together explain 67.47% of total item variance.

The analysis of the pattern and structure matrix (see appendix 2) showed the components allowed content-related interpretation. Items with a large component loading, but a content-wise ill-fit have not been combined in total scores with the other items of the component (e.g. 'Course book based activities').

### **Components found**

In general, the components found are logical combinations of items or represented language teaching activities. When two items have high loadings on the same component, this means that teachers who say they use an activity described in one of these items, also tend to use the activity described in the other item more often. After inspection of item loadings on components, two components appear to consist of activities that mainly deal with one of the language skills writing (component 3) and listening (component 5). The other components were named after some other category of activities they represent, in all

resulting in the following four components: 'Coop-learn' (component 1), 'Song and game' (component 2), 'Grammar' (component 6), and 'Adapt' (component 4). Because some important aspects were missing in the aforementioned components the categories speaking, reading and tests have been added.

The components found are described with their given names below, including the items that make up the component. Also, items that have been summed as indications of speaking, reading and tests, are presented. The items in the component are only listed when they logically belong together based on their shared aim or activity. Surprising, or seemingly deviant items are discussed on the basis of their coherence with the other items that make up the component.

The first component with high loadings of the items 'Cooperative learning activities', 'Group work', 'Projects' and 'Thematic activities' is named 'Cooplearn' as it contains activities in which students are expected to cooperate. The second component has high loadings of the items 'Singing', 'TPR' and 'Drama'. These items are summed into a score named 'Song and game'. The 'Writing' component (3) is mainly determined by items that deal with writing, such as: 'copying', 'processing', 'creative writing' and 'filling in forms', together with 'Creating posters or brochures'. 'Creating posters or brochures' has been made part of this component, although this activity could also be seen as belonging to 'Coop-learn'; the product aimed for is still largely based on the students processing the language acquired in previous teaching moments through writing the brochure or editing words and texts for use in a poster. The fourth component 'Adapt', has high loadings of the items 'Activities geared to the student perception of their environment', 'Activities geared to student language skills levels', 'Activities adapted to student wishes or suggestions' and 'Activities adapted to the students' English language teaching experiences of previous years'. The fifth component, 'Listening' is not only determined by items concerned with the listening skill: 'Listening to stories in English', 'Listening for information in English' and 'Listening to authentic English listening materials'. The items 'Using spoken English for all interactions during the lesson' as well as 'watching Videos / YouTube / TV (recordings)' also show high loadings on this component. In both cases the student activity required is 'listening'. When a teacher uses spoken English for all interactions during the lesson, for interactions to be effective, students need to listen and although watching TV or videos offers additional information and support, 'listening' is still the main receptive skill trained. 'Grammar',

component 6, has high loadings of the items 'Grammar – explaining/discovering rules' and 'Grammar – practice'.

Sums that are not based on the PCA but on conceptual reasoning are 'Speaking', 'Reading' and 'Tests'.

'Speaking' contains the items 'Speaking - dialogues/conversations' and 'Speaking - presenting/performing monologues – show and tell'. 'Reading', consists of the items: 'Reading stories in English' and 'Reading informative texts in English'.

'Tests' is constructed by summing the items 'Diagnostic tests', 'Formative tests' and 'Summative tests'.

The pattern matrix of the PCA is given below in table 3.3, the structure matrix can be found in appendix 2.

**Table 3.3: Pattern matrix of principal components analysis. Highest loading of an item contained in a component found in bold, highest loading of an item not contained in a component found in bold and italics. Items contained in a component are indicated by light grey shading.**

items	Component							
	1	2	3	4	5	6	7	8
English is the instruction lang.	-.146	.154	.073	.211	<b>-.471</b>	.189	-.335	-.376
Listening to English stories	.073	.081	.028	-.046	<b>-.790</b>	-.057	.069	-.093
Listening for info. in English	.073	-.068	.094	-.108	<b>-.802</b>	-.008	.137	-.233
Listening to auth. English materials	-.068	-.023	-.043	-.046	<b>-.874</b>	-.099	-.074	.177
Watching Video / YouTube / TV	.155	.093	-.159	.225	<b>-.340</b>	.236	.078	.316
Reading English stories	.149	-.292	.128	.160	<b>-.421</b>	.199	.258	.028
Reading informative texts in English	.020	<b>-.528</b>	.164	.222	-.347	.242	.260	.130
Speaking - dialogues/conversations	.215	.286	<b>.368</b>	.052	-.253	-.230	-.061	-.304
Speaking - monologues – show and tell	<b>.407</b>	.075	.182	.252	-.048	.018	-.233	.141
Singing	.048	<b>.714</b>	.202	.035	-.113	-.262	.097	-.013
Writing - copying	-.124	.288	<b>.753</b>	-.165	.002	.264	.008	-.053
Writing – practice	-.033	.138	<b>.755</b>	.006	-.044	.065	-.082	-.199
Writing – creative writing	.232	-.416	<b>.455</b>	.154	-.134	-.013	-.152	.173
Writing – filling in forms	.129	-.345	<b>.700</b>	.046	-.041	-.127	.153	.107
Making a Poster or brochure	<b>.349</b>	-.072	.263	.156	-.033	.027	.018	.299
Grammar – explaining rules/constructs	.044	-.091	.033	.050	.115	<b>.904</b>	.064	-.085
Grammar – practice	.034	-.138	.087	.015	.061	<b>.890</b>	.110	-.079
Total Physical Response activities	.094	<b>.640</b>	.134	.066	-.069	.056	.360	.167
Drama	.095	<b>.514</b>	.051	.427	-.081	-.166	.192	.201
Playing (language) games	.113	.250	.118	<b>.424</b>	-.078	.031	.201	.129
Drills and choral work	.004	.269	-.014	.146	-.139	.151	<b>.670</b>	-.030
Cooperative learning activities	<b>.756</b>	.105	-.031	-.028	-.016	-.027	.281	-.387
Group work	<b>.837</b>	-.020	.033	.017	.048	-.123	.106	-.159
Project based activities	<b>.740</b>	-.058	-.045	.010	-.077	.139	-.098	.234
Theme based activities	<b>.562</b>	-.013	.005	.164	-.126	.087	.007	.148
Activities based on course books	.085	-.111	.124	.205	-.067	.137	.011	<b>-.699</b>
Activities adapted to l. perception	.024	-.013	-.109	<b>.907</b>	.047	.027	.036	-.074
Activities adapted to l. language skills	-.089	-.100	-.100	<b>.973</b>	.001	-.047	-.030	-.177
Activities adapted to l. wishes/sugg.	.083	.072	.018	<b>.727</b>	.034	.022	.076	.049
Activities adapted to l. experiences	.054	.006	.123	<b>.714</b>	.092	.104	-.070	.022
Diagnostic tests	<b>.599</b>	.049	.008	.085	.005	.259	-.174	.028
Formative tests	<b>.430</b>	.066	-.012	.082	-.196	.367	-.255	.068
Summative tests	.097	.210	.057	.094	-.196	<b>.384</b>	-.231	.132

Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization. a. Rotation converged in 29 iterations.

As can be seen in table 3.5, in some cases an item with a high loading on a component has been left out of the sum score for that component based on the meaning of the item. Component 1 (Coop-learn), for instance, contains the items: Cooperative learning activities, Group work, Project based activities and

Theme based activities. Speaking (monologues), Based on reasons of content, Making a Poster or brochure, Diagnostic tests and Formative tests have been left out of component 1. Making a poster had a high enough loading to be added to component 1, but fitted component 3 better content-wise.

Of the components found, the first six were clear enough to sum based on their loadings. Two loose items (Drills and choral work and Activities based on course books) showed high loadings in the seventh and eighth component respectively. Instead of using these loose items, sums have been made for the components Reading and Tests. So, the ultimate sums have been based on the content-wise merits of item combination and subsequently calculated Alphas.

### **Differences in language teaching methodology used between the primary and secondary education sample**

For six of the eight components found, the items characterizing the component, as described above, were summed resulting in six variables. Each sum is divided by the number of items, so that sums can be interpreted on the original Likert scale. For each of these sums, Cronbach's alpha was calculated as indication of reliability (see table 3.4). Apart from Speaking, the reliability of each of the sums presented is larger than .6, the lower limit for research on group level (Field, 2009). The reliability of the Speaking component was too low to separately construct a sum.

**Table 3.4: Reliabilities of sums language teaching methodology items (based on PCA)**

<b>Variable (PCA component between brackets)</b>	<b>N</b>	<b>items</b>	<b>Cronbach's alpha</b>	<b>Range RIT's</b>
Coop-learn (1)	187	Cooperative learning act., Group-work, Projects, Thematic act	.792	.708-.765
Song + game (2)	189	Singing, TPR, Drama	.784	.634-.755
Writing (3)	188	(Writing) - Copying, - Processing, Creative writing and Creating posters/brochures	.695	.600-.685
Adapt (4)	179	(Activities geared to) - Student perception of their environment and - Language skills levels., (Activities adapted to) - Students wishes or suggestions and – student English language teaching experiences of previous years	.861	.794-.861
Listening (5)	190	Using spoken English for all interactions during the lesson, (Listening to) - Stories in English, - information in English, - Authentic English listening materials and Watching Videos / YouTube / TV(recordings)	.754	.655-.774
Grammar (6) speaking	187	(Grammar) – Explaining and/discovering rules and Practise	.951	.907
	191	Speaking - dialogues/conversations, Speaking - presenting/performing monologues – show and tell	.507	.347
Reading	190	Reading - stories in English and - informative texts in English	.748	.598
Tests	188	Diagnostic-, Formative- and Summative tests	.705	479-.806

Subsequently for each of the eight sums with sufficient reliability (so not for speaking) and for each individual item not contained in a sum, t-tests were conducted to verify whether mean scores differ for primary and secondary education. The items not contained in a sum are the speaking items 'Speaking – dialogues', 'Speaking – presenting' and the other items 'Playing language games', 'Drills and choral work' and 'Course book based activities'. T-tests were also conducted for these individual items. Preceding the independent t-tests on the group means for primary and secondary education on the 8 sums and remaining individual items, Levene's test has been used to check whether the assumption of homogeneity of variances holds. If the assumption is violated, results are corrected for lack of homogeneity.

The assumption appears to have been violated for 'Song+game' ( $F_{(1, 185)}=17.729$ ;  $p=.000$ ), 'Adapt' ( $F_{(1, 175)}=9.337$ ;  $p=.003$ ) and 'Reading' ( $F_{(1, 186)}=10.489$ ;  $p=.001$ ) and appears to not have been violated for 'Coop-learn' ( $F_{(1, 183)}=.003$ ;  $p=.957$ ), 'Writing' ( $F_{(1, 184)}=.068$ ;  $p=.795$ ), 'Listening' ( $F_{(1, 186)}=.807$ ;  $p=.370$ ), 'Grammar' ( $F_{(1, 183)}=2.283$ ;  $p=.133$ ) and 'Tests' ( $F_{(1, 184)}=.456$ ;  $p=.500$ ). In the t-test, corrections were applied where the assumption of homogeneity of variances was violated. For the individual items, the homogeneity of variances assumption was violated for 'Speaking' (monologue) ( $F_{(1, 187)}=14.336$ ;  $p=.000$ ) and 'Drills and choral work' ( $F_{(1, 183)}=20.936$ ;  $p=.000$ ) and appears to not have been violated for 'Speaking' (dialogue) ( $F_{(1, 187)}=.005$ ;  $p=.942$ ), 'Playing language games' ( $F_{(1, 186)}=3.154$ ;  $p=.077$ ) and 'Course book based activities' ( $F_{(1, 184)}=.627$ ;  $p=.430$ ).

The t-tests showed there are significant differences between the means for teachers in primary and secondary education for the variables 'Song+game' ( $t=10.127$ ;  $df=115.667$ ;  $p=.000$ ;  $r=-.628$ ), 'Reading' ( $t=4.537$ ;  $df=134.894$ ;  $p=.000$ ;  $r=.329$ ), 'Grammar' ( $t=-6.948$ ;  $df=183$ ;  $p=.000$ ;  $r=.457$ ) and 'Tests' ( $t=-2.869$ ;  $df=184$ ;  $p=.005$ ;  $r=.207$ ). The means, standard deviations and standard errors are presented in table 3.5, below. The differences in means for teachers in primary and secondary education are not significant for Cooperative learning activities ( $t=-.786$ ;  $df=183$ ;  $p=.433$ ), 'Writing' ( $t=-.021$ ;  $df=184$ ;  $p=.983$ ), 'Adapt' ( $t=-1.533$ ;  $df=127.605$ ;  $p=.128$ ) and 'Listening' ( $t=1.067$ ;  $df=186$ ;  $p=.287$ ).

For the individual items, significant differences between teachers in primary and secondary education were found for 'Speaking' (dialogue) ( $t=6.259$ ;  $df=187$ ;  $p=.000$ ;  $r=-.416$ ) and 'Drills and choral work' ( $t=2.937$ ;

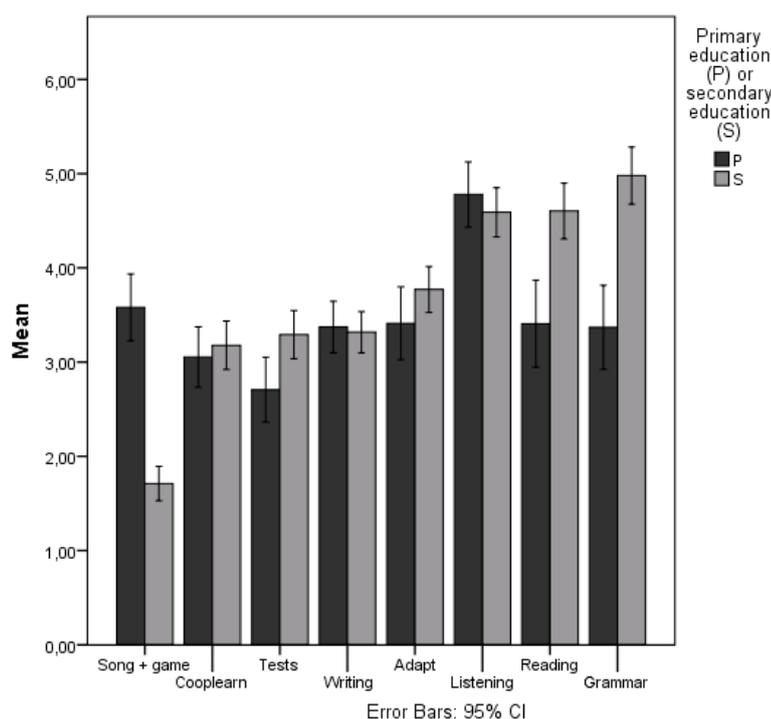
df=118.114; p=.004; r=-.227), and not significant for 'Speaking' (monologue) (t=-.494; df=136.192; p=.638), 'Playing language games' (t=1.153; df=186; p=.250) and 'Course book based activities' (t=1.127; df=184; p=.261).

**Table 3.5: Means of sums and items concerning language teaching methodology activities (sums based on PCA). Means differing significantly between primary and secondary education in bold (1=(almost) never; 4= half of the meetings; 7= (almost) every meeting).**

Variable	education	N	Means	SD	SE Mean
Song+game	Primary	76	<b>3.65</b>	1.437	.165
	Secondary	111	<b>1.77</b>	.907	.086
Coop learn	Primary	76	3.06	1.235	.142
	Secondary	109	3.20	1.248	.120
Tests	Primary	75	<b>2.76</b>	1.395	.161
	Secondary	111	<b>3.33</b>	1.241	.118
Writing	Primary	76	3.34	1.063	.122
	Secondary	110	3.35	1.062	.101
Adapt	Primary	73	3.48	1.577	.185
	Secondary	104	3.81	1.199	.118
Listening	Primary	76	4.76	1.372	.157
	Secondary	112	4.55	1.287	.122
Reading	Primary	77	<b>3.46</b>	1.856	.212
	Secondary	111	<b>4.60</b>	1.420	.135
Grammar	Primary	75	<b>3.41</b>	1.731	.200
	Secondary	110	<b>5.07</b>	1.507	.144
Speaking. dialog	Primary	77	<b>5.43</b>	1.576	.180
	Secondary	112	<b>3.95</b>	1.615	.153
Speaking. presenting monologues	Primary	77	2.61	1.623	.185
	Secondary	112	2.71	1.262	.119
Language game	Primary	77	2.92	1.604	.183
	Secondary	111	2.68	1.315	.125
Drills and choral work	Primary	74	<b>3.14</b>	2.050	.238
	Secondary	111	<b>2.33</b>	1.403	.133
Course book based activities	Primary	76	5.42	1.798	.206
	Secondary	110	5.14	1.617	.154

Graph 3.6 below represents the mean scores of the items measuring frequency of activities used for language teaching as discussed above. These are sorted from lowest frequency to highest frequency used in secondary education (in grey bars). The adjacent black bars represent the mean scores for primary education.

**Graph 3.6: Frequency of activities used for language teaching.**



The significant differences found exactly reflect the difference in approach of English teaching between primary and secondary education as described in the introduction and also found in the literature review. Singing songs and playing language games are almost exclusively used in primary education. Conversely, teaching grammar, reading and testing play a dominant role in secondary education.

The eight sums have again been analysed with a PCA (again with Direct Oblimin rotation) to check for 2<sup>nd</sup> order factors. Again the sample size was big enough (KMO=.811) as for each true-score (variable), as the MSAs range between .585-.857. The correlation between the 8 variables was sufficient (Bartlett's test of sphericity, chi-squared=433.207; df=28; p=.000) for a PCA.

The PCA showed 2 components with an eigenvalue larger than 1, that together explain 61.094% of the variance in the 8 variables. Item variance is reasonably explained, as communalities are between .486-.801. Only 'Song+game' loads high on the 2<sup>nd</sup> component, 'Grammar' loads high on component 1, like all 6 other variables load on the first component. This clearly shows that 'Song+game', as a component of clustered activities, is completely different from the other clustered activities. Teachers who use the activities

clustered in 'Song+game' in their classroom do less with grammar, those who focus on other activities also pay more attention to grammar.

The above findings not only reflect the difference in approach to English teaching between primary and secondary education as described in the introduction and in the conclusion of the chapter 2 literature review, it also reconfirms the importance of grammar to the repertoire of English language teaching activities in secondary education.

### **3.5. Conclusion of the questionnaire survey on the differences in English language teaching in primary and secondary education in The Netherlands.**

The explicit importance given to grammar by the Dutch teachers (of English) in the sample, shows itself in the time they spend on the topic as well as in their choice of activities. These findings hold despite the low response level and the range of time spent in practice as was found in the standard deviation in relation to the means. The situation is illustrated in graph 3.6 above, in which the SD is given in the vertical lines at the top of each bar. The picture this paints of the activities used for English language teaching in Dutch education does not reflect the developments language teaching has gone through in the past century. Although speaking has won a place in the methodological repertoire like TPR and other activities, grammar still takes a considerable amount of time.

Generally speaking, it seems clear from the results that in primary education considerably more time is spent on speaking than in secondary education, where conversely considerably more time is spent on grammar.

Grammar as language teaching activity is pivotal in the sense that, as described in chapter 2, language teachers have tried to move away from grammar teaching for more than a century, but it reappears in each new method and approach presented and shows itself in this questionnaire research as a consistent part of English language teaching in secondary education.

Singing songs and playing language games are almost exclusively used in primary education. Conversely, teaching grammar, reading and testing play a dominant role in secondary education.

The found difference in approach to English teaching also shows, on a more abstract level, that the communicative approach is more clearly applied in primary education. Secondary education is moving in the direction of this approach, helped by CEFR labelled course books and final exams, but it currently does not show in the reported classroom activities of this research; there is still a difference in focus.

This difference in focus between primary and secondary education might be explained by the way teachers have been trained in their respective institutes. It might be explained by the history of the developments language teaching has gone through and the different moments in time English language teaching was introduced in the respective levels of education in The Netherlands. It might even be the vicious circle described by Hermans-Nymark (2006) because of which new teachers conform their methodological repertoire to the activities used by senior staff in their departments.

Apart from these explanations, the position of grammar teaching remains a strong one. A position hardly challenged by teaching reform but one that itself might even be a challenge to reform. The time claimed by grammar teaching, explanations and practice, as reported by the sample of secondary school teachers in this study, is around 15 minutes per meeting of 50 minutes. This means 30% of all English language teaching time is spent on grammar. This investment in grammar weighs heavily on, if not limits, the repertoire of possible language teaching activities used in every day teaching.

The effect of time investments in grammar and the reasoning, or beliefs, behind the reported situation are beyond the scope of this study. The reason this finding has been elaborated upon is because teachers are expected to invest time in a language teaching programme based on differentiated instruction, as will be reported in chapter 6.

**4. Literature study on differentiated instruction: Historical context(A), Theoretical foundations(B), building blocks (C), the role of teachers (D) and the effectiveness of differentiated instruction (E).**

**4.1. Introduction**

In this literature study a literature review is presented to answer sub question **3) *Which aspects should a language teaching programme based on the principles of Differentiated Instruction for the first year of secondary school include, to ensure knowledge and skills acquired in primary school are used to their full potential as well as enable execution of activities on different levels?***

The scope of this literature review is rather wide, as a lot has been published about differentiated instruction, the reasons for using this approach in education and the components of the educational process that can be differentiated. An extensive review is essential because differentiated instruction is rather complex, as will be shown in 4.3.2.1. and there are quite a number of myths about differentiated instruction that would hamper its implementation (Struyven, Coubergs, Gheysens, Engels, & Smets, 2016). To structure the literature found, the results are divided in five parts. The first part (A - the historical context of differentiation) describes how differentiation won a place in education and only moved from streaming and ability grouping to whole class practice in the past fifty years. The second part (B - the theoretical foundation of differentiation) reviews the research on which differentiated education is based and publications on differentiated instruction in general. This section is called 'theoretical foundation of differentiated instruction' as it aims to explain what differentiated instruction is by how it is defined and give a rationale for why this approach is needed. The reasons for the implementation are dealt with subsequently. In the third part (C - the building blocks of differentiated instructions) the literature dealing with the aspects that together make up differentiated instruction is reviewed. The label 'building blocks' is used in concord with studies that use the same terminology (Rock, Gregg, Ellis, & Gable, 2008; Subban, 2006; Tomlinson & Allan, 2000), as it contains all aspects teachers need to take into consideration to differentiate instruction for all students. Tomlinson (2001) has embedded differentiated instruction in the complexity of educational practice. This results in a wide range of aspects to be taken into account when a teacher wants to apply differentiated instruction; aspects that will only build up towards real differentiated

instruction when stacked together, hence the term 'building blocks'. The third and fourth part (D – teachers, and E – the effectiveness of differentiated instruction) review literature on the role of the teacher and the effectiveness of differentiated instruction, as found in previous studies, respectively. Each of the sub-chapters below will be concluded with a review of the consequences for the constructed language teaching programme based on differentiated instruction, referred to in sub-question 3, as mentioned above. This language teaching programme based on differentiated instruction is used as treatment in the experiment needed to answer the second part of the main research question (*How and to what extent does English teaching in the final year of primary school and first year of secondary education differ from each other and what is the effect of differentiated English language teaching on the attitude towards learning English and the English language skills development of students in the first year of secondary education?*)

At the end of each sub-section that offers insights, prerequisites, or information, the implications, which can be used in the development of a language teaching programme based on differentiated instruction, are described in italics.

The outcomes, as described in the conclusion of this chapter, consist of the prerequisites for a language teaching programme based on the principles of Differentiated Instruction for English language teaching in the first year of secondary education in the Netherlands. In the following chapter (5) the development and form of the language teaching programme based on differentiated instruction will be described, based on the prerequisites formulated in the conclusion of chapter 4 and the use of this programme as treatment in the experiment is described in chapter 6.

#### **4.2. Method used to perform the literature study on differentiated instruction.**

A literature search was conducted in the journals, dissertations, articles and books in the 398 databases of CataloguePlus of Amsterdam University. The use of this approach helped limit the hits to publications that are of methodological quality. The keywords used in the search concerned publications on differentiation of education ('differentiation' or 'differentiating' or 'differentiated' and 'education' or 'instruction' or 'curriculum' or 'teaching' and 'English' or 'language'), which yielded 5841 hits. Synonyms used .like

'adaptive' or 'differentiation' or 'differentiating' or 'differentiated' and 'education' or 'instruction' or 'curriculum' or 'teaching' and 'English' or 'language'), yielded 2931 hits. The greater part of these hits concerned differentiation in general sense. The keywords 'adaptive' and 'education' together yielded 6844 hits, although after addition of 'language' or 'teaching', 88 hits remained and after addition of 'English' 78 hits remained.

In addition to the digital library search, the so-called snowball method was used, in which key documents lead to citations or references of other sources on the same subject. The list on the theoretical foundations for differentiated instruction and the aspects that make up the 'building blocks' of differentiation found in the initial search and the follow-up search for publications on these topics are an example of this approach.

To determine which articles were included for further analysis, the following criteria were used:

- The study needed to report on differentiation teaching in general. Studies on differentiation, or differentiated instruction that did not have an international character, i.e. outside of the Dutch perspective, were excluded.
- The study needed to report on differentiation of (language) teaching in primary and secondary education, as this is the scope of this study.
- Both quasi-experimental and case studies and both quantitative and qualitative studies were included, as long as the method of differentiation was described in a sufficiently elaborate and transparent manner.

Applying these criteria to the initial set of articles resulted in a selection of 54 titles. As this study is not a systematic review of literature at a specific moment in time, with an added extensive snowball method effect, the literature used has been extended in the five-year duration of this study.

### **4.3. Results of the literature study on differentiated instruction.**

The results will be presented in five parts, as described in the introduction of chapter 4 above:

- A** – reviews research that provides the historical context of differentiation;
- B** – reviews research that forms the theoretical foundation of differentiation;
- C** – reviews research dealing with aspects that make up the building blocks of differentiated instructions;
- D** – reviews research on the role of teachers in differentiating instruction;
- E** – reviews research reporting on the effectiveness of differentiated instruction.

### **4.3.1. Results (A) - The historical context of differentiation**

#### **4.3.1.1. Differentiation in nineteenth century education**

In the nineteenth century, mainstream education in the western world moved away from the one room schools with students of all ages in the same room and one teacher, to an organisation where students were placed in grades according to their age. Home-schooling of individuals, or small groups, by subject teachers was limited to rich families who could afford to pay specialists. Although this form of education was successful and probably quite differentiated, it was beyond the reach of the masses because of the costs involved. A teacher per year group was the best governments could do at the time in The Netherlands as well as in the rest of the western world. Education around the western world was organised based on the principles of mass industrial production. Grading students according to age and having them go from grade to grade was obviously the most efficient and practical approach (Winter, 1985). This classification of pupils into grades had already become accepted practice in the second half of the 19th century and has survived until today as the predominant form of school organisation (Stoll Lillard, 2005).

Within this development of the school organised in year groups there is one systematic approach of interest to this research.. This system, 'the Monitorial System', employed during the first decades of the 19th century, was based on abler pupils being used as 'helpers' to the teacher. Helpers passed on information they had learned to other students. It was also known as "mutual instruction" or the "Bell-Lancaster method" after the British educators Dr Andrew Bell and Joseph Lancaster who both developed it independently of each other (Rayman, 1981). Although terminology like: 'abler pupils', 'helpers' and 'passed on to other students' might sounds like differentiated and cooperative learning, the 'helpers' mentioned only passed on what information the others needed to learn by heart. The method itself did not survive the 19th century, although it established the principle that students can learn from one another and that the teacher does not need not give all the instruction personally (Grittner, 1975). At the time, the Monitorial System was an exception to the rule. In general, all students worked in 'lock-step', meaning that all pupils in a class worked at the same rate in the same book and were required to master the same amount of material at the same level of thoroughness (Stauffer, 1959).

Due to industrialisation, the world had changed in the 19th century and The Netherlands had changed with it. Dutch society had changed from agricultural to industrial and this was reflected in mass public schooling that resembled a factory model in which children are taught in year groups where all receive the same treatment (Stoll Lillard, 2005).

#### **4.3.1.2. The Progressive Movement**

The Progressive Movement started in the last decades of the 19th century and also had an impact on education in the United States (Rollins, 2014). The drive to modernise education in the US included the application of scientific methods used on the study of education itself. This modernisation concerns the general ideas and practice of education and goes beyond the language teaching ideas of the Reform Movement (History of Education, 1992, p. 118). The Progressive Movement consisted of the whole spectrum of small scale, one man's reactions to flaws in the system, up to nation-wide publications on new ideas and theories to be applied. One example of an individual reactive attempt to change teaching practice that was perceived as flawed, was that of Principal Preston Search. In 1888, Search was the first educator to completely reject the lockstep method of mass instruction. Preston Search endeavoured with his teachers to make it possible for each child to learn at his own natural rate in each subject and he eliminated the concept of non-promotion, failures, grade-repetitions and grade skipping (Washburne, 1922; Winter, 1985). His work in Pueblo, Colorado lasted until 1894 and his approach disappeared with his departure, or as Washburn put it: "...continued only as long as Search's dynamic personality aroused the necessary enthusiasm." (Washburne, 1953, p. 140).

People like Montessori and Dewey, on the other hand, published and lectured about their far-reaching ideas, and their principles influenced educators throughout the world. Dewey, for instance, developed a philosophy of education to meet the needs of a changing democratic society. To him education was a miniature version of life, not merely a preparation for life. Dewey's writings, combined with the dissatisfaction with the existing schools, ignited the foundation of a number of experimental schools like the University of Chicago Laboratory School, founded in 1896, directed by John Dewey himself (History of Education, 1992, p. 81).

Underlying Montessori education is a model of children as active students. The children choose what they want to learn about, based on what interests them. To facilitate that, they are granted an enormous amount of freedom within a carefully designed, ordered structure. The structure, offered by the teacher, assisting children toward independence, guarantees that learning is always in context and activities are always part of the wider curriculum. The children learn with and from peers, work and talk with children of their choosing, in a classroom with three-year age groups (Stoll Lillard, 2005).

As a reaction to the year group system, and based on ideas like those of Montessori and Dewey, a number of plans, systems and approaches originated in the United States with the intention to promote individualized instruction. The relative success of these 'plans' led to a huge push for individualized instruction in the mid-1920s (Grittner, 1975).

Two of the most famous plans were those which became known as the Winnetka and Dalton plans, named after the cities in which they were developed (i.e. Winnetka, Illinois and Dalton, Massachusetts). The Winnetka Plan, originated by Washburne, had the teacher decide the exact amount of knowledge and skill to be mastered in the individualized subjects. Then each student was given a course of study for each subject in the programme with provisions for continuous progress including promotion of pupils on a non-graded basis. The plan attempted to expand educational focus to creative activities and emotional and social development. It was based upon the ideas of Burk used in the training school at San Francisco State College from 1913 to 1924 (Grittner, 1975).

The Dalton Plan was first developed by Parkhurst in 1919 in a school for physically disabled children, and, in 1920, in the Dalton, Massachusetts High School. The Dalton Plan was student-centred in the sense that the student needed to motivate himself, budget his own time, and evaluate his own learning pace, progress and performance. While in some versions the plan even allowed differentiated assignments through contracts agreed upon in advance (Grittner, 1975).

Despite the fact that these "Plans," have not been turned into regular mainstream education, it is clear that genuine efforts were made to implement them in many schools in the first quarter of the twentieth century. An analysis of bibliographies by Boegelein in the 1925 NSSE Yearbook shows nearly 500 articles could be documented on the subject of individualization between 1910 and 1924 (Grittner, 1975).

Language teachers were also involved in this movement. Handschin, for example, wrote an article in the *Modern Language Journal* in January 1919 with the following opening lines: "Little has been done in modern language teaching towards adapting work to suit individual differences. This is, however, one of the vital movements in our present day and the teacher of modern languages should not lag behind in regard to it" (Handschin, 1919, p. 158).

These general pedagogical ideas that belong to the Progressive Movement also had an impact on education in The Netherlands. The largest impact of these pedagogical ideas was to be found in primary education in The Netherlands. Between 1890 and 1940 quite a number of primary schools, still in existence today, were founded based on the ideas of people like Rudolf Steiner, Maria Montessori, Peter Petersen, Helen Parkhurst and Celestin Freinet (Imelman & Meijer, 1986).

Maria Montessori's first visit to The Netherlands in 1914, for instance, led to the founding of several schools specifically based on her pedagogical and educational ideas. On her return in 1923 she even managed to move the Dutch minister of education to create legislative dispensation for teaching outside the subsidiary conditions of the time. Almost a century later the Dutch Montessori Society (Nederlandse Montessori Vereniging) lists 42 nursery groups, 175 accredited primary schools and 19 secondary schools (<http://www.montessori.nl/>, 21-04-2016).

Endeavours to change education did not succeed overnight, nor did they take off without resistance.

Although a scientific approach to educational development might seem like a logical step forward from our present point of view, at the turn of the century some people still needed to get used to the idea of measuring the effectiveness of their educational practice.

Ayres (1912), for example, describes the difference in reaction of American school superintendents when they were confronted with tests and measurements of educational efficiency between 1897 and 1912. In 1897 the effectiveness of teaching spelling was questioned because data showed that after eight years of education, children did not spell any better whether they had studied spelling ten minutes a day, or forty minutes a day. The school superintendents of America did not take the findings well. In fact, as Ayres reports, the data presented "...threw that assemblage into consternation, dismay, and indignant protest...", not because of the data itself, but because of the investigator, "...who had pretended to measure the

results of teaching spelling by testing the ability of the children to spell.”. The school superintendents even unanimously “... voiced the conviction that any attempt to evaluate the teaching of spelling in terms of the ability of the pupils to spell was essentially impossible and based on a profound misconception of the function of education.” (Ayres, 1912, p. 300)

Fifteen years later the change in education is evident as, during the 1912 convention, the same association of school superintendents devoted forty-eight addresses and discussions to tests and measurements of educational efficiency. According to Ayres (1912) it was not just a shift in mind-set of the school superintendents. People at all levels started to measure what exactly happened in education. Ayres reports, for instance, how in three years’ time the number of cities in the US that kept track of the school histories of their children, by keeping individual record cards, had risen from 29 to 216. The number of city school systems using uniform records of accounting to enable mutual comparison of results rose from 15 in 1911 to 418 in 1912. Scientific data collection, testing and measuring greatly influenced the development of (language) education in general and more specifically the development of educational psychology.

#### **4.3.1.3. Aptitude Treatment Interactions**

American educational psychology, in which behaviourism was the prevalent discipline, was challenged at the end of the 50s. Cronbach (1957) pleaded in an article for the confluence of the experimental and correlational disciplines within psychology. He described how, since the 1920s, a schism had grown between the experimental (behaviourist, perception and learning) psychologists and correlational (personality, social, and child) psychologists. Experimental psychologists had stripped individual differences of their scientific importance because generalisations based on experiments with controlled stimuli and measured responses “could make and unmake individual differences at will” (Cronbach, 1957, p. 673). Correlational psychologists, on the other hand, posed that “...the individual and his motives, desires, wants, ambitions, cravings, aspirations...” (Cronbach, 1957, p. 673) is the real ruler of the psychologists’ domain of study. After describing how the two disciplines of scientific psychology historically differed, as well as describing how, in research questions, experiments and statistical instrumentation, the two disciplines are

related, he goes on to state that the disciplines themselves had started to turn the tide of separation in psychology.

In fact, Cronbach called for the combined application of experimental and correlational methods, “in order to find for each individual the treatment to which he can most easily adapt” (Cronbach, 1957, p. 681).

This development would lead to “...an educational psychology which measures readiness for different types of teaching and which invents teaching methods to fit different types of readiness. In general, unless one treatment is clearly best for everyone, treatments should be differentiated in such a way as to maximize their interaction with aptitude variables.” (Cronbach, 1957, p. 682). Measurement of the effectiveness of different treatments was referred to as ‘Aptitude Treatment Interaction’(ATI) research.

However, weak methodology, inappropriate hypotheses, and lack of replication made a lot of ATI research inadequate (Cronbach & Snow, 1969). It seems ATI research was not easily performed in educational research. In 1992, Snow described how in the 1980s many educational psychologists ignored or rejected ATI research because of the complexity of the problem and failure to heed methodological conclusions published. He even reports ATI as being listed as just one of the many fads educational research had lived through and discarded (Snow, 1992, p. 11-12).

ATI studies are still being performed today, albeit less frequently than might be expected. Vatz, Tare, Jackson and Doughty (2013) report, in a review of ATI studies, that ATI studies are often complex, and “a well-designed study can yield interesting and potentially useful results regarding how individual differences affect language learning” (Vatz, et al., 2013, p. 286).

#### **4.3.1.4. Differentiation in the final decades of the twentieth century**

In the United States, struggling and advanced students had been ‘helped’ in the past by pulling them from general education to group them homogeneously in special classrooms. This changed in 1975 when the Education for All Handicapped Children Act (EHCA), guaranteed:

“...that students with disabilities would receive as much of their education as possible with students who are not disabled” (Stronsnider & Lyon, 1997, p. 611).

Op 't Eynde (2004) describes how in Flanders, Belgium, the call for differentiation within the classroom started at the time of the 'Renewed Secondary Education' (Vernieuwd Secundair Onderwijs). This was the first half of the 1970s. He refers to publications at the time in Impuls Magazine containing pieces on Mastery Learning (Op 't Eynde, 2004).

Unfortunately, the substantial investments and developments in primary schools and teacher training mentioned above do not reflect the entire situation in The Netherlands. The variety in language skills acquired by primary school students has grown, but secondary school teachers have not responded by differentiating their language teaching accordingly.

Differentiation of education in general remains a recurrent theme in publications in the last decades. Blok (2004) described the introduction of 'adaptive education' in The Netherlands in the mid-80s. The term came from the United States where, in the 1970s, it replaced 'individualized instruction'. This term in turn indicated adaptive programmes like IPI (Individualized Prescribed Instruction), IGE (Individually Guided Education), LFM (Learning for Mastery) and ALEM (Adaptive Learning Environment Module) (Blok, 2004). These programmes were aimed at matching or adapting instructional variables to individual differences between students. In the 1990s in the United States the Individuals with Disabilities Education Act (IDEA), increased schools' responsibility for inclusion of students with special needs in general education classrooms. Students who before were previously in special classes and or remedial teaching classrooms and programmes now began to make up part of the regular classrooms, that became more heterogeneous (Gibson, 2005). The IDEA requirements of mainstreaming education applied to all exceptional students, because: "...there is no part-time solution to a full-time need: an hour a day or a half-day a week in a specialized program is not powerful enough to make enough difference in the learning of most students...." (Tomlinson, 2000, p. 4).

The educators at the time were looking for solutions and approaches that would help them deal with these heterogeneous classrooms. Differentiation could possibly be an answer to the needs of the students in heterogeneous classrooms although the term 'differentiation' and what it might mean had not yet been settled according to Oakes, Gamoran and Page (1992). The term 'curriculum differentiation' in their publication refers mostly to tracking; separating students by academic ability into groups for all subjects or

certain classes, and ability grouping; putting students in small, informal groups within a single classroom.

Tracking has a permanent and official character, while ability grouping can change per subject and students can move between ability groups over time.

The lack of clarity on differentiation was not a new phenomenon in 1992 because: "Despite more than 70 years of research and discussion, the United States, in particular, remains perplexed about curriculum differentiation" (Oakes, Gamoran & Page, 1992, p. 572). Some people were even against the use of differentiation: "Thus, ethical and pedagogical arguments have been mounted for a common curriculum, even if individual differences among students prevent them from benefiting equally." (Oakes, Gamoran & Page, 1992, p. 597).

In The Netherlands adaptive education received a boost in the 1990s because of the 'Weer Samen Naar School' government plans, meaning: 'together to school again'. Children with all sorts of learning disabilities, who had been placed in special educational needs schools and programmes, should go back to regular primary schools (Blok, 2004; Bulterman-Bos, 2007). In the decade and a half after the term had been introduced, adaptive education has shifted in meaning from 'adapting instructional variables to individual differences' to inclusive education, in which special educational needs students became part of regular classrooms. To realise this level of inclusive education teachers were expected to adapt their teaching to the possibilities and needs of the children. They needed to differentiate by fine-tuning their goals, methods and demands to what children could do. Differences between students needed to be approached positively by addressing students on their specific level of needs.

With the above described developments, clarity on the use of the terms 'differentiation' and 'adaptive education' in The Netherlands has not improved. Difference in meaning, although both terms accentuate the importance of dealing with differences within the classroom, are lost, as the terms have been used interchangeably by educators and the government. (Blok, 2004; Bulterman-Bos, 2007).

At the heart of these endeavours to work with heterogeneous classrooms are the experiences of teachers in regular education. These teachers had hands-on experience with the diversity of the students in their classrooms and found ways to make their teaching responsive to that diversity. In the United States one of those teachers, Carol Ann Tomlinson, has published extensively about her experiences and studies on the

potential of differentiated education. She started her teaching career in the early seventies and in her first teaching years she realized there were a number of prerequisites to effective education. These prerequisites have been formulated by Tomlinson and Imbeau (2010) as follows:

- when students are engaged they have no motivation to misbehave;
- if students understand that you see them as worthwhile people with significant potential, it opens doors to learning;
- an effective teacher teaches content to human beings with different starting points;
- each individual deserves a legitimate opportunity to grow as much as possible from his or her starting point;
- classroom management is the process of figuring out how to set up and orchestrate a classroom in which students sometimes work as a whole group, small group, and as individuals;
- the goal would be to have everyone work not only on things they all need to do in common, but also on things that were of particular importance for their own individual growth.

These ideas formed the basis for her later publications (e.g.: Tomlinson & Kalbfleisch, 1998; Tomlinson, 1999; 2001; Tomlinson & Allan, 2000; Tomlinson, Brighton, Hertberg, Callahan, Moon, Brimijoin, Conover & Reynolds, 2003; Tomlinson & Imbeau, 2010).

Just like Preston Search in 1888, Tomlinson changed classroom teaching practice to improve the situation based on her experience and view of what the children needed. The main difference between Search and Tomlinson was that Search's approach did not have following, despite being mentioned and discussed in publications. Tomlinson published about her endeavour to differentiate instruction, had following and continued publishing about her activities and that of others when she became an Assistant Professor at the University of Virginia in Charlottesville.

Guild and Garger (1998) were among the first to publish on differentiated instruction as a new topic in education towards the end of the twentieth century: 'What Is Differentiated Instruction? Marching to Different Drummers'. The body of research and publications on differentiated instruction has been growing ever since.

#### **4.3.1.5. Differentiation in the first decades of the twenty-first century**

In the meantime, in the United States the No Child Left Behind (NCLB) legislation of 2001 added an incentive for schools to actively document improvement for every child tested within the educational system. Data gathered from tests had to be reorganised to show advancement per subgroup as soon as a population of 30 or more students were tested. Results had to account for the advancements of the following subgroups: disadvantaged (socio-economic), disabled (special education), limited English speaking, migrant, gender, and ethnicity. As a result of this explicit organisation of data per subgroup, many districts realized that certain populations of their students were underperforming (Koeze, 2007). The data unequivocally showed teachers the effectiveness of their current teaching practice to heterogeneous classrooms. The heterogeneous classroom would not go away and teachers were not only expected to work with them, they also had to prove that every child received the education it needed through consistently improving test scores of individual students. Differentiated instruction was accepted, although not widely, as a possible solution.

Wide range implementation of methodological or pedagogical reform, like differentiated instruction, as an answer to the heterogeneous classroom, would need the broadly based support of governments, boards, teachers and publishers.

The real professional dilemma for educators and policy makers in the United States, according to George (2005), is acceptance of the task at hand. Although involved professionals accept the heterogeneous classroom and its teaching requirements (like differentiation), they are daunted by the next task of changing the way they teach. Despite the dilemma, George believes "...there are realistic strategies that can help teachers create diverse classrooms where authentic human learning is served, and where all students are successfully and meaningfully challenged" (George, 2005, p. 191-192).

From New Zealand Subban (2006) observes that student populations across the globe are changing significantly. Around the world, students bring more and diverse languages, cultural backgrounds and (dis)abilities into the classroom. Teachers, however, do not yet appear to have adjusted their methods.

With contemporary classrooms becoming increasingly diverse, educational authorities, teachers and school

administrators are looking to teaching and learning strategies that cater for a variety of learning profiles (Subban, 2006).

#### **4.3.2. Results (B) - The theoretical foundation of differentiation**

The literature found on the theoretical foundation of differentiated instruction has been structured to first build towards a workable definition. The subsequent paragraphs deal with research that forms the base for a change of current practice to a new educational approach. These reasons for a new educational approach consist of: the diversity of students present in regular classrooms, the effects of not responding to the diversity found, brain research that says something about diversity in students and learning and, to round off, what research says about learning styles.

##### **4.3.2.1. The theoretical foundation of differentiation: A workable definition of differentiated instruction**

As described in the historical review presented in chapter 4.3.1, it has taken some time for differentiated instruction to reach mainstream education. The 19th century saw the start of mainstream education in the western world. Education became available to all children, albeit based on the principles of mass industrial production and not differentiated. Students were grouped according to age, received whole class instruction and moved through the system from grade to grade.

At the start of the 20th century, ideas of pedagogical reformers such as Dewey, on how to differentiate education through 'individualized instruction', were radical at the time but did not influence public education as a whole.

When 'individualized instruction' was introduced more than half a century later, in the 1960s, differentiated education was only offered through ability grouping and streaming based on perceived capacity.

The first definition found for differentiated instruction, conceptually based on 'individualized instruction' is from the first half of the 1970s: *"Allowing differences develop between parts (e.g. schools, departments, year groups, subgroups, individual students) of an educational system (e.g. national teaching institutes,*

*college, department, year group) with respect to one or multiple aspects (e.g. aims, contact hours, instructional approaches)” (Own translation<sup>19</sup>; De Koning, 1973, p. 3)*

The importance of the above quote is that De Koning actually suggests a system-wide change, although in his definition the roles of the teachers and students do not receive direct attention. In the decades following the above publication, education has been influenced by constructivist theories (Hall, Vae, Strangman & Meyer, 2002; Tomlinson et al., 2003; Van Schooten & Emmelot 2004; Krol, Janssen, Veenman, & van der Linden, 2004; McTighe & Brown 2005; Subban 2006; Koeze 2007; Huebner, 2010). This influence explains the change from a mechanical view on differentiation, found in De Koning’s publication, towards the social view in which the student plays an important role.

Instead of giving a definition, Tomlinson and Allen (2000) described what in their view differentiated instruction entails. They described what, according to them, the required ingredients are that make up differentiated instruction as follows:

“Differentiation is simply attending to the learning needs of a particular student or small group of students ... The goal of a differentiated classroom is maximum student growth and individual success ... Demonstrating clarity about learning goals, both teachers and students understand that time, materials, modes of teaching, ways of grouping students, ways of expressing learning, ways of assessing learning, and other classroom elements are tools that can be used in a variety of ways to promote individual and whole-class success ... . In a differentiated classroom, a teacher sees everything a student says or creates as useful information both in understanding that particular learner and in crafting instruction to be effective for that learner... In a flexibly grouped classroom, a teacher plans student working arrangements that vary widely and purposefully over a relatively short period of time. Such classrooms utilize

---

<sup>19</sup> “Het doen ontstaan van verschillen tussen delen (b.v. scholen, afdelingen, klassen, subgroepen, individuele leerlingen) van een onderwijssysteem (b.v. nationaal schoolwezen, scholengemeenschap, afdeling, klas) ten aanzien van één of meerdere aspecten (b.v. doelstellingen, leertijd, instructie-methoden.)”(De Koning, 1973, p.3)

whole-class, small-group, and individual explorations ... All students consistently work with “respectful” activities and learning arrangements ... While the teacher is clearly a professional who diagnoses and prescribes for learning needs, facilitates learning, and crafts effective curriculum, students in differentiated classrooms are critical partners in classroom success.” (Tomlinson & Allan, 2000, p. 4-11).

Later researchers who tried to synthesise a definition from the available research and publications came to a comparable formulation (e.g.: Subban, 2006; Rock, Gregg, Ellis, & Gable, 2008; Huebner, 2010).

Huebner (2010) states that across the literature (Algozzine & Anderson, 2007; Rock, Gregg, Ellis, & Gable, 2008; Tomlinson, 2000) experts suggest a number of guiding principles to support differentiated classroom practice. These principles are (Huebner, 2010, p. 81):

- “Focus on the essential ideas and skills of the content area, eliminating ancillary tasks and activities;
- provide choice and keep options open for influence in the planned process and product;
- respond to individual student differences (such as level of engagement and readiness, prior knowledge, interests, and learning profile preferences);
- group students flexibly by shared interest, topic, or ability;
- integrate ongoing and meaningful formative assessments with instruction, or use them as instruction;
- continually assess; reflect; and adjust content, process, and product to meet student needs.”

These principles formulated seem to put the teacher-student relationship at the heart of differentiated instruction. The focus is more on the student and the process needed for students to become partners in the educational process. Despite the wider social paradigm of the principles described above by Huebner, the definition of De Koning still stands. De Koning describes the effect of differentiated instruction, Tomlinson and the others describe what needs to happen to differentiate instruction. Since differentiating instruction means changing the entire complex educational process, including all contributing factors, a definition is not sufficient. The long description of required ingredients by Tomlinson and Allen given above, has been condensed by Tomlinson (2014) when she discusses the practice of teachers that

differentiate instruction because they: "... strive to do whatever it takes to ensure that struggling, advanced, and in-between learners, students with varied cultural heritages and children with a broad array of background experiences all grow as much as they possibly can each day, each week, and throughout the year." (Tomlinson, 2014, p. 2)

The shortest description found are the guiding principles Huebner mentions, although each principle represents quite an arduous task for the teacher when applied conscientiously. Unfortunately, this list of guiding principles, as given on the previous page, is also not an easy recipe for success, because as Rock, Gregg, Ellis and Gable (2008, p. 4) say:

"Although teachers express a desire to meet the needs of all of their students, often excessive workload responsibilities, demands for substantial content coverage, and negative classroom behaviour make the challenge seem insurmountable."

#### **4.3.2.2. The theoretical foundation of differentiation: A research base for a new educational approach**

##### **4.3.2.2.1. Diversity of students**

##### **4.3.2.2.2. Effects of (not) responding to the diversity**

##### **4.3.2.2.3. Brain Research**

##### **4.3.2.2.4. Learning style**

According to Subban (2006), the current educational system offers inadequate possibilities to treat every student with respect, does not make it possible to offer every student an engaging learning experience or to help every student to reach his or her potential. Teachers often work with a one-size-fits-all approach in which they focus on exposing and remedying deficits. By doing so, a number of students will experience a pattern of failure. Subban's position is that the alternative approach is to accept and build on the basis that students are all essentially different.

The research base for the ways in which students differ, is dealt with below (4.3.1.2.1.). Apart from academic diversity and effects of (not) responding to diversity in teaching (4.3.1.2.2.), the physiological differences of students' brains (4.3.1.2.3.) and the way they learn (4.3.1.2.4.) are discussed.

#### **4.3.2.2.1 The theoretical foundation of differentiation: Diversity of students**

Research from around the globe states that today's classrooms are typified by academic diversity (Tomlinson et al., 2003). In The Netherlands, Bosker (2005) claims it is generally safe to say that by halfway primary education there is a difference of approximately 4 school years between the ten percent of best performing students and the ten percent least performing students with regard to language skills (Bosker, 2005, p. 4, own translation<sup>20</sup>). The diversity of students is not, however, limited to their academic advancement. Sprenger (2003) tells us of teachers faced with the challenge of teaching to a diverse group of students in the modern Australian classroom. He states that in this situation one-size-fits-all lesson plans are no longer feasible for students who have different cultural backgrounds, home lives, and learning styles. In The Netherlands, Oostdam (2009) reports that according to many, the school population is becoming more diverse and the current generation has more specific interests and needs than before. A changing society and a youth culture in which new means of communication play a dominant role, have contributed to this diversity (Oostdam, 2009).

Apart from academic and cultural diversity, students also differ individually in the role their personality plays in the educational process. In her article, Horwitz (2000) provides a historical overview of how in the past century the partnership of teachers and students has evolved. Although she cites Handschin's 1919 article on adapting work to individual differences, in her view the prevailing images of students changed over time. Horwitz found a marked change in labels used for students in the 1970s: "The terms good and bad, intelligent and dull, motivated and unmotivated have given way to a myriad of new terms such as integratively and instrumentally motivated, anxious and comfortable, field independent and field sensitive, auditory and visual" (Horwitz, 2000, p. 532).

---

<sup>20</sup> "Meer in het algemeen kunnen we stellen dat tussen de top tien procent presterende leerlingen en de onderste tien procent presterende leerlingen halverwege het basisonderwijs een verschil van ruim vier schooljaren aanwezig is op het gebied van taal." (Bosker, 2005; 4)

As Ellis (2004) puts it, the main purpose of research into individual differences shifted from the selection of students to receive foreign language instruction, towards the prediction of student's language acquisition success and, around the end of the century, towards an explanation of why some students succeed more than others. The applied version of this research matches students to different types of instruction in aptitude–treatment interactions. The top five individual differences recognised in research in the early years of the 21st century were, according to Ellis (2004, p. 528): Language aptitude (1), Learning styles (2), Motivation (3), Anxiety (4) and Personality (5).

Ever since the 1970s, the driving force behind individual difference research is the quest for a “Grand Unified Theory of Individual Differences” (Dewaele, 2012, p. 160). This unified theory acknowledges the complex interplay of independent variables of the student's psychological profile and particular context from a dynamic perspective.

In a 2015 review of the changes and developments in individual differences research in the past decade, Dörnyei and Ryan confirm the abovementioned ‘dynamic turn’ the field of applied linguistics has been undergoing. Instead of investigating distinct variables in isolation, scholars have taken greater interest in examining the dynamic, often nonlinear interactions, between individual difference components (Dörnyei & Ryan, 2015). In their review, research on **Language aptitude** is, like in 2015, still in transition. Aptitude is researched as a complex measure, comprising all factors affecting the student's capacity to learn a language, making aptitude research more pedagogically relevant. Research on **learning styles** had not progressed much between 2005 and 2015. Although mainstream educators have embraced the importance of learning styles, researchers are far less positive; styles research and theory being “...a field based more on hope than substance...” (Dörnyei & Ryan, 2015, p. 107). Learning styles are even listed as a science myth because of, as Scudellari (2015) reports, a disturbing lack of credible evidence.

According to Dörnyei and Ryan (2015), the advantage of the renewed focus on cognitive style, as field of research prior to learning styles, is that it allows for a ‘purer’ definition, as it is devoid of any educational and situational/environmental interferences. Cognitive style is deemed to be a stable and internalized dimension related to the way a person thinks or processes information. Dörnyei and Ryan (2015) conclude however, after discussing current leading models of cognitive styles and their assessment (Dörnyei & Ryan,

2015, p. 114-123), that learning styles and cognitive styles could be seen as a situated and interactive dynamic concept to be explored and researched, as they “...constitute an as yet unrealized potential.” (Ushioda & Dörnyei, 2009, p. 1).

**Motivation** research is currently in the process of being radically reconceptualised in the context of ‘self’ and ‘identity’. The L2 Motivational Self System, as first presented by Dörnyei (2005), is a self-specific motivation construct that mediates and controls ongoing behaviour, in which people view themselves in three dimensions. These three dimensions consist of the Ideal L2 Self (the person people would like to become in speaking an L2), the Ought-to L2 Self (the attributes in speaking an L2 that people believe they ought to possess) and the L2 Learning Experience (‘executive’ motives people have related to the immediate learning environment and experience).

**Personality** traits are consistent patterns of human behaviour of which the most widely used taxonomy in dimensions is the ‘Big Five’: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism (Dewaele & Al-Saraj, 2015).

Education has the complex task of dealing with the diversity found in students. As discussed above, this diversity reveals itself in quite a number of dimensions like intelligence, personality, social circumstances, etc. Schools are required to offer education and care where needed and to consistently work on fine tuning the relationship between the educational environment and student characteristics (Sligte, Bulterman-Bos, & Huizinga, 2009).

#### **4.3.2.2.2 The theoretical foundation of differentiation: Effects of (not) responding to the diversity**

Conformation and limitation of all educational activities to textbooks or standardised tests, creates a number of pitfalls. Ignoring the aforementioned differences in one-size-fits-all education not only deprives students of suitable instruction and activities, it also requires the teacher to invest more time and effort in the need to keep all students on the same track. According to Subban (2006), single-paced lessons with a singular instructional approach disregards differences between students present in all classrooms. In his opinion:” Ignoring these fundamental differences may result in some students falling behind, losing

motivation and failing to succeed. Students who may be advanced and motivated may become lost as the teacher strives to finish as much of the curriculum as possible.” (Subban, 2006, p. 938-939).

The quite sizeable group of students in the middle will play an important part when teachers start to differentiate. Choices for the approach used to differentiate can have extensive implications. How teachers differentiate and to what end may differ greatly, as Deunk, Doolaard, Smalle-Jacobse and Bosker (2015) described when they reported on convergent differentiation and divergent differentiation.

Reezigt (1993) already found that on average, performance of all students will increase after a period of undifferentiated education. When teachers are mainly focussed on students with specific educational needs and/or deficits, the differences between student achievement will be less varied at the end of the educational period. The initially underachieving students will make relatively great progress, as they receive more attention to help them meet the minimal attainment targets. Students who have already met, or will easily meet the minimal attainment targets are not challenged and will make relatively less progress than the aforementioned students with specific educational needs and/or deficits. As this approach brings the achievement of both of students closer together, teachers use convergent differentiation.

Education differentiated to the extent that the most optimal situation is created for all students, leads to the situation in which all students make a significant leap forward and the spread of achievements might even become wider. Because of this effect it can also be labelled as divergent differentiation (Op ’t Eynde, 2004). The effect of a wider spread of achievements and abilities is part of the practical consequences of differentiated instruction. In fact, introducing differentiated instruction might possibly have consequences for almost every part of the educational system.

The choice, proposed as a dichotomous situation by Subban (2006), is to differentiate teaching towards the needs of the students or to have the students work towards the standard. When all students are offered the same teaching, the teacher must work hard to keep all students ‘on track’. Reezigt (1993) however, already warned that when student differences are addressed through differentiation, a wider spread of achievements and abilities might need to be accepted. If students no longer work towards the same standard, the assessments and grading system used might also need to be differentiated. In turn, it is possible that changes in the way schools grade their students would influence the streams and year groups

currently used in Dutch secondary education and even the final exams. As Tomlinson (2014) explains in 4.3.1.1, for example, differentiating instruction is more than just an expansion of the methodological language teaching repertoire; it requires a change in the practice of (language) teaching itself. This means that the choice to differentiate education needs to be taken in awareness of the possible consequences.

#### **4.3.2.2.3. The theoretical foundation of differentiation: Brain Research**

Brain research also endorses the idea that educators need to take student differences into account. This discipline is part of the 'Mind, Brain Education' mix Tokuhamo-Espinosa (2008) used in her study, which was constructed as a mixed-method design, followed by a Delphi expert survey in 2008. According to Tokuhamo-Espinosa, the Delphi expert panel agreed that all brains are unique. "This uniqueness implores educators to differentiate their practice and to treat students individually in order to aid them to reach their learning potential" (Tokuhamo-Espinosa, 2008, p. 22). The experts in her Delphi-panel agreed that emotions are a critical factor to learning, to detecting patterns and to decision-making. Positive emotions and challenges enhance learning in the same way negative emotions and threats inhibit learning (Tokuhamo-Espinosa, 2008), which means that a safe learning environment is also important from a neuroscientific view on education.

Students differ internally, brain-wise at least as much as might be observed on the outside, as described in 4.3.2.2.1. Diversity of students. The need to deal with the diversity found in students is underpinned by the research above (e.g. Tokuhamo-Espinosa, 2008), requiring education that offers appropriate challenges, possibilities to make sense of the ideas and skills through significant association and a safe learning environment to do so.

#### **4.3.2.2.4. The theoretical foundation of differentiation: Learning style**

Although a causal interpretation of the relationship between personality and educational attainment is seen as problematic (Van Eijck & De Graaf, 2004), the Big Five framework of personality traits, consisting of scales on Agreeableness, Conscientiousness, Emotional Stability, Extraversion and Openness, has been referred to as a robust and parsimonious model for the relationship between personality and academic

behaviour (Komarraju, Karau, Schmeck, & Avdic, 2011). Although there is evidence of personality types having an effect on how people learn (Miller, 1991; Fallan, 2006), according to Poropat (2009), the role of personality in academic performance "...is also subtle, complex and in need of much further exploration." (Poropat, 2009, p. 40). The complexity is evident from the widespread factors linked to learning as, in the theory on learning styles referred to, learning is impacted by factors like ambience, time of day, classroom organisation, pedagogical approach, student mobility, freedom and demand for concentration (Tomlinson et al., 2003). No explicit recommendations for language teaching methodology based on personality traits have been offered, but despite some opposition against working with learning styles (e.g. Rock et al., 2008; Landrum & McDuffie, 2010), mainstream educators seem to find the use of learning styles efficacious. Tomlinson et al. (2003) warns that in her opinion no particular approach to learning is superior to another and that there is great variance in learning preferences among cultural and gender groups.

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment, and for which this literature study had been initiated, has a rationale for its implementation. The diversity found in students, the adverse effects of a one-size-fits-all approach to teaching, the advances in brain research and the knowledge of the different ways in which students learn, indicate a clear, foundational need for differentiated instruction.*

### **4.3.3. Results (C): Building blocks of differentiated instructions (DI)**

#### **4.3.3.1. Students as partner in the educational process**

#### **4.3.3.2. Readiness, Interest and Learning Profile**

**Readiness**

**Interest**

**Learning Profile**

#### **4.3.3.3. Emotions**

#### **4.3.3.4. Motivation**

#### **4.3.3.5. Content, process and product**

**Content**

**Process**

**Product**

#### **4.3.3.6. Grading and assessments**

#### **4.3.3.7. Feedback**

#### **4.3.3.8. Grouping students**

This section takes a closer look at the evidence available on the guiding principles that support differentiated classroom practice. As can be seen from the list of topics for 4.3.2., the ‘building blocks’ that make up differentiated instruction, these aspects are concerned with the differences between students as well as the practical curricular requirements teachers have to work with. As explained in 4.3.1.1., differentiated instruction is best described as a review of prerequisites and guiding principles. The following section looks at the literature available on these prerequisites and principles that make up differentiated instruction ‘Building block’, by ‘building block’. The chapter starts with the general idea of having ‘students as partner in the educational process’, as they focus on the essential ideas and skills together with the teacher, who provides choice and allows autonomy to engage students towards optimal involvement in the educational process where possible. The next topic is student differences in ‘readiness, interest and learning profile’, the three main student characteristics to which teachers are expected to respond. These are followed by ‘emotions’ and ‘motivation’ because of the key role they play in the effectiveness of teaching activities and their interaction with the previous topic. The final topics are: ‘content, process and product’ as this is the part of the educational process that can be differentiated; the ways in which ‘feedback’ and ‘grouping students’ helps to make teaching more effective; rounding off with ‘grading and assessments’. As explained in the introduction, each section is rounded off with the implications of the findings for a language teaching programme based on differentiated instruction in italics.

#### **4.3.3.1. Building blocks of DI: Students as partner in the educational process**

Having students as partners in the educational process is not a new idea. Early pedagogical experiments on student autonomy after the Second World War arose from the 'counter-cultures' of late 1960s Europe (Gremmo & Riley, 1995). Practically speaking, these experiments focused on self-directed learning through self-access centres.

Student autonomy at the time was defined as "the ability to take charge of one's own learning" (Holec, 1981, p. 3). According to several literature reviews (e.g. Dickinson, 1995; Lai, 2011), substantial evidence can be found for student autonomy having a positive effect on educational achievements. Motivation increases when students are allowed to make decisions about their own work. When students are given choices they are more likely to be interested in the work, show more perseverance and other self-regulated learning behaviours. In fact, "learning success and enhanced motivation is conditional on learners taking responsibility for their own learning, being able to control their own learning and perceiving that their learning successes or failures are to be attributed to their own efforts and strategies rather than to factors outside their control." (Dickinson, 1995, p. 174). These effects were also found with students in self-directed learning programmes, who obtained a more positive view on their abilities, success expectations, accomplishments and the educational challenges offered (Lai, 2011).

According to Tomlinson and Allan (2000), continuous involvement of the students in making decisions about the learning process and products that show their knowledge, understanding and skills will result in students becoming more independent as students (Tomlinson & Allan, 2000). This partnership of sharing the responsibility for what happens in the educational process will help to engage students. Engaged students find added value in the learning process, or as Subban (2006) puts it: "Engaging students actively in the learning process and in the content allows them to see patterns developing, to see the overlap between disciplines, to see learning as a cumulative whole" (Subban, 2006, p. 941). Engaging students in purposeful, active, and inquiry-driven teaching and learning activities will ensure students learn best, say McTighe and Brown (2005). "The more learners are situated at the centre of their own learning process, the greater the extent of their understanding and mastery of desired outcomes will be" (McTighe & Brown, 2005, p. 236). A sizeable amount of research points towards the need for a more active role of the student to make the educational process more effective. In this educational process it is vital to achievement and

student satisfaction for students to be allowed to make choices and see the link between their learning and personal interests (Koeze, 2007). This move towards involvement also comes from the students themselves. The Dutch educational inspection service (2006) reports one of the challenges found to be exactly this situation. "Students and students increasingly put their teachers under pressure to be taken more seriously in their personal educational demands" (Own translation<sup>21</sup>, Inspectie van het Onderwijs, 2006, p. 232).

When self-directed education is proposed as means and aim for education, the stress lies on the student as active partner in the educational process" (Own translation<sup>22</sup>, Op't Eynde, 2004, p. 6). This partnership goes beyond being allowed to choose every once in a while or knowing what is expected. McTighe and Brown (2005) say students not only need to understand the learning goals but also see them as meaningful and personally relevant. students in their view: "...must own the learning goals for which they are responsible, and demonstrate growing capacity to articulate the connectivity of what they learn to their world beyond the classroom" (McTighe & Brown 2005, p. 240).

One of the ways in which students can be made aware of the relationship between their personal language skills levels and the learning goals is through the use of rubrics (Panadero, Tapia & Huertas, 2012).

Airasian and Russell (2008) explain rubrics as a set of clear expectations or criteria teachers as well as students use to help focus on what is valued in a subject, topic, or activity. A rubric shows what aspects will be assessed and a description of the criteria used to assess each aspect, including a scale for grading the different levels of achievement and a description for each qualitative level (Panadero, Tapia & Huertas, 2012).

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment, it seems clear students need to, at least, be offered choice in both the educational process as well as the product to help assess learning. Students should also have the*

---

<sup>21</sup> "Leerlingen en studenten oefenen een toenemende druk uit op hun docenten om hen meer serieus te nemen in hun eigen leervragen." (Inspectie van het Onderwijs, 2006; p.232)

<sup>22</sup> "Zelfgestuurd leren wordt naar voren geschoven als middel en als doel voor het onderwijs. De klemtoon komt hierbij te liggen op de leerling als de actieve partner in het onderwijsgebeuren" (Op't Eynde, 2004, p6)

*opportunity to offer contributions from their personal interests. On top of this influence offered to the students, they become partner in the educational process when goals are set in a dialogue between general institutional demands and a student's personal educational demands. This dialogue can use rubrics of language skills levels as a starting point. Students can, for instance, indicate their personally perceived level and teachers can indicate the required, expected, or desired level to reach within a chosen timespan, after which the dialogue can revolve around the effort needed to bridge the gap between the two indicated levels.*

#### **4.3.3.2. Building blocks of DI: Readiness, Interest and Learning Profile**

##### **Readiness**

The readiness to work on something is one of the student characteristics for which teachers can differentiate. (Tomlinson & Allan, 2000).

" Challenges, however, must be at the proper level of difficulty in order to be and to remain motivating: tasks that are too easy become boring; tasks that are too difficult cause frustration." (Bransford, Brown & Cocking, 1999, p. 61). Readiness is also influenced by the extent to which teachers offer scaffolding (Tomlinson & Allan, 2000). By devising support and material to support all students, teachers create an atmosphere for success for all students. (Subban, 2006). *Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment needs to help teacher estimation of student advancement. Teacher assessment together with student self-assessment can help the teacher relate expected individual learning gains to individual achievements. This is an important factor of the partnership with students in the educational process, as mentioned in 4.3.3.1. This approach makes learning gains relative to individual student capacity and commitment to which challenge and scaffolding can be measured.*

##### **Student Interest.**

"The varying student interests found in our classrooms can become effective tools to support learning, because most students, even struggling students, have aptitudes and passions. Giving students the

opportunity to explore and express these interests through activities built around them, will allow the curriculum to become more meaningful to them” (Subban, 2006, p. 941).

In general, research seems to support the idea that teaching in response to student interest contributes to a sense of competence and self-determination in students, positive learning behaviours like the willingness to accept challenge (Csikszentmihalyi, Rathunde, Whalen, 1993; Fulk & Montgomery-Grymes, 1994; Vallerand, Gagne, Senecal, & Pelletier, 1994; Zimmerman & Martinez-Pons, 1990) as well as a positive attitude about learning itself (Amabile, 1996; Runco & Chand, 1995; Torrance, 1995).

One way to work with different students’ interests is to link subject matter to personal experiences. This approach is reported to be effective for second language learners (Kasper, 1997; Echevarria & Short, 2000). Teaching activities linked to student interests will hold their attention, although, it is important to remember students learn best when there is a change of person, place, topic or activity every 20 minutes. This is of course because the average student has an attention span of between 10 and 20 minutes. For the students to reflect on new information teachers should allow for “down time”, the same way they should give all students a minute to think about questions before demanding an answer (Tokuhama-Espinosa 2008).

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment should endeavour to link subject matter to personal experiences and interests through varied activities.*

### **Student Learning Profile**

The learning profile is what teachers need to know about their students to effectively differentiate instruction. A learning profile is, according to Tomlinson and Imbeau (2010), shaped by learning style, thinking preference, gender, culture and personality. Tomlinson and Allan (2000) published examples of how to work with different learning profiles in the classroom. They suggest using flexible spaces, grouping, learning options, activities and the use of auditory, visual, and kinaesthetic modes for students to explore information and ideas. These activities, which consist of a choice of competitive, cooperative, and independent learning experiences, can be done alone or with peers. In this way, educators balance varied

perspectives on an issue or topic while providing authentic learning opportunities in various intelligence or talent areas.

*Given the above research and publications, the treatment should at least contain some cooperative learning activities, and choices, as these are reiterated suggestions. The suggestion to work flexibly with spaces, grouping, learning options, modes might be beyond the possibilities of this research and the treatment planned.*

#### **4.3.3.3. Building blocks of DI: Emotions**

As an undercurrent of readiness levels, interests and learning profile, students bring different emotions into the classroom. Emotions and feelings influence the effectiveness of teaching (Damasio, 2000; Immordino-Yang & Damasio, 2007). Although research on the impact of emotions on learning is relatively new (e.g., LeDoux, 2008; Reeve & Jang, 2006), it is quite clear that students who feel enthusiastic, appreciated and safe will be more engaged and willing to learn than students who are in distress, or anxious (Skinner, Kindermann, & Furrer, 2008). Dörnyei and Ryan (2015) describe how past research on student characteristics has suffered from a general 'emotional deficit'. Emotions and reason were seen as completely separate and research on language acquisition had no place for something as irregular and unpredictable as emotions.

Emotions, however, are part of the educational process in general, or as Pekrun, Frenzel, Goetz and Perry (2007) put it: "Because of their subjective importance, educational settings are infused with intense emotional experiences that direct interactions, affect learning and performance, and influence personal growth in both students and teachers" (Pekrun et al., 2007, p.13). Emotions also play an important role in language learning. They are called the driving force behind foreign language learning (Dewaele & Al-Saraj, 2015) and deemed to be at the heart of the teaching foreign languages (Kramsch, 2009). In Kramsch's view this includes students and teachers. Her advice is to find something to love or hate about what we are required to teach, because teacher's indifference leads to student's boredom (Kramsch, 2009).

Pekrun et al. (2014) published about Control-Value Theory, which assumes a general functional mechanisms of human emotions through which "...achievement emotions can be influenced by changing

subjective control and values relating to achievement activities and their outcomes.” (Pekrun et al., 2014, p. 31-32).

*Given the above research, the language teaching programme based on differentiated instruction to be used as treatment needs to use recommendations like: Keep in mind that enthusiasm can be reciprocally induced. Educational activities need to be cognitively stimulating; as much as possible self-regulated; in cooperation with peers. Goal structures need to be individualistic and pertaining to mastery of goals. Feedback needs to be given on both success and failure (at learning). The long-term consequences of achievement need to be positive and lead to student self-regulation of their control and value appraisals.*

#### **4.3.3.4. Building blocks of DI: Motivation**

Interest levels and emotional states are linked to positive motivation (Halpern & Hakel, 2002) and the psychological claim that motivation impacts learning is well documented according to Tokuhama-Espinosa (2008). It is clear that motivation, like success experience and interest in learning goals promote learning (Krapp, Hidi, & Renninger, 1992; Turner, 1993; Alexander, Kulikowich, & Jetton, 1994; Tobias, 1994).

The implication for educators is that they bear a substantial part of the responsibility for the motivation of their students. Students’ level of motivation brought into the classroom will always be influenced (positively, or negatively) by the teacher (Halpern & Hakel, 2002). So, teachers can influence student motivation by what information they offer and how. Students will be more motivated to process knowledge that makes sense for them to learn and is delivered at their level, offering a success experience. Students will not only become more motivated; research into 'depth of processing' claims that the presentation of coherent and meaningful information also leads to a deeper form of processing (Anderson, 1990; Barsalou, 1992; Stillings et al., 1987).

Motivation to learn is not only influenced by what students receive in the educational process. Students need to produce language before they are able to ‘notice the gap’. This ‘noticing the gap’ means students experience first-hand, while producing language, what it is they do not yet know. This experience clarifies for the student the need to learn and work with language. According to Ellis (1994) ‘noticing the gap’, or students’ observation of personal deficiency (Swain, 1995; 2001) is essential to learning a language.

For students to come to a deeper form of processing when acquiring a language, the work needs to be coherent and meaningful and lead to production for them to observe their personal deficiency and notice the gap.

Cenoz (2003) reports that in her longitudinal study students' motivation to learn a foreign language decreases over the years they progress from primary to secondary education due to their growing rejection of the school system in general as well as to the change in language teaching methodology. This goes for education in general, as Davies and Brember (2001) observed, where the more years students spend studying a subject, the more disenchanted they become with traditional language learning and teaching. The Basque students in the study Cenoz reports on, who had enjoyed the oral-based approach in primary education, had, as Cenoz claims, less positive attitudes towards language learning after experiencing the more teacher-centred grammar and vocabulary teaching methodology in secondary education (Cenoz, 2003).

This literature study, as reported on in chapter 4, focuses on regular English language teaching in the first year of secondary education in The Netherlands; content and language integrated learning (CLIL) is beyond the scope of this study. The CLIL approach is, however, still worth mentioning from a motivational point of view. The 'coherent and meaningful information' used in education that would lead to a deeper form of processing, as mentioned in the first paragraph, can also be related to subjects other than English. CLIL, as presented by Lasagabaster and Sierra (2009), seems to yield positive results in the sense that it helps to keep students interested in learning foreign languages. Or, as they put it: "Learners' motivation to learn content through the foreign language may foster and sustain motivation towards learning the FL itself." (Lasagabaster & Sierra, 2009, p. 14). In the Dutch context, CLIL applied in the highest tier secondary schools offering a bilingual (roughly 50%) programme, has consistently yielded positive effects (Huibregtse, 2001; Verspoor, Schuitemaker-King, Van Rein, De Bot, & Edelenbos, 2010).

Historically speaking, motivation has basically been divided into intrinsic and extrinsic motivation. Intrinsic motivation comes from within the student who enjoys the process and feels the language is interesting. Students who are extrinsically motivated participate in the educational process to avoid punishment or gain rewards. Language learning motivation research has progressed enormously beyond this simple dichotomy.

For more than half a century in both fields of second language acquisition and educational psychology study, goals, attitudes and self-related beliefs have served as central constructs for models of motivation. Reviews of which can be found in several publications (e.g. Kormos, Kiddle & Csizér, 2011; Dörnyei & Chan, 2013).

As Moskovsky, Alrabai, Paolini and Ratcheva, (2013) report, the number of research papers on the practical approach to implementing motivational strategies is small. In their research on the effects of teachers' motivational strategies on student motivation, they used the following 10 strategies:

- 1) Break the routine of the classroom by varying learning tasks and the presentation format;
- 2) Show students that you care about their progress;
- 3) Show students that you accept and care about them;
- 4) Recognize students' efforts and achievements;
- 5) Be mentally and physically available to respond to your students' academic needs in the classroom;
- 6) Increase the amount of English you use in the language classroom;
- 7) Make learning tasks more attractive by adding new and humorous elements to them;
- 8) Remind students of the importance of English as a global language and the usefulness of mastering the skills of this language;
- 9) Relate the subject content and learning tasks to the everyday experiences and backgrounds of the students; and
- 10) Consistently encourage students by drawing their attention to the fact that you believe in their effort to learn and their capabilities to succeed (Moskovsky et al., 2013, p. 41-42).

The implementation of these strategies resulted in a significant positive change in student language learning motivation.

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment should try to lead students to a deeper form of processing. This could be achieved by production of language through which they can observe their personal deficiency and notice the gap, because of which, ensuing language production and revision would become more coherent and meaningful. Furthermore, it is interesting to see that numbers 2, 3, 4, 5 and 10 of Moskovsky are actually*

*about the positive pedagogical relationship. The other strategies that deal with methodology can be implemented, but strategy 9 stands out. Generally speaking, it is logical to build on the relationship between the students' everyday life and the language learning activities. Maybe this should be taken one step further and also explicitly show the students how the language learning activities relate to the language skills and abilities students are working towards.*

#### **4.3.3.5. Building blocks of DI: Content, process and product of the educational process**

##### **Content**

When content is discussed it includes what students need to learn as well as how they gain access to the desired knowledge, understanding, and skills (Tomlinson & Allan, 2000). Rethinking the approach to unlocking curricular content for a diverse student population should be successful. By defining the organizing principles of the content to be taught, teachers can move from isolated acts to connections between and among ideas. Exploration of implicit connections and underlying principles of a topic can facilitate transfer of learning and ultimately support student performance on standardized tests (Bransford, Brown, & Cocking, 2000). Teachers should help students to come to understand the big ideas found wrapped in the curriculum standards and revisit those multiple times during the course of their education. Offering decontextualized facts and skills in a fragmented curriculum will never maximize student achievement (McTighe & Brown 2005).

*Given the above research and publications, for the language teaching programme based on differentiated instruction to be used as treatment it is important to ask the teachers to stress the connections between the language learning activities and the abilities students work towards; they should, in fact, contextualize facts and skills towards communication in a foreign language.*

##### **Process**

Differentiating the educational process refers to type, form and sequence of activities with which the students come to understand and assimilate facts, concepts, or skills (Algozzine & Anderson, 2007; Tomlinson & Allan, 2000). A teacher has to make sure students work on essential skills to come to

understand essential ideas, and that they are clearly focused on a learning goal. He can pace support, offer choices about how to express what students have learned (Huebner, 2010). This does not imply that each student is offered a different task, but just enough flexibility in how complex the task is, how students are allowed to work and present their achievement (Tomlinson & Allan, 2000). Lawrence-Brown (2004), for example, explains in his publication how to adapt the classroom curriculum and utilize all different teaching aids available to address student variety.

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment should allow time for revision and discussions on progress with peers and the teacher. This implies assignments that extend well beyond lesson level.*

### **Product**

In education products are usually the scores of tests. There are, however, other ways for students to demonstrate the knowledge, understanding, or ability they have gained. This could be more than just a student portfolio, exhibition or project. "A good product causes students to rethink what they have learned, apply what they can do, extend their understanding and skill, and become involved in both critical and creative thinking" (Tomlinson & Allan, 2000). Or, as Anderson (2007) puts it: "Differentiating the performance measure or product component of a lesson means affording students various ways of demonstrating what they have learned from the lesson or unit of study" (Algozzine & Anderson, 2007, p. 51).

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment should work towards products related to the ability level expected of the individual student and related to student's previous achievements. In that way the appreciation of the products can be fully differentiated.*

#### **4.3.3.6. Building blocks of DI: Grading and assessments**

As Stiggins, Arter, Chappius and Chappius (2004) describe, assessment of learning is done summatively, after learning is supposed to have happened and to measure if it did. The outcome is usually presented in

grades (scores, percentages, numbers/letters). The assigned grades evaluate the quality of student work, are used to communicate progress and achievement, motivate how students study, assess their focus and involvement in the course and effort on the final educational goal, mark transition or bring closure to (part of) an educational programme. Summative assessments are used in classrooms for unit and chapter tests, nationwide standardized exams and anything in between.

Assessment *for* learning is used formatively, while students are learning. The outcome is used to diagnose progress and needs, give feedback to help students on their way to achievements. For that reason, usually these assessments are not graded; they are concerned with how students are doing instead of how they have done.

In 2005 Tomlinson wrote an article about the compatibility of quality differentiation and quality grading (Tomlinson, 2005) and showed the perceived incompatibility of differentiating instruction and grading to be based on misconceptions. Assessments are central to curriculum design and instruction, but only useful when they inform teachers and students (Wiggins & McTighe, 1998). In their extensive review of research on activities that inform and help modify teaching and learning activities, Black and Wiliam (1998) discuss the efficacy of formative assessment based on the link between formative assessment and positive student outcomes. They also found many studies illustrating that teachers rarely use formative assessment systematically to provide information on their teaching.

Curriculum, instruction, and assessment need to be tightly aligned to make sure assessments do reflect the actual learning as planned (Moon, 2005). This alignment is also important for the effect the tests themselves have on the students, teachers and the educational process itself. This is called the 'washback' effect. The 'washback' effect is described by Buck (1988) as: "...a natural tendency for both teachers and students to tailor their classroom activities to the demands of the test, especially when the test is very important to the future of the students, and pass rates are used as a measure of teacher success" (Buck, 1988, p. 17). Or, as Swain (1985) puts it: "It has frequently been noted that teachers will teach to a test: that is, if they know the content of a test and/or the format of a test, they will teach their students accordingly" (Swain, 1985, p. 43). This effect can be harmful or beneficial when tests, or testing techniques,

are at odds with the learning goals or classroom activities. Ideally the educational content, goal and activities are in a partnership, aligned for mutual reinforcement (Hughes, 2003).

The starting point of this alignment is the need for clear and specific learning goals (Guskey & Bailey, 2001; Marzano, 2000). These learning goals have to be clear in order to effectively assess growth and direct learning or to measure and grade achievement. The logical next step is to move away from normative grades, because when content and process are differentiated it does not make sense to base grading on a year group norm (Guskey & Bailey, 2001; Wiggins, 1993). When students are presented with goals, translated in indicators of student success and then measured according to the described criteria, the grades will reflect the learning of individual students.

In the Dutch educational system, the ministerial goals for English language skills are not specific and are unavailable for the separate years in secondary education. Schools are effectively free to choose the way in which they help their students acquire the knowledge and language skills needed to pass the nation-wide written exams in the final year of secondary school. In this situation, according to Panadero, Tapia and Huertas (2012), rubrics might help to offer a roadmap of intermediate steps that build up towards a desired language skills level. This could help students focus on developing personal language skills, instead of on grades. Instead of grades, students should receive frequent and immediate feedback on their performance in such a way that they can use the feedback in their learning. Teachers could offer positive, neutral, and corrective statement to vary the feedback they give (Brosvic, Dihoff, Epstein, & Cook, 2006; Chard, Vaughn, & Tyler, 2002). This support of learning through assessments, or feedback based on formative assessments *for learning* is very effective. According to Popham (2011), recent reviews of research clearly show well-implemented formative assessment can essentially double the speed of student learning.

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment should avoid grading if possible. If awarding grades cannot be avoided, the grade should reflect the relationship between achievement, expected personal growth and curricular requirements. A rubric with clear intermediate steps from starting to use the English language towards the required language skills level, could help students to focus on their personal language skills development. In that sense, it would be best to create more formative assessment moments. Instead of summatively*

*rounding off a small portion of the language skills at hand, students might start to see the formative assessment moments as part of their personal language skills development.*

#### **4.3.3.7. Building blocks of DI: Feedback**

Feedback, in the sense of telling someone how well they did, is not always effectively used in the educational process. Giving negative feedback orally turned out to be ineffective in drills and repetitive exercises in non-communicative language teaching (Ellis, 1994; Lightbown, 1983; Lightbown, Spada & Wallace, 1980). Van den Branden (1997), however, shows that commenting on just correctness and appropriateness of utterances can be very effective, like DeKeyser (1993) who stated that receiving feedback on a specific problem was more effective than feedback on all mistakes made. Corrective feedback and focus on form also give positive results in second language teaching in primary and secondary education (Spada & Lightbown, 1993; White 1991; White, Spada, Lightbown & Ranta, 1991) and adult education (Williams & Evans, 1998). Lyster and Ranta (1997) found teachers often ask for clarification after an utterance, not because of a lack in comprehension, but as corrective feedback intended to trigger the student into 'noticing the gap'. Pica, Holliday, Lewis and Morgenthaler (1989), like Nobuyoshi and Ellis (1993), found this type of question often leads to improved output of the second language. In general terms, Long (1996) found positive effects of giving comprehensible input and negative feedback. When learning a language, students can come a long way by receiving input, but feedback is necessary when the second language learner is unable to recognise his own mistakes (e.g. Rutherford & Sharwood Smith, 1985, 1988; White, 1987). Some researchers see feedback as a form of (correct) comprehensible input (Mackey, 2002). Feedback is even seen as a means to enhance metalinguistic awareness (Brinton, Snow & Wesche, 1989). The need for focus on form related to content is claimed to be necessary in second language acquisition and this claim is built on a Vygotskian, social-cultural approach to education. (e.g. Lantolf & Pavlenko, 1995). In this sense feedback is also claimed to be necessary to get students to revise their personal output (Schmidt & Frota, 1986; Long, 1996). This feedback needs to come from a teacher, because

second language learners are no longer corrected by peers once they have receptively mastered a language and are reasonably able to convey a message (Williams, 1997; Garcia Mayo & Pica, 2000)

Feedback should preferably be used in situations with meaningful interaction (Lyster, 1994; Lyster, Lightbown & Spada 1990; Lochtmann, 2002; Pica & Washburn, 2003).

Apart from focus on form, recasts may also be used to teach attention to style, rhetoric and register when utterances are grammatically correct, but could be rephrased more precisely, concisely, or more beautifully (Mohan & Beckett 2001).

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment needs to contain structural moments for feedback on the work in progress. Feedback should be from peers to make it as meaningful as possible and from the teacher more specifically for focus on form. When students give each other feedback a rubric is necessary to help focus on the aims and a form to record the feedback and help students keep track of their process.*

### **Building blocks of DI: Grouping students**

Grouping students must be done flexibly if diverse students are grouped according to their differences in readiness levels, interests and learning profiles. This can have a positive influence on the learning environment, help students to become engaged students and support them as they construct new knowledge (Dion, Morgan, Fuchs, & Fuchs, 2004; Fuchs, Fuchs, Al Otaiba, Thompson, Yen, McMaster, Svenson & Yang 2001; Bond & Castagnera, 2006; in Rock, Gregg, Ellis, & Gable, 2008). Conversely, there are negative effects when ability grouping is not used flexibly and students remain in the same ability group (Deunk, et al., 2015). Or, as Op 't Eynde (2004) puts it: "Steady, structurally used grouping (e.g. ability grouping) is best avoided because of the negative side effects this has, for example, on the weaker students." (Own translation<sup>23</sup>, Op 't Eynde, 2004, p. 10).

---

<sup>23</sup> "Vaste, structureel ingebouwde groeperingsvormen (bijvoorbeeld:niveaugroepen) worden het best vermeden, gezien de negatieve 'side-effects' voor bijvoorbeeld de zwakkere leerlingen." (Op 't Eynde, 2004; 10)

Grouping does have an effect on the students with regard to their respective perceived ability levels and the status derived from the grouping itself. This side effect does not make it easier on teachers.

“Dealing with differences [in a group] is much more complicated than usually assumed. Because of the tendency to imitate and the pressure for uniformity [within the group] differentiation for individual students always needs to be accompanied by integration into the whole-group process. Furthermore, it seems that teachers need to have a certain authority to curb the aggressive aspects of internal rivalry.” (Own translation<sup>24</sup>, Bulterman-Bos, 2007, p. 12).

Apart from how groups are organised, research is also available on working in a group, or cooperative learning. Many studies underpin the positive effects of cooperative learning in general (Slavin, 1995; Pica, Lincoln-Porter, Paninos & Linnell, 1996) and list the following characteristics of effective cooperative learning: structured objectives, group objectives and awards, individual responsibility and an equal chance for success for each group member.

In language learning, working in pairs and groups improves correct production in reading and writing, but seems to be less suitable for achieving proper pronunciation and grammatically correct production (Swain, 2001). This is the reason Swain (2001) concludes that cooperative learning should be accompanied by corrective feedback to prevent mistakes becoming entrenched. Swain (2001) also states that receptive knowledge and skills will help students to pinpoint mistakes, regardless of whether they are able to come to correct production themselves, which in itself speaks for the use of cooperative learning.

*Given the above research and publications, the language teaching programme based on differentiated instruction to be used as treatment should use grouping flexibly. At least it needs to be ensured that students do not end up in the same ability group every time. Students should be grouped flexibly, based on ability, interest and preferably for cooperative learning activities.*

---

<sup>24</sup> “Omgaan met verschillen is dus veel gecompliceerder dan doorgaans wordt aangenomen. Door de imitatietendens en de uniformiteitsdruk moet differentiatie m.b.t. individuele gevallen gepaard gaan met integratie in het geheel van de groep. Het lijkt erop dat de leraar bovendien een bepaald gezag moet hebben om de agressieve aspecten van onderlinge rivaliteiten in toom te houden.” (Bulterman-Bos, 2007;12).

#### **4.3.4. Results (D): The role of Teachers in differentiated instruction**

Teachers have a leading role in the implementation, quality and continuation of differentiated instruction. This role can be subdivided into the knowledge and skills teachers need to differentiate instruction, i.e. teacher's professional development on the one hand, and the implementation process on the other hand.

##### **Professional Development**

In their review study Van Driel, Meirink, Van Veen and Zwart (2012) reported consensus on the core features of effective professional development programmes. The six core features they report upon are based on five features listed by Desimone (2009) which "...are critical to increasing teacher knowledge and skills and improving their practice, and which hold promise for increasing student achievement."

(Desimone, 2009; 183).

According to van Driel, et al. (2012), the six core features of a professional development (PD) programme focus on classroom practice and contemporary, research-based knowledge of teaching and learning (1). For them to learn within a programme, teachers need to be actively engaged in inquiring into the professional development they go through and in their daily practice (2). To increase the effectiveness and usefulness of a programme, the learning needs to be a collaboration between teachers of the same school, grade or subject with permanent access to expertise (3). A programme will be more effective when a substantial amount of time is invested in the sense of duration as well as actual hours (van Driel et al. (2012) report a range found in reviews between 14 and 80 hours).

Furthermore, the invested time in the initial PD programme should be sustained by continued support (4).

A PD programme needs to be in coherence with teachers' knowledge, beliefs and problems experienced daily and be consistent with reforms and policies at school and national levels (5). The final feature consists of the need for PD programmes to take 'school organisational conditions' into account (6). Although their study concerned a science teacher, the general advice is to take the daily school reality into account as it seems "...too important to be ignored if PD aims to make a lasting impact on school practice." (van Driel et al. (2012, p. 154). This reconfirms the findings of an earlier review study on effective characteristics of

teacher professionalization interventions, in which Van Veen, Zwart, Meirink & Verloop (2010) reported how lack of attention for the transfer and reciprocity between professional development and the consequences for daily school reality greatly impede implementation of intended reform (Van Veen, Zwart, Meirink & Verloop, 2010).

Fixsen, Naoom, Blasé, Friedman, and Wallace (2005) defined implementation as “a specified set of activities designed to put into practice an activity or program of known dimensions” (p. 5). The specified set of activities were defined as two distinct sets: (A) intervention-level activities, which are well-defined, purposeful, and measurable delivery of the intervention, and (B) implementation-level activities, which are actions required by the organisation or setting to ensure effective intervention delivery. According to Fixsen et al. (2005) implementation went through the following stages:

1. exploration and adoption, in which the need for an intervention is observed and researched,
2. program installation, in which everything is prepared for the implementation of the intervention,
3. initial implementation, in which first successes and barriers are used to further development,
4. end, in which the intervention is implemented with adequate treatment integrity and
5. innovation and sustainability, in which the intervention becomes business as usual (Fixsen et al., 2005).

Teachers can use this structure of the implementation process, but they also influence the quality of deliverance as they will be confronted with implementation problems along the way. When teachers are responsible for implementation of interventions, the extent to which interventions are implemented as planned, influences how much students benefit from the intervention. This concept is known as ‘treatment integrity’ (Long, Sanetti, Collier-Meek, Gallucci, Altschaeffl & Kratochwill, 2016). Research on how treatment integrity in schools is influenced or hampered by implementation barriers has only recently begun to grow (Long et al., 2016). Implementation barriers can be defined as: “Variables that obstruct efforts to implement an intervention, often reducing its impact” (Long et al., 2016).

In order to organise and advance the literature on implementation barriers, several systematic reviews on intervention research have formulated structures or frameworks (e.g. Feldstein & Glasgow, 2008). Sanetti

and Kratochwill (2009), in turn have used these reviews to bring the hypothesized implementation barriers down to four categories, consistent with the work of Feldstein and Glasgow (2008):

Intervention barriers, which have to do with how compatible and easy implementation is, like required additional time investment.

Organisation barriers, which have to do with in-school resources, climate, leadership and school administration.

Implementer barriers, which have to do with the teacher's skills, self-efficacy, belief in- and support of – the intervention.

External environment barriers, which largely have to do with coordination between external organisations or schools; the broader context in which an intervention is conducted. According to Sanetti and Kratochwill (2009), these categories contain 37 specific types of implementation barriers (see appendix 3.1.).

### **Teachers and differentiated instruction**

Research on this topic shows a mixed picture of success. Tomlinson (1998) showed how some teachers were initially opposed to differentiated instruction because of insecurity about the changes in the curriculum as well as insecurity about the different role the teacher would have in the classroom. One of the persistent reasons for opposition is teachers' uneasiness about the increased planning time (Robison, 2004). In a questionnaire survey of a representative stratified random sample of 1988 middle schools, Tomlinson, Moon and Callahan (1998) found that most teachers had problems related to the need to deal with the diversity of the students in their classroom while very few teachers would take student interests, learning profile or cultural differences into account when planning lessons. The ability of teachers to differentiate instruction was not influenced by their age. In fact, those teachers who started to implement differentiated instruction demonstrated that the teacher's attitude towards change was a more decisive factor for successful implementation than age (Tomlinson, Moon and Callahan, 1998). When the start of the implementation was successful teachers would be more likely to persist (Tomlinson, 1995). Along the same lines, Johnsen (2003) reported his undergraduate teachers had a success experience when they used differentiating instruction to suit different ability levels. Affholder (2003) concludes her case study by

posing the hypotheses that teachers who intensively use differentiated instruction show improved individualised perception of the individual learning needs of students, an enhanced sense of efficacy and responsibility for student growth. According to Affholder (2003), the use of higher levels of differentiated techniques seem to make teachers feel more effective and open to try new instructional approaches (Affholder, 2003).

Based on a large-scale longitudinal survey research to investigate the growth of, and link between, teachers' instructional behaviour and motivation components, Opdenakker and Van Damme (2006) found that it is important for teachers to differentiate instruction to the full extent to optimize classroom practice.

Hattie (2009), who fell outside of the original scope of this literature research because he did not report on differentiated instruction itself, concluded that not every teacher will make a difference. He said: "... teachers with high expectations for all students, and teachers who have created positive student–teacher relationships are more likely to have the above average effects on student achievement" (Hattie, 2009, p. 126).

*Given the above research and publications, the language teaching programme based on differentiated instruction should have the following six critical core features: The programme should:*

*focus on classroom practice and research-based knowledge of teaching and learning (1);*

*actively engage teachers into their professional development related to daily practice (2);*

*have teachers learn in collaboration with colleagues and experts (3),*

*invest a substantial amount of time (timespan and actual hours), sustained by coaching (4);*

*be in coherence with teachers' knowledge, beliefs and daily experienced problems as well as consistent with reforms and policies on school and national levels (5);*

*take 'school organisational conditions' into account (6).*

*When the programme contains these features, it can best be implemented through the following stages:*

- 1. exploration and adoption, in which the need for an intervention is observed and researched;*
- 2. program installation, in which everything is prepared for the implementation of the intervention;*
- 3. initial implementation, in which first successes and barriers are used to further development;*

4. *full operation, in which the intervention is implemented with adequate treatment integrity; and*
5. *innovation and sustainability, in which the intervention becomes business as usual.*

*The teachers implementing the programme should have a positive attitude towards the (educational) process and the students. Teachers with high expectations of all students, positive student–teacher relationships and a positive attitude towards change seem to be more effective. So the teacher and his, or her, attitude to students and education are a decisive factor in the implementation of differentiated instruction.*

*Finally, when teachers start implementing the programme, they will possibly run into implementation barriers. It is important to keep in mind these Intervention-, Organisation, Implementer, and External environment barriers, can hamper the implementation process.*

#### **4.3.5. Results (E): The effectiveness of differentiated instruction**

Differentiated instruction as described in section 2.3.3-2.3.4 above is a rather difficult to research. This is probably why research on differentiated instruction is limited (Tomlinson & Allan, 2000; Algozzine & Anderson, 2007; Hall, et al., 2002) and why differentiated instruction did not appear in Hattie's (2009) synthesis of over 800 meta-analyses relating to the influences on achievement in school-aged students. The research that *is* available shows that differentiated instruction can lead to improved test scores (e.g. Rock, Gregg, Ellis, & Gable, 2008; McQuarrie, McRae, & Stack-Cutler, 2008; McQuarrie & McRae, 2010), as well as to positive affective outcomes, such as raised levels of engagement, motivation, and more positive attitudes to school (e.g. Tieso, 2001; Johnsen, 2003; Beecher & Sweeny, 2008).

In a meta-analysis that summarizes studies that investigated the effects of teaching on student learning from the decade preceding 2007, Seidel and Shavelson (2007) found differentiation (with an effect size of .04) to be among the variables that yielded the highest effect sizes (Seidel & Shavelson, 2007).

Despite the limited availability of research on differentiated instruction, research is available on a number of the building blocks, theories and principles that make up differentiation, such as: effective classroom management procedures; promoting student engagement, motivation and readiness; flexible grouping and teaching to the zone of proximal development (Tomlinson & Allan, 2000; Ellis & Worthington, 1994).

**4.4. Conclusion of the literature study on differentiated instruction: its theoretical foundations (B), building blocks (C) The role of Teachers in differentiated instruction (D) and prerequisites for a language teaching programme based on the principles of Differentiated Instruction.**

The findings of the above chapters have been summarised below, as stepping stone towards a description of the required characteristics of a treatment in which language teaching is differentiated.

The starting point for an experimental treatment intended to help teachers differentiate their language teaching must be the acceptance that all students are essentially different and that there is a need to differentiate education to respond to these differences (4.3.2.2.). The logical consequence is that all students receive the learning experience they need to achieve the steepest learning curve possible (4.3.3.5.).

In order to offer the learning experience needed, teachers are required to work with their students' learning styles (4.3.2.2.4.) and how these affect student's motivation (4.3.3.4.), attitude to learning, and performance.

To a large extent, motivation is the responsibility of teachers since it promotes learning (4.3.3.4.). The main reason that motivates students to attend class, for instance, is the teacher's enthusiasm for the subject matter. Student performance is influenced by what students think their teachers think about them; on an emotional level, students will mirror what teachers bring into the classroom (4.3.3.3.). Any positive influence in this direction should be employed since feelings precede thinking and emotions highly influence decision-making and consequently learning. A last point to make is of course the safe learning environment required for students to venture into working in a foreign language (4.3.3.3. & 4.3.3.4.).

The students, on the other hand, need to become partners in the educational process (4.3.3.1.). The first step towards this partnership is for the educational process itself to be as transparent as possible. Starting with the learning goals which must be:

- crystal clear to the students,
- communicating high expectations of students by their teachers and
- seen as meaningful and personally relevant to the students.

Besides these learning goals, students need to understand they will be offered choices to influence the process and products (4.3.3.5.) towards their personal interests and/or learning style (4.3.3.2.) when and where possible. Choice in the process itself will be possible as long as students work on essential skills to come to understand essential ideas, while focused on a learning goal. Choice in the product will be possible as long as it clearly demonstrates newly won knowledge, ability and understanding.

Another important point related to the shared learning goals and standards is a clear description of what evidence of learning is measured (4.3.3.5.); how this is measured, assessed and graded. The (final) product at hand should afford students various ways of demonstrating their understanding and/or ability learned from the lesson or unit of study.

When assessment is used, the best approach is to use assessment 'as' instruction, where the assessment is used formatively to help students become aware of their ongoing progress (4.3.3.6.).

When grading is needed it is an end-point judgment about students' achievement, preferably relative to his or her personal growth and involvement. It does not make sense to offer normative grades and compare students to others in class when content and process are differentiated (4.3.3.6.).

Tasks offered need to be in response to student interest, at the proper level of difficulty, and always related to the 'bigger picture' of the higher order learning goals. Teachers should help students to come to understand the big ideas found wrapped in the curriculum standards and revisit those ideas in practice multiple times during the course of their education. Exploration of implicit connections and underlying principles of a topic can facilitate transfer of learning and support student performance (4.3.3.1.).

In language teaching, the tasks offered need to ensure students produce language often, because production leads to observation of personal deficiencies and 'noticing the information gap' is imperative to learn a language. Helping students to notice the gap is usually done through feedback.

Students should receive frequent and immediate feedback on their performance in such a way that they can use the feedback in their learning (4.3.3.7.). Corrective feedback and focus on form are very effective, especially when given on a specific problem in situations with meaningful interaction. Asking, for instance, for clarification after an utterance triggers the student into 'noticing the gap', which often leads to improved output of the second language.

The findings summarised in the first part of this conclusion lead to the following list of elements that need to be used in the development language teaching programme based on differentiated instruction:

1. students need to be offered choice in both the educational process as well as the product to help assess learning;
2. students should have the opportunity to offer contributions from their personal interests;
3. students need to become partner in the educational process by discussing and personalising goals;
4. teacher assessment needs to go together with student self-assessment, as this can help the teacher relate expected individual learning gains to individual achievements;
5. learning gains need to be relative to individual student capacity and commitment;
6. educational activities need to be self-regulated as much as possible;
7. educational activities need to be in cooperation with peers;
8. goal structures need to be individualistic and pertaining to mastery of goals;
9. feedback needs to be given on both success and failure (at learning);
10. the long-term consequences of achievement need to be clear;
11. students need a deeper form of processing by production of language through which they can observe their personal deficiency and notice the gap;
12. there needs to be a positive pedagogical relationship;
13. the relationship between the students' everyday life and the language learning activities needs to be built upon;
14. explicitly show the students how the language learning activities relate to the language skills and abilities students are working towards;
15. it is important to ask the teachers to stress the connections between the language learning activities and the abilities students work towards;
16. it is important to allow time for revision and discussions on progress with peers and the teacher;

17. teaching should work towards products related to the ability level expected of the individual student and related to student's previous achievements. In that way the appreciation of the products can be fully differentiated;
18. avoid grading if possible. If awarding grades cannot be avoided, the grade should reflect the relationship between achievement, expected personal growth and curricular requirements;
19. a rubric with clear intermediate steps from starting to use the English language towards the required language skills level, could help students to focus on their personal language skills development;
20. create more formative assessment moments;
21. students might start to see the formative assessment moments as part of their personal language skills development;
22. needs to contain structural moments for feedback on the work in progress;
23. feedback from peers to make it as meaningful as possible and feedback from the teacher;
24. a rubric is necessary to help focus on the aims;
25. a form to record the feedback and help students keep track of their process;
26. students should be grouped flexibly, based on ability, interest and preferably for cooperative learning activities;
27. teachers with a positive attitude towards the (educational) process and the students;
28. teachers with high expectations of all students;
29. positive student–teacher relationships;
30. a positive attitude towards change;
31. teachers who understand that when they start to use the programme, they will run into implementation barriers;
32. the programme should be implemented through 5 stages;
33. the programme should have six critical core features.

These 33 points are used to build the differentiated language teaching programme to be used as treatment in the experimental phase. The development of this treatment is described in chapter 5 below.

## **5. The development of a language teaching programme based on differentiated instruction.**

### **5.1. Introduction**

Although the differences between researchers and practitioners have been described as being so fundamental they might be seen as a gap (Tyler, 2009), the teachers in this research have been treated as important partners in the efforts to differentiate language education, in contrast to other educational reform (Van Veen, 2008). This partnership is based on the view on differentiating education as described in chapter 4.3.3.1. and the partnership between teachers and students in the educational process (Horwitz, 2000). When students are placed at the centre of their own learning process through continuous involvement in making decisions on process and products, it helps them progress and become more independent students (Tomlinson & Allan, 2000; McTighe & Brown, 2005; Lai, 2011).

One of the reasons for this research into a programme based on differentiated instruction for English language teaching was an appreciation for the practical situation of secondary school English teachers faced with an enormous variety in English language skills levels. In an experiment intended to offer a solution to that situation, it is crucial to involve these teachers as partners in the choices in process and products involved in the development of a language teaching programme based on differentiated instruction. Only when the research-based recommendations for a language teaching programme based on differentiated instruction are enhanced and adjusted through consideration of the practical limitations of everyday language teaching in the Dutch secondary school context, will the experience be transferable to other (English) language teaching classrooms.

### **5.2. The foundation of a language teaching programme based on differentiated instruction.**

The development of a language teaching programme based on differentiated instruction, to be used as treatment in the experiment as reported upon in chapter 6, is based on the results of chapter 4. The conclusion of chapter 4, the literature study on differentiated instruction, resulted in a list of elements that need to be used in the development. As can be seen in the list of elements for the treatment (Chapter 4.3 above), the results say something about the aims and content, teacher disposition and activities as well as student involvement. The list has been re-structured to show how the elements relate to each other and

function in concord. Below, the elements have been sorted according to what literature has to say about the aim of the programme and the programme itself, followed by what is expected of the teacher, to conclude with the role that needs to be given to the student.

As you can see below, the aim of the programme is to differentiate instruction in order to involve students as partner in the educational process (3) by:

- having students focus on their personal language skills development by using rubrics (19, 24);
- having students see formative assessments as part of their personal language skills development (21, 3);
- having students see goals theirs and part of their personal language skills development (8).

The programme itself should have the following six core features (33):

- focus on classroom practice and research-based knowledge of teaching and learning (33.1);
- actively engage teachers in their professional development in relation to daily practice (33.2);
- have teachers learn in collaboration with colleagues and experts (33.3);
- invest a substantial amount of time (timespan and actual hours) sustained by coaching (33.4);
- be in coherence with teachers' knowledge, beliefs and daily experienced problems and be consistent with reforms and policies on school and national levels (33.5);
- take 'school organisational conditions' into account (33.6).

The programme is implemented through 5 stages:

- exploration and adoption (32.1);
- programme installation (32.2);
- initial implementation (32.3);
- full operation (32.4) and
- innovation and sustainability (32.5).

The whole endeavour up to this point in the dissertation can actually be viewed as part of the first stage (exploration and adoption, 32.1).

The teachers are the ones who need to make this all happen, but they are not to be seen, or treated as mere vehicles for the implementation of the programme. Literature shows (cf. chapter 4) that in order to successfully plan, organise and implement the programme, teachers should be seen as partners (3) in the

process of differentiating education. In addition to their willingness to initiate that partnership, they need to:

- have a positive pedagogical relationship (12, 29);
- have high expectations of all students (28);
- have a positive attitude towards change (30).

Besides these personal qualities, teachers need to work to ensure students:

- see language learning goals as theirs and working towards the mastery language skills (8);
- understand explicitly how language learning activities relate to students' everyday life and the mastery language skills (13,14,15);
- understand how the assessment of products relates to the ability level expected of the individual student, student's previous achievements and educational standards (4,17).

Apart from these general and abstract conceptualisations, the literature also shows that teachers in practice need to organise:

- time for revision and discussions on progress with peers and the teacher (16);
- avoid grading if possible (18), or have the grade reflect the relationship between achievement, expected personal growth and curricular requirements (18);
- work towards self-regulated educational activities (6);
- create formative assessment moments (20);
- organise structured and formalised (peer) feedback on success as well as failure (9, 22, 25);
- group students flexibly based on ability and interest (26);
- organise cooperative learning activities (26).

The students, as the last group are by no means the mere recipient of the programme. In this programme students are expected to:

- have choice and influence in process & product (1);
- contribute to content, process and product from their personal interests (2);
- cooperate with peers (7);
- self-assess their personal educational process and achievement (4);
- produce language for a deeper form of processing and for them to 'notice the gap' (11);
- focus on their personal language skills development by using rubrics (19,24).

All of this, everything that is desired and expected from teachers and students, all the changes in daily practice, should lead to language teaching based on differentiated instruction. Language teaching that is responsive to student differences, with students as partners in the educational process, so that the language teaching is effective for all students.

The tactic of placing numbers next to the elements of the programme to indicate the origin from the list of 33, as used above in 5.2., will be continued to clearly mark the relationship between theory and practice.

### **5.3 The development of a language teaching programme based on differentiated instruction**

In order to effectively have teachers agree to, develop and implement a language teaching programme based on differentiated instruction, the next stage (programme installation, 32) was initiated by discussing the list of elements for the treatment (5.2 above) with the secondary school English teachers involved (3, 33). The rubrics (appendix 3.2.), as tangible and useful materials, were developed in close cooperation (19, 24) during a meeting in July 2013.

The English language teachers of all 12 schools participating in the discussions preliminary to the experimental phase were enthusiastic about the possibilities that a language teaching programme based on differentiated instruction might offer (33.5). All teachers involved indicated they were experiencing discontent with the current way in which they offered English. They were aware of the enormous variety in language skills levels within one classroom and of the discrepancy between individual English language learning needs of students and the in-class language learning offered. This meant they were open to discuss all the possible interventions and changes to teaching practice as proposed by literature and previous studies. Throughout the academic year preceding the experimental phase (2013-2014), the teachers were consulted on the possibilities differentiation could offer (3).

With the English teachers and their respective schools as partner, the discussions were on how far they were willing to take the idea of differentiated instruction (33.2). A school-wide differentiation of the curriculum was out of the question and also beyond the scope of this research. Although the English department of the secondary schools involved were committed to contributing to this research, the other subject teachers as well as school leaders thought a school-wide differentiation of the curriculum was too

radical an educational reform to base solely on the proposed experiment. The same applied to the English teaching curriculum. As not all teachers within the departments of the respective schools welcomed the attempts to differentiate language teaching, schools decided to have the experiment take place in the first year only (33.6).

### **Contextual limitations**

This limitation meant that the language teaching programme based on differentiated instruction to be used in the experimental phase of this research was tied to a number of organisational prerequisites. Because the treatment was to be used in only one class of one year group, all agreed (33.5) that, to allow differentiated instruction to become part of the language teaching, the treatment needed to be:

- a) embedded in the daily practice,
- b) in concord with the themes and assignments offered in the course- and text-books,
- c) limited in its duration relative to the subsequent teaching activities.

In order to meet these prerequisites (33.6) while adhering to the recommendations found in literature, as described in chapter 4.4 above, the language teaching programme based on differentiated instruction generally speaking:

- focusses on the productive skills, writing and speaking (11, 33.1);
- offers (writing and speaking) assignments at the beginning of a course book chapter and students are supposed to work on these assignments the duration of the course book chapter (chapter-wide assignments) (16);
- has teachers relate the quality of the (speaking and writing) products to a rubric during instruction, feedback and assessment (19, 21, 24, 3, 8);
- has teachers plan weekly feedback moments during cooperative learning activities (20, 21, 22, 25, 26, 3, 4, 7, 8, 9, 16, 17);
- has teachers use feedback forms to help students keep track of their learning (9, 22, 25);
- has teachers offer choice in the process or product of learning (1, 2);

- has teachers relate other language teaching activities to the writing and speaking assignments of the language teaching programme based on differentiated instruction (13, 14, 15).

#### **5.4 The language teaching programme based on differentiated instruction**

The above qualities have been combined into a differentiated approach to writing and speaking. This approach has been described below in a procedure for English language teaching, to use during the academic year of the experiment.

##### **The language teaching programme based on differentiated instruction's pedagogical conditions**

The starting point for the programme consisted of 4 pedagogical conditions, that are exactly the same for both speaking and writing. To involve students as partners in the educational process (3) the following conditions have to be met. **The first condition is** the consistent use of the writing/speaking skills rubric (19, 24). The rubric must be an accessible document for the students. When this condition has been met, students, as well as teachers, will become aware of each student's level of competency in writing/speaking. The insight gained enables the teacher to individually challenge students to perform above their current level of writing/speaking skills (3, 4, 8, 17, 21). **The second condition is** the consistent use of feedback (9, 22, 25). During practice moments, students give each other compliments, tips and suggestions based on the rubric (7, 16). The teacher does the same for individual students. During the experiment this is structured through the use of a form. This form helps students collect written feedback where every writing/speaking product is to be used as a reference and documented survey of the student's progression. When this condition has been met the student gains tangible proof and insight into his progression in writing/speaking skills (4, 8, 17). **The third condition is** the possibility to influence the process or product (1, 2). Students must be allowed to adapt the form or content of the writing/speaking product based on personal preference or interest. When this condition has been met the writing/speaking assignments will become more meaningful to the students and their writing/speaking products will become more personal expressions (8, 17). **The fourth condition is** the explicit coherence of educational activities. Teachers must explicitly relate the general English language teaching activities such as listening, speaking, reading, vocab work, grammar, exercises and explanations, to the writing/speaking product that the students are working

on (13, 14, 15). When this condition has been met students will perceive the whole educational process as coherent and find added value in activities since they relate to the writing/speaking process (4, 8, 17).

### **The language teaching programme based on differentiated instruction's rubrics**

After the formulation of the pedagogical conditions, the rubrics (19, 24) must be created. The rubrics required for the treatment have been based on the ECML rubric for writing skills of the AYLLIT project (Hasselgreen, 2012) and the speaking skills rubric of the 2009 handbook relating the CEFR<sup>25</sup> with exams and assessments. The CEFR was originally designed for adults and has six levels, A1, A2, B1, B2, C1 and C2, which represent a lifelong span. The AYLLIT project team attempts to provide teachers of young language students (9-13 years) with the necessary tools for assessing students' writing, using criteria linked to the 'portion' of the CEFR relevant to this age group (A1 to B1). The AYLLIT project team adapted levels to the age appropriate language-use situations and to reflect progress; in-between stages have been defined. As can be seen in the appendix (3A), the grid used by the AYLLIT project team introduced an "approaching A1" descriptor, to describe the earliest efforts of some students who would score below A1. The team also introduced intermediate levels 'A1/A2', 'A2/B1' and 'above B1', to reflect progress. The rubrics, as developed in cooperation with the English teachers involved for the treatment of this study, reflect the approach by the AYLLIT project team. As can be seen in the appendix (3B), the rubrics used in the treatment also used an "approaching A1" preliminary level, as well as intermediate levels to reflect progress. The rubric of this study is based on the ECML rubric for writing skills of the AYLLIT project because the methodology offered in the treatment should reflect the 'can do' attitude of the CEFR. The CEFR was based on the communicative approach to fulfil the communicative needs of students, as well as to elaborate and promote the concept of autonomy in foreign language learning. These concepts form the foundation of the treatment discussed in this chapter in this study as well.

Research shows that, at least in Europe, secondary education still struggles with the 'can do' attitude of the CEFR on one hand and standard, accepted, grammar syllabus progression on the other hand (Kedde,

---

<sup>25</sup> For additional information see 'The Common Europe Framework of Reference' in chapter 2.3 above.

2004). The same goes for publishers who label their textbooks and tests according to the CEFR levels, while the content, topics and organisation by language function do not differ from what they offered ten years ago (Figueras, 2012).

The rubric should offer students, as well as teachers, an instrument to assess language skills progress from a communicative point of view. An important point of view, because it tells students what they 'can do', instead of how many mistakes were made (Burmeister & Piske, 2008).

### **The language teaching programme based on differentiated instruction's practical procedure**

The procedures for the two language skills have been described as one procedure, as they share a number of communalities and expectations. Differences in the procedure will be highlighted in bold.

In the language teaching programme based on differentiated instruction, the rubrics are used by students for self-assessment and measure of progress, instead of being only a tool for teachers.

Both rubrics, for writing and speaking, have been translated in Dutch and adapted to the language used by the target group of first year secondary school students, in cooperation with the teachers involved in the experimental phase (33.2, 33.3). Teachers are invited to take ownership of the rubrics and expand or abbreviate at their professional discretion. These and further adaptations by teachers are welcomed as long as the changes aid teachers and students in giving feedback on products and performance.

### **The language teaching programme based on differentiated instruction practical procedure: start of a new chapter.**

With each introduction of a new chapter from the course book in use, teachers give a writing/speaking assignment to be completed at the end of the chapter (33.5). This writing/speaking assignment is based on the content of the course book chapter and may originate from the course book itself, or be given as additional assignment (33.6).

The instruction of the writing/speaking assignment is offered together with the rubric for writing/speaking skills (19, 24). Students indicate at what level of the rubric they aim to produce their writing product (4). Teachers indicate expected minimal achievement on the rubric (4, 17) and plan (at least two) '*practice moment(s)*' in the intervening months for students to work in class with peers on both the writing and the speaking products (20, 21, 22, 25, 26, 3, 4, 7, 8, 9, 16, 17).

### **The language teaching programme based on differentiated instruction's practical procedure: working on growth**

The very first time the above instructions are given, students should be informed of the difference between the writing and speaking assignments.

**The writing products** assigned are cyclic in nature, as each new product is added to previous work during the academic year and earlier work can be revised based on newly gained knowledge, skills and understanding. Because of the option to physically go back and rewrite, the writing products are expected to a **have cyclic growth**. At the end of each chapter the writing product is assessed and possibly even marked but in later chapters earlier writing products may be revised and improved. It must also be clear to students that the teacher, as well as their peers, will read their writing and give feedback on the writing during the planned '*practice moment(s)*'.

**The speaking products** assigned are accumulative in nature, as each new product is expected to be better than previous work produced during the academic year. The very first time the above instructions are given, students should be informed that the speaking assignments will recur at each course book chapter. Just as for the writing assignments, they will receive feedback on their (half-finished) performance, on a form, during the planned '*practice moment(s)*' and performance. Students are required to collect and keep these feedback-form(s) to use again as starting point for following preparations and performances. New performances or achievements will be related to previous performance. As said above, every student is expected **to show accumulative growth** (improvement) in each subsequent performance.

#### **Practical procedure: Time investment**

During the experiment teachers are expected to spend a minimum of 20 minutes per week on writing activities and 20 minutes per week on speaking activities. These activities could, for example, be the '*practice moment(s)*' in which feedback is given on the writing/speaking (half-finished) products/performances that students are working on, or other activities directly linked to the writing product or process, as described in the fourth pedagogical condition above.

### **The language teaching programme based on differentiated instruction's practical procedure: feedback and feedback-form**

Between the initial instruction and submitting the writing/speaking product for assessment, students should have at least one moment in which they receive feedback on their (half-finished) writing/speaking product.

It is important to limit the feedback to only one or two parts of the rubric. When students receive feedback from their peers or teacher, they write the feedback on the feedback form. The form itself helps limit the amount of feedback by allowing a maximum of three feedback points, one of which must be a compliment. Discussing the product with their peers and/or teacher provides students with written recommendations to improve their writing/speaking. Through the feedback forms students document their ongoing progression, related to the rubrics as described appendix 3. These consistent references to the rubrics in the points mentioned in the feedback helps students review their writing/speaking on specific points instead of on revisions based on corrections of all mistakes found. The collection of recorded feedback on language production functions as a work of reference on their personal language skills advancement. It shows all the compliments and other feedback they have received on products they have been working on throughout the academic year.

The feedback forms used had the following fields:

**Writing**

Name:

Assignment:

Date:

I received compliments and points to work on from:

Compliment:

Work on:

Work on:

1. In my writing I changed the following things...
2. I have learned next time to immediately...
3. The most important thing I learned is...
4. In my writing I changed the following things...
5. I did not change this because...

**Speaking**

Name:

Assignment:

Date:

I received compliments and points to work on from:

Compliment:

Work on:

Work on:

1. In my speaking I will change the following things...
2. I have learned next time to immediately...
3. The most important thing I learned is...
4. In my speaking I will not change the following things...
5. I will not change this because...

**The language teaching programme based on differentiated instruction's practical procedure: assessment**

The writing and speaking skills rubrics are used as an assessment instrument when the assignments are rounded off at the end of each course book chapter. Concordance with grading or marking in figures is left to the teachers' discretion. It is, however, important that, at the moment of instruction, students hear how

teachers will mark or grade relative to the levels in the rubric reached, or progression through levels of the rubric.

In the assessment of a final writing or speaking product at the end of the academic year, the teacher also uses the rubric. In addition, students can also be offered an assessment of their progression in writing and/or speaking skills in the first year of secondary school.

### **5.5 Conclusion: The implementation of the language teaching programme based on differentiated instruction as experimental treatment**

The language teaching programme based on differentiated instruction, as described above, is the product of a match between research-based suggestions (the 33 elements listed in 4.3 and restructured in 5.2 above) for differentiated education on the one hand and the practical opportunities and limitations of the Dutch secondary schools involved in this research on the other (described in 5.3 as Contextual limitations). As shown above in 5.2, the programme and its development is in compliance with both the six core features of a PD programme (Van Driel et al., 2012) as well as the list of required 33 elements, as summarised in the conclusion of chapter 4.

The matchmaking involved has not led to a full-scale, whole school, curriculum differentiation of education in the Netherlands. It has, conversely, led to a programme that seems practicable, with enough solid footing in the daily practice of the teachers involved for them to feel the proposed programme lies within their competence.

The cooperative nature of the development of the language teaching programme based on differentiated instruction as described above does not naturally lead to consistent implementation as intended, or in other words, treatment integrity. In order to sustain congeniality and offer scaffolding to the teachers of the experimental group, three trainings sessions were organised in July and August 2013. The information offered and discussed during these sessions was recorded and published in a blog <sup>26</sup>that could be used as

---

<sup>26</sup> <http://meijnen.com/myclassroom/>

forum to exchange information, questions and suggestions on for the teachers to form a professional learning community.

In all, the discussions, adherence to the practical limitations, training sessions, and planned visits during the pre-experimental phase, offered as much of a foundation as possible for a language teaching programme based on differentiated instruction. All the described activities above are in compliance with the first two stages of implementation according to Fixsen et al. (2005): exploration and adoption (in which the need for an intervention has been observed and researched) and programme installation (in which everything has prepared for the implementation of the intervention). The next chapter (6) deals with the next stage: initial implementation.

## **6. Quasi experiment to assess the effect of a language teaching programme based on differentiated instruction on the attitude towards learning English and the English language skills development of students in the first year of secondary education.**

### **6.1. Introduction**

The historical developments in English language teaching in The Netherlands, as described in chapter 2.4, clarify the differences in approach to English language teaching offered in primary and secondary education in The Netherlands. English teaching in Dutch secondary schools has, generally speaking, been moving towards a communicative approach and the teaching of grammar is, despite a century of reforms, new methods and approaches, still a fundamental part of teaching English as a foreign language. English teaching in Dutch primary schools, on the other hand, was introduced 30 years ago with an explicit choice for the communicative approach. Speaking and listening skills applied in communication through standard dialogues generally make up the curriculum.

These historically explicable differences in current English teaching approaches also show themselves in the results of the questionnaire survey described in **chapter 3**. The differences found reflect the difference in approach of English teaching between primary and secondary education as described in the introduction and as found in the literature review. Singing songs and playing language games are almost exclusively used in primary education, while teaching grammar, reading and testing play a dominant role in secondary education. The explicit importance given to grammar by the Dutch secondary school teachers of English in the sample, is reflected in the time they spend on the topic as well as in their choice of activities.

The differences described above in the approach and choice of language teaching activities in primary and secondary education are not helpful when it comes to the problem of the growing variety in language skills levels of students entering secondary education. As described in **chapter 2**, due to the rise of VVTO schools, students that fill the first year classrooms of secondary schools have a growing range of language skills levels and the one-size-fits-all approach to English language teaching is no longer feasible. The language teaching programme based on differentiated instruction, in this research used as treatment and described in **chapter 5**, might offer a solution to this problem.

The quasi-experiment as described in this chapter was set up to answer sub-question **4) To what extent does a language teaching programme based on differentiated instruction effectively increase the attitude towards learning English and increase the learning gains of first year students in secondary education?** To ascertain the effect of the treatment, the relative progress in English language skills of students has been measured with the help of the following language aspects as relevant variables: listening/spelling, reading, vocabulary and writing.

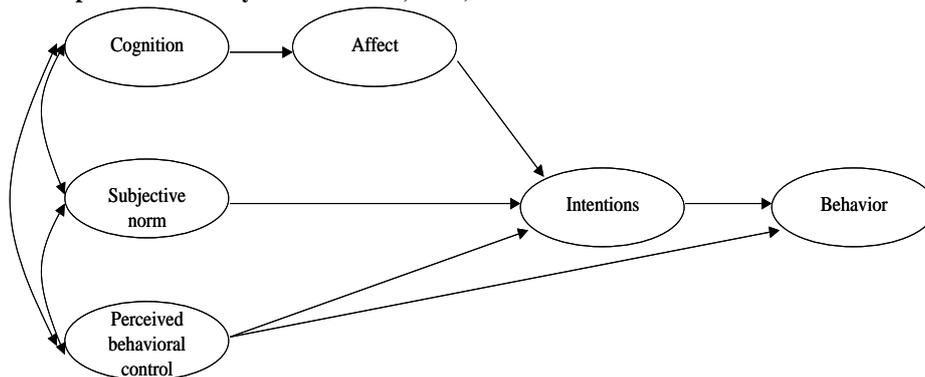
Besides affecting the language skills, it is interesting to see whether the treatment positively influences the student's attitude towards English language learning. The influence of the treatment on student attitude towards English language learning is important for two reasons:

- a) a positive attitude towards learning is an aim of education in itself and
- b) students would progress more when their (English language learning) behaviour has been positively influenced by their attitude.

The causal chain between attitude and behaviour is theoretically underpinned in the Model of Planned Behaviour (MPB). The MPB (Ajzen, 1985; 1988; 1991) is a widely accepted model that has repeatedly been validated in empirical research and is frequently used in social research (e.g. Van der Pligt & De Vries, 1998; Van Schooten, 2005; Fretschner, 2014). In the MPB, attitudes are defined as positive or negative feelings one has about performing the behaviour in question and this 'Affect' is seen as a consequence of the ideas an individual has about the expected consequences of performing the behaviour, also called 'Cognitions', and whether these consequences are perceived as positive or negative. According to the MBP, Affect in turn influences the 'behavioural intentions', which are seen as the intentions to perform the behaviour in the near future, and these behavioural intentions are hypothesized to influence the actual performing of the behaviour. So, the causal chain in the MPB contains a construct consisting of 'behavioural Intentions' and the model hypothesizes causal relations between Cognitions and Affect, between Affect and Intentions and between Intentions and Behaviour. Next to Affect, the MPB postulates two other determinants of behavioural Intentions: the 'Subjective Norm' and the 'Perceived Behavioural Control'. The Subjective Norm is defined as the perceived social pressure to perform or not perform the behaviour combined with the motivation to comply with these norms. The Perceived Behavioural Control (PBC) is defined as the

available options to perform the behaviour under consideration (perceived resources and opportunities), combined with the perceived facility with which a person thinks he or she can realise these options. The PBC is also thought to have a direct influence on behaviour, separate from Intention. A lack of opportunities can hinder the actual performance of the behaviour, even though the intention to perform the behaviour is, or remains, strong. The MPB and its postulated causal structure is shown in Figure 1:

**Figure 6.1: MPB as operationalized by Van Schooten, 2005)**



The possible effect of the treatment on student's language proficiency might be caused by the changes in teaching offered that create a richer learning experience. Possible (language) learning gains could also be due to a change in study behaviour of the students. This change in study behaviour might be caused by changes in different components of the MPB, which in turn might lead to more intensive study, which in turn yields larger language learning gains.

The MPB allows us to ask questions about how high students score on the different constructs of the MPB and which components of the MPB show the strongest relationship with the intention to study English as a school subject. Insight into how high students score on these constructs, gives an indication of how positive or negative students estimate the usefulness of learning English (Cognition) to be, how much they like it (Affect), whether they think they should make an effort to learn English (Subjective norm), the amount of behavioural control they perceive they have over learning English (PBC) and in to what extent they intend to make an effort to learn English in the near future (Intention). Insight into whether these intentions are mainly driven by, for instance, affect or perceived behavioural control (the relationships between constructs of the MPB) can help to plan specific measures to stimulate learning behaviour. In the quasi-experimental design used in this study, it is important to verify whether or not the attitude scores between

the experimental and the control group of students differ significantly at the pre-test. An insignificant difference in attitude is an indication that the attitudes of the experimental and control group of student are comparable at the onset of the study. A significant difference in attitude at the pre-test between the experimental and the control group could, conversely, make it more difficult to attribute changes in attitude to the effect of the offered treatment. This is because it is uncertain whether using the pre-test as a covariate will sufficiently adjust for pre-test differences between both groups.

Changes in MPB constructs might precede changes in English language skills. Effects on (parts of) the MPB might be found as result of the treatment even though effects on student's language skills are not (yet) found. The model presupposes that the determinants of behavioural intentions must change before the behavioural intentions change and also the behavioural Intentions are supposed to change before the actual behaviour changes. If the assumption of a causal chain of the MPB is correct, higher scores for MPB constructs will eventually lead to greater language proficiency.

For a comprehensive answer to research question **4) *To what extent does a language teaching programme based on differentiated instruction effectively increase the attitude towards learning English and increase the language learning gains of first year students in secondary education?*** the following sub-questions are added:

- a) *Has the measurement of the MPB been reliable and valid?***
- b) *Do pre-test scores of experimental and control students differ at the start of the experiment?***
- c) *Which constructs of the MPB best predict the intention to study the English language?***
- d) *If the MPB has been reliably and validly measured, what are the effects of the language teaching programme based on differentiated instruction on changes in the components of the MPB: Cognition, Affect, Subjective Norm, Perceived Behavioural Control and Intention?***
- e) *Has the measurement of the language skills been reliable?***
- f) *If the measurement of the language skills has been reliable, what are the effects of the language teaching programme based on differentiated instruction on changes in student's English I) vocabulary, II) spelling skills, III) reading skills and IV) writing skills?***

## **6.2 Method used to assess the effect of the language teaching programme based on differentiated instruction on the development of different components of the MPB and the English language learning gains of first year students in secondary education in The Netherlands**

### **6.2.1 Design**

In this study a quasi-experimental design, including a pre-test and a post-test with an experimental and a control group, has been used. The treatment lasted a whole academic year, from September 2013 to July 2014. The independent variables in this study are the variable indicating whether a student was a member of the experimental or the control group, the pre-test scores on the MPB-constructs and language skills and the background variables age, the ethnic background of parents and child, the language the parents speak with their child and the amount of time spent watching television, playing computer games and reading English books for leisure. The dependent variables measured in the post-test are the different constructs of the MPB and the language skills in vocabulary, spelling, and reading and writing in the post-test. The pre-test scores have been used as first covariate to operationalise the gains between pre-test and post-test (Ancova approach).

### **6.2.2 Instruments used to assess the effect of the language teaching programme based on differentiated instruction**

For the data collection an attitude test and a language skill level test was prepared, of which the first version, offered in September 2013, included a small set of questions to obtain background information. The attitude test consisted of 26 statements which operationalized each of the different constructs of the MPB, accompanied by 5-point Likert-scales (1 = I completely agree, 2 = I partially agree, 3 = I do not agree, nor disagree, 4 = I partially disagree and 5 = I completely disagree). The statements formulated and used to measure the MPB constructs in this research were based on the definition of the MPB constructs as given by Ajzen and Fishbein (1980). The complete test can be found in the appendix (4); for 'Cognition' six

statements were used (e.g. 'Learning English is important for everyday life.'<sup>27</sup>), for 'Affect' five statements were used (e.g. 'Learning English is fun.'<sup>28</sup>), for 'Subjective Norm' five statements were used (e.g. 'Learning English is part of a good education'<sup>29</sup>), for Perceived Behavioural Control six statements were used (e.g. 'I have a quiet place to study English.'<sup>30</sup>) and for 'Intention' five statements were used (e.g. 'My next homework for English I will do as soon as possible'<sup>31</sup>).

The language skills tested were: vocabulary, listening/spelling, reading and writing.

Vocabulary was tested (appendix 5) through a multiple choice version of items from The Peabody Picture Vocabulary Test—III (Dunn & Dunn, 1997). The Peabody Picture Vocabulary Test is used to test receptive vocabulary through individual interaction with an examiner. The examiner presents an item of four pictures to the testee, speaks a word which describes one of the pictures and asks the testee to point to the correct picture or indicate the corresponding number. The items in the Peabody are meant to be used with an age range tested from 2.3 to 90 years old in 17 sets of 12 items, in which each item consists of four pictures, as previously mentioned.

Although the original Peabody test provides a quick estimate of the testee's verbal ability, the time needed to administer it (15-20minutes) per student would exceed the scope of this research. Instead of testing all students individually, the students were offered a multiple choice version in which the English word was given after the Dutch instruction: 'Put a cross in the square that belongs to the English word...', as shown in the example below.

---

<sup>27</sup> Engels leren is belangrijk voor het gewone dagelijks leven.

<sup>28</sup> Engels leren is leuk.

<sup>29</sup> Het leren van Engels hoort bij een goede opleiding.

<sup>30</sup> Ik heb een rustige plek om mijn Engels te leren.

<sup>31</sup> Mijn volgende huiswerk voor Engels ga ik zo snel mogelijk doen.

0. Kruis het vakje aan wat hoort bij het Engelse woord:		ball	
			
		1	X 2
			
		3	4

To allow for differentiation of all possible language skills levels, the items offered in the multiple choice test consisted of items of all levels (1-17) as defined in the Peabody III. Additional level 4 items were offered as these were of the approximate language skill level of students in the first year of Dutch secondary education. The lowest levels leading up to level 4 were left in the test to allow for lower language skills levels and to help prevent test anxiety (Miller & Stoeckel, 2015), by offering initial test items that are easy to answer.

Listening/spelling was tested (appendix 6) through a dictation test with 10 sentences constructed from the 5 most often used English teaching course books in primary school English language teaching<sup>32</sup>. The sentences increase in complexity and spelling difficulty (e.g. (1) *'Hello, my name is Peter'*, (10) *'My teacher says I am good at Geography because I read a lot'*). During the test all students had to listen to a recording of the dictation to ensure consistency in exposure. Each sentence was read out in full, then repeated in 3 parts and finally read out in full once more.

Reading was tested (appendix 7) through an editing test. This editing test consisted of a text in which testees had to find and underline superfluous words. In this research the test is referred to as 'editing test', as it resembles the test procedure described by Bowen (1978) who coined the phrase. The procedure has been given different names, like: 'cloze-edit procedure' by Valette (1967); 'timed test of reading speed' by Davies (1975); 'Passage Correction test' by Odlin (1986) and 'Cloze-elide test' by Manning (1987), but the

---

<sup>32</sup> Top 5 best sold English teaching coursebooks for primary education in The Netherlands (Titel and publisher): Take it easy (ThiemeMeulenhoff), Real English (ThiemeMeulenhoff, derde editie), The Team (Noordhoff, tweede editie), Hello World (Malmberg, tweede editie), Backpack gold (Pearson Education) (Toorenborg & Bodde-Alderlieste, 2003).

phrase 'editing test' has been retained in this research since Bowen (1978) performed the first empirical study of the test procedure's reliability (Sattarpour & Ajideh, 2014).

To prevent recognition of the text and additional words in the repeated measurement after the experiment, two versions of the editing test were produced. The A version consists of five self-constructed paragraphs based on English language teaching themes from the primary and secondary school course books in use in the Netherlands. The B version is a one-page summary of a graded Penguin reader 'The Fireboy' (Rabley, 2008). The readability of both texts was calculated by the Flesch Readability Formula<sup>33</sup>. The Fire Boy (393 words) and the paragraphed text (388 words) had a Flesch Reading Ease score of 90.1 and 93.8 respectively, meaning both text were deemed: 'very easy to read'. In both texts, words were added randomly and students were asked to underline all words not belonging in the text<sup>34</sup>.

The placement of the words in the texts was organised through an online random integer generator<sup>35</sup>. The 75 random integers generated, with a value between 5 and 15, both inclusive, were used as the word count between intervals to ensure random placement. The additional words placed there were selected from the list of 1200 most common words (Sitton, 1995). As the first 100 words were mainly function words, the second 100 were used (100 – 200) from this list. Again the online random integer generator was used to generate 50 random integers, this time with a value between 100 and 200, both inclusive. These numbers were then applied to the 100-word list mentioned above to ensure random selection of words to add to the texts for editing. Words were not used if the added word produced a new syntactically and/or semantically valid sentence. This procedure produced sentences like: "He lives in a put big house with his brothers and sisters, in which the added word in this case has been underlined to indicate its position. Writing was tested through a writing assessment (appendix 9) in which students were asked to review a movie they had seen in the past six months by describing in English: the title and type of movie, the story of

---

<sup>33</sup> The specific mathematical formula is  $RE=206.835 - (1.015 \times ASL) - (84.6 \times ASW)$ , in which RE means Reading Ease, ASL means Average Sentence Length and ASW means Average number of Syllables per Word.

<sup>34</sup> The Dutch instruction said: "In de tekst staan woorden die er niet thuishoren en de opdracht is deze te onderstrepen."

<sup>35</sup> <http://www.random.org/integers/?num=50&min=7&max=12&col=5&base=10&format=html&rnd=new>

the movie, the people (or animals) that were in it, the best, funniest, most exciting, scariest moment and anything else they wanted to share.

The assignment allowed repeated use as the interval between pre- and post-test (9 months) exceeded the timeslot used in the requested description (a movie seen in the past 3 months).

Background information of all students was collected through 16 questions posed on the first page of the language skills test package (vocabulary, listening/spelling, reading and writing). The background information collected consisted of personal details that might influence a student's English language skills levels (appendix 10). Questions 1, 9 and 10 (name, current school and current class, respectively) were used to create unique numbers for each participant to enable multi-level analysis and the merging of data files. The remaining questions served as possible variables on the language skills levels and development in the academic year, divided into: (A) personal details (name, date of birth and gender); (B) possible influence of mother tongue and language(s) spoken at home (country of birth, mother's/father's country of birth, language usually spoken with mother/father); (C) previous education and English offered in primary school, (original primary school, Cito - final-test score<sup>36</sup>, number of years primary school English was offered); (D) out of school exposure to English (time spent on watching English spoken TV, playing English computer games, frequency of reading English books).

Unfortunately, the SES information on the student population in the sample was unavailable. Secondary education in the Netherlands does not use the student 'weight' SES indicator customary in primary education. Instead of using SES information, the possible influence of mother tongue and language(s) spoken at home was acquired through the questions on personal background as explained above.

---

<sup>36</sup> This standardized Dutch achievement test, called CITO test after the Dutch acronym for the National Institute for Educational Measurement of the Netherlands, is used in the final grade of 93% of Dutch primary schools and is part of the primary school's advice toward the choice of secondary education (Polderman, Bartels, Verhulst, Huizink, Van Beijsterveldt & Boomsma, 2010).

### **6.2.3 Instrument distribution used in the quasi experiment to assess the effect of the language teaching programme based on differentiated instruction**

The attitude test and the language skills level test have been administered twice. The attitude questionnaire was offered halfway through and at the end of the academic year, in week 8, February 2014 and week 27, July 2014. The language skills level test was offered to all students at the beginning (pre-test) and end of the academic year (post-test). The pre-test in week 35, September 2013 and the post-test in week 27, July 2014. The two versions that originated from the use of an A and B version of the reading test were distributed randomly over all students in both the experimental and the control group in the pre-test. Only afterwards were students linked to an A or B version to ensure students received the appropriate alternative version in the post-test (all pre-test A-students received a B version in the post-test and vice-versa).

During the study, which took 40 weeks (week 35, 2013 to week 27, 2014), the classes in the experimental group were offered the language teaching programme based on differentiated instruction as described in chapter 4.4. In all three secondary schools involved in this research, all English teachers confirmed the use of course books, classroom activities and tests to be largely the same in the parallel first year control groups. Prior to the treatment offered, none of the teachers used rubrics, chapter-wide assignments, cooperative learning activities or peer-feedback in their regular teaching, which were part of the experimental treatment. The teachers in the control group affirmed, during the experiment and preliminary contacts for the Delphi study as reported in **chapter 7**, that their teaching during the experiment concerned 'business as usual' and that none of the experiment-specific methodology was used.

#### **6.2.4 Sample of schools and teachers used in the quasi experiment to assess the effect of the language teaching programme based on differentiated instruction**

All students in the sample of students for the experiment were first year students in secondary education<sup>37</sup>.

The study was conducted with 127 students studying in first year groups of three secondary schools. The three schools are located in cities the Netherlands: in Delft, Rotterdam and The Hague respectively.

All three schools are public secondary schools that offer general pre-vocational education, higher general education, and pre-university education respectively. The sizes of the schools are around 600, 500 and 1500 students respectively. Apart from the size of the schools, the only difference was the aforementioned educational stream.

The students in the sample of the experiment were in two year-one groups of the pre-vocational education stream, two year-one groups of a mixed pre-vocational/higher general education stream and two year-one groups of a pre-university stream respectively.

There were 68 students in the experimental group, 37 boys and 31 girls, and 59 students in the control group, 29 boys and 30 girls. In the first week of September 2013, the mean age was 12.5 (sd = .5). In the experimental group all but 2 students were born in The Netherlands, against all but 5 students in the control group. The picture is different for their parents, although for some students these numbers are in a one parent family situation.

Half, 34 of the 68, of the mothers and fathers of the students in the experimental group were born outside The Netherlands. The control group is similar, with 28 of the 59 mothers and 36 of the 59 fathers having been born outside the Netherlands. This means students may be familiar with multiple languages. In the experimental group, 10 of the 68 students usually speak a language other than Dutch with their mother and 6 students usually speak a language other than Dutch with their father. The control group shows approximately the same pattern as 3 students usually speak a language other than Dutch with their mother

---

<sup>37</sup> From the age of four, the Netherlands offers eight years of primary education and four, five or six years of secondary education. Secondary education is stratified into four mainstream school types: VMBO/MAVO – 4 years, HAVO – 5 years, VWO/Gymnasium – 6 years. For additional information see 3.1.

and 7 usually speak a language other than Dutch with their father. 12 of the students in both groups speak as much Dutch with their mother as they do another language; when it comes to language spoken with the father, this is true for 9 and 10 of the students in the experimental group and control group respectively. 46 of the 68 mothers in the experimental group speak Dutch as opposed to 43 of the 59 in the control group. The numbers are comparable for the fathers, with 51 of the 68 in the experimental group as opposed to 37 of the 59 in the control group.

Looking at the amount of English teaching received in primary school: in the control group 15 students received English in the final grade only, as opposed to 10 in the experimental group. 33 students of the control group had English in the final 2 grades of primary school, as opposed to 43 in the experimental group. 20 students of the control group received English in the final 4 grades of primary school, as opposed to 17 in the experimental group. The smallest numbers are found in the group that were either taught English throughout primary school, or not at all. In the control group one student had 8 years of English language teaching in primary school, as opposed to 3 in the experimental group. 2 students in the control group never received any English, as opposed to 1 in the experimental group.

#### **6.2.5 Fidelity check of the implementation of the language teaching programme based on differentiated instruction**

The cooperative nature of the applied treatment in the experiment implied that the fidelity check in the experiment had the character of consultation about the practical implementation of the principles and activities. During the monthly visits and intermediate e-mail contacts the teachers in the experimental group shared their practical problems with the planned programme and it became clear they were not able to maintain all aspects of the differentiated activities as intended and described in chapter 5. In general, they only completely applied the principles and planned activities in the first four months. The organisation of cooperative learning and feedback moments was found to be most difficult, together with the rubric use. Their attention to offering students choice in process and product as well as relating the planned curriculum to student interests was sustained for the duration of the experiment. These problems with the

implementation of the treatment led to the Delphi study as described in **chapter 7**, in order to gain insight into the causes, reasons or reasoning behind the incomplete implementation of the treatment as intended.

### **6.3 Analyses of the data won through the instruments used to assess the effect of the language teaching programme based on differentiated instruction**

To verify the validity of the measurement of the different constructs of the MPB, as part of the answer to sub-question: **a) *Has the measurement of the MPB been reliable and valid?*** confirmatory factor analyses (CFA; using maximum likelihood estimation) were conducted using version 7 of the Mplus program (Muthén & Muthén, 1998–2006). The fit of the MPB to the data was evaluated by means of several fit indices. The chi-square statistic ( $\chi^2$ ) tests exact fit, which is too strict a criterion for the social sciences (MacCallum, Browne, & Sugawara, 1996). Also the  $\chi^2$  statistic is highly sensitive to sample size. Therefore, the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), the Standardized Root Mean Square Residual (SRMR) and the Root Mean Square Error of Approximation (RMSEA) were also examined. Generally, the model fit is considered acceptable (or ‘fair’) when TLI and CFI are larger than .90 and good when above .95 (Bentler, 1992; Hu & Bentler, 1999). Furthermore, for a good fit, SRMR should not exceed .08 (Hu & Bentler, 1999). In addition, RMSEA values below .05 are considered indicative of close fit, values between .05 and .08 indicate fair fit, values between .08 and .10 indicate mediocre fit and values above .10 poor fit (Hu & Bentler, 1999; MacCallum, Browne, & Sugawara, 1996). Modification indices and standardized residuals were used to locate the items that caused misfit, and after that, the content of these items was reviewed. Decisions to remove items were made on both empirical and theoretical grounds. In subsequent steps, CFAs were performed again with adjusted models.

First the a priori factor structure (i.e. 26 items representing five factors: Cognition, Affect, Subjective Norm, Perceived Behavioural Control and Intention) was tested, both for the pre-test data and the post-test data. In these models only the measurement model was specified, allowing correlations between all latent constructs by default. After removal of items causing misfit, a final model was constructed.

Both for the pre-test and post-test data, the final model was fitted with the MPB as structural model (see Figure 1). Separately for pre-test and post-test the fit of both models (with and without MPB as structural

model) is compared to verify whether the more parsimonious model based on the MPB is not a significantly worse fit. The comparison of the nested models (models with and without the MPB as structural model) is done by means of the difference in chi-square. This difference has a chi-square distribution, with the difference in degrees of freedom of both models as the new number of degrees of freedom. When this chi-square is not significant, the most parsimonious model is preferred. When the chi-square is significant, the loss of degrees of freedom results in a significantly better model fit and so the less parsimonious model is preferred.

Since the scores on the items used as variables in the confirmatory factor analyses are categorical, the fit of the final models for both pre- and post-test data is also established using the option 'categorical' in Mplus. With this option, a different estimator (WLSMV) is used and nested models are compared using the option 'Difftest' that MPLUS offers for comparing nested models when the WLSMV estimator is used. This option also gives a chi-square, numbers of degrees of freedom and p-value, to verify the significance of the differences in fit between nested models.

As indication of the reliability of the sums that were created for each MPB-construct, Cronbach's alpha was calculated for each of the constructs. The alphas found need to be at least .60 for research at group level; .70 is 'a satisfactory level' (Bryman, 2012, p. 170). For constructs of the MPB of which items were removed in the CFA, alpha was calculated twice: once with and once without the aforementioned items, to check whether the alphas were still high enough after items had been discarded.

To answer research question **b) Do pre-test scores of experimental and control students differ at the start of the experiment?**, first all scores were recoded. The MPB questionnaire contains positively and negatively worded items. Before summing, items are recoded to ensure that on all items higher scores indicate a more positive attitude.

The means of the sums of the items measuring each of the MPB constructs are presented and interpreted. These item sums are divided by the number of items, so that the means on sums vary on the same range as the Likert scale used. This is done to facilitate interpretation of the means.

For all dependent variables used in the regression analyses (MPB constructs and language skills) differences between pre-test scores of experimental and control group students are tested by means of multi-level

regression analyses with pre-test scores as dependent variables, to verify whether both groups are comparable at the start of the experiment.

To answer research question *c) 'Which constructs of the MPB best predict the intention to study the*

*English language?'* the standardized paths of the structural model fitted are inspected. Since a path model is a form of regression model, suppressor effects are possible. Therefore, also the zero order correlations between constructs of the MPB are presented and compared to the standardized paths.

The answer to sub question *d) 'If the MPB has been reliably and validly measured, what are the effects of the language teaching programme based on differentiated instruction on changes in the components of*

*the MPB: Cognition, Affect, Subjective Norm, Perceived Behavioural Control and Intention?'* is obtained

through regression analysis, in which the post-test score of each MPB-construct is consistently used as dependent variable and the pre-test score as predictor in the first regression model. As the data is won

from students in year groups, it is first ascertained whether in a model with only the pre-test score as predictor, the intra-class correlation is significantly larger than zero for each construct of the MPB. If so,

regression analyses are conducted multi-level. The next step is to check whether the addition to the regression model of the variable, indicating student's placement in the experimental or control group,

provides a significant improvement of model fit. Additionally, for each construct it is checked whether

background variables significantly correlate with the changes between pre- and post-test. If so, again it was

checked whether the addition of the variable indicating student's placement in the experimental or control group explains the change between pre- and post-test, albeit after controlling for all background variables

that show to be significant. Continuous independent variables are grand mean centred<sup>38</sup> before including

them in the regression analyses, so that the intercept can still be seen as the mean score of students

scoring zero on all dummy variables and average on all continuous predictors.

Only the background variables that have a significant effect will be included in the tables shown.

---

<sup>38</sup> 'Grand mean centred' means that each of the score is linearly transformed by subtracting the mean of each score. Thus the new mean will be equal to zero. One of the results of grand mean centring predictors is that the intercept can be interpreted as the predicted value of the outcome variable for students that have a zero score on all predictors (which is equal to the mean for all grand mean centred variables).

Testing the significance of the model fit improvement after addition of a variance level to the regression model is done by means of the chi-square distributed difference in deviance ( $-2 \cdot \log\text{likelihood}$ ). The number of degrees of freedom for this test is equal to the difference in the numbers of parameters estimated in both models. So, when testing the significance of fit improvement after adding one variance level, the number of degrees of freedom is one. This difference should be tested 1-sided since variances are always positive (Hox, 2010). Therefore, when the chi-square distributed difference in deviance is used for testing the significance in model fit improvement after adding a variance level to the model, the chi-square is tested at .10 (which is comparable to testing one-sided at .05). The chi-square is significant at 10% (=5% one sided) with one degree of freedom when larger than 2.706.

For comparison of nested models after adding predictors the difference in deviance can also be used. Again this difference is chi-square distributed with the difference in the number of parameters estimated as the number of degrees of freedom. So, when testing the model fit improvement after adding one predictor to the model, the chi-square is significant at 5% when larger than 3.841. The significance of regression weights is calculated by dividing the weight by its standard error (Wald-test) resulting in a t-score with the number of observations minus the number of predictors minus one as degrees of freedom. For predictors that vary at student level, the number of observations is equal to the number of students. For predictors at class level, Hox (2010) recommends a strict criterion, namely using the number of classes as the number of observations (Hox, 2010). Hox's advice will be followed, unless the number of predictors exceeds the number of classes minus 1. When this is the case, following Hox's criterion will lead to a negative number of degrees of freedom. Therefore, the significance of adding a predictor in those instances is only evaluated by means of the difference in fit between nested models.

When looking at the language skills measurement, the first instrument discussed is the editing test. As indicated in the instruments section, two different versions of the editing tests were used. For the pre-test both versions were randomly distributed among students. As post-tests student received the version of the test they did not receive for the pre-test. Both these editing tests resulted in two scores: the number of correctly edited words (EdCor) and the number of incorrectly edited words (EdIncor). It was assumed that both versions of the test measure the reading skill: the ability to make correct revisions and the ability to

not make incorrect revisions. Correlations were calculated to verify whether the scores ('EdCor' and 'EdIncor') are measurements of different traits of the reading skill. For the scores of the two versions of the correction test of both traits to be comparable over all students, for pre- and post-test, the scores of each of both versions were first converted separately into z-scores. This resulted in four z-scores for the pre-test and four z-scores for the post-test: correctly edited words separately for pre- and post-test and incorrectly edited words separately for pre- and post-test, both for the A and B version. Since tests were administered randomly to students at pre-test, it can be assumed that the two z-scores for one trait both at pre- and post-test can be seen as comparable indicators of the same trait. The final variables used are constructed by combining the z-scores indicating the same trait at the same measurement moment (e.g. correctly edited words at pre-test) coming from different versions of the editing test.

Of the language skills measurements, research question **e) 'Has the measurement of the language skills been reliable?'** only the reliability of vocabulary and spelling was estimated by means of Cronbach's alpha. For reading and writing skills the reliability could not be estimated. Reading skills are measured by means of a revision test. This test has no items, since students are required to underline superfluous words in a text. For writing skills, a holistic judgement of writing quality given by two cooperating assessors was used. The last research question, **f) 'If the measurement of the language skills has been reliable, what are the effects of the language teaching programme based on differentiated instruction on changes in student's English I) vocabulary, II) spelling skills, III) reading skills and IV) writing skills?'** is answered in the same way as research question d) by means of regression analyses.

#### **6.4 Results of the analyses of the data won through the instruments used to measure the effect of the language teaching programme based on differentiated instruction programme**

In order to answer the question: **a) 'Has the measurement of the MPB been reliable and valid?'**

Confirmatory Factor analyses (CFA; maximum likelihood) were conducted as described above.

The results of these analyses can be found in table 6.1. The table (6.1) shows that for both the pre- and post-test well-fitting measurement models were found after the removal of four of the 26 items, of which one (#16) had no variance in the pre-test or the post-test (all the students gave the same answer to

question 16 both on pre- and post-test administration of the questionnaire). As indicated in the above, models were fitted with and without the option 'categorical'. The results of the CFAs without the option 'categorical' shows a fair (RMSEA<.08) fit in model 4, but not an exact fit (Chi-square is significant). Also the value of SRMR indicates a good fit (<.08). However, CFI and TLI are below .90, indicating a less than acceptable fit. The same goes for the fit based on the post-test data. Furthermore, it can be seen that both for pre- and post-test the MPB is the preferred model when compared to a model where all possible correlations between constructs are permitted. Both for pre- and post-test the difference between the fit of model 4 with and without MPB as a structural model is not significant. The fit indices obtained when the option 'categorical' is used gives a much rosier picture. Exact fit does not hold for pre- and post-test, but RMSEA indicates a close fit for the pre-test ( $\leq .50$ ) and a fair fit ( $\leq .60$ ) for post-test data; CFI and TLI values also indicate a fair fit ( $>.90$ ). Again, both for pre- and post-test adding the MPB as a structural model does not significantly deteriorate model fit, meaning that the MPB is the preferred more parsimonious model both for pre- and post-test. Furthermore, all 22 items in the models 4 have significant factor loadings on the factor they are supposed to measure. The conclusion is that the MPB model shows a fair fit for pre- and post-test data, as the fit-indices for the models with the MPB as structural model indicate a 'fair' to 'good' fit for both the pre- and post-test and do not fit significantly worse than the less parsimonious model without the MPB as a structural model.

In the pre-test model the estimated correlation between Cognition and Subjective Norm is larger than 1, which is possible in an estimated structural model. This might be seen as an argument for combining these two constructs into one. In the post-test however the correlation between these constructs is substantially lower than 1 (.844). Since the aim is to detect whether the experiment favourably changes the scores on constructs of the MPB, Cognition and Subjective Norm constructs have not been combined into one construct for pre-test data. In order to answer questions about the effect of the treatment, identical models (sums based on the same items) for pre- and post-test are preferred.

As indication of the reliability of the pre- and post-test sums for each MPB-construct, Cronbach's alpha has been calculated. In the model fits, 4 items have been removed; items 3 and 6 have been removed from Cognition, item 7 from Affect and item 16 from Subjective Norm. For both Cognition and Affect the alphas

have been calculated with, as well as without, the removed items, to establish whether the removal of the items impairs the reliability of the sums. This has not been done for the Subjective Norm as item 16 did not show any variance and therefore had to be excluded from the pre- and post-test sums for Subjective Norm. The alphas obtained can be found in table 6.2. The table (6.2) shows all sums have a sufficiently high alpha (>.6) and, apart from three, a 'satisfactory' (.7) reliability or better. As the alphas for Cognition and Affect are clearly higher when including the items that had been removed in the CFA, the multi-level regression analyses have been performed with sums including all original items (excluding 16, as it showed no variance). The correlations between sums with and without aforementioned items are for the pre- and post-test of Cognition respectively .945 (N=133) and .971 (N=127) and for the pre- and post-test of Affect respectively .981 (N=133) and .977 (N=127).

**Table 6.1.: Results confirmative factor analyses Model of Planned behaviour for pre- and post-test data (N pre-test=133; N post-test=127)**

model	$\chi^2$	df	p	RMSEA (90% C.I.)	CFI	TLI	SRMR	Deleted items
Pre-test model 1	490.809	265	.000	.080 (.069-.091)	.791	.764	.086	16
Pre-test model 2	422.493	242	.000	.075 (.063-.087)	.824	.799	.081	3, 16
Pre-test model 3	357.564	220	.000	.069 (.055-.081)	.845	.822	.068	3, 7, 16,
Pre-test model 4	308.808	199	.000	.064 (.050-.078)	.864	.842	.066	3, 6, 7, 16,
Pre-test model 4 plus MPB	312.236	202	.000	.064 (.050-.078)	.863	.843	.067	3, 6, 7, 16
Difference pre-test mod 4 with and without MPB	3.428	3	n.s.					
Post-test model 1	601.093	265	.000	.100 (.089-.111)	.753	.720	.108	16
Post-test model 4	355.101	199	.000	.079 (.065-.092)	.843	.817	.074	3, 6, 7, 16
Post-test model 4 plus MPB	356.611	202	.000	.078 (.064-.091)	.844	.822	.075	3, 6, 7, 16
Difference post-test mod 4 with and without MPB	1.510	3	n.s.					
<b>Categorical CFA</b>								
Pre-test model 4	265.036	199	.001	.050 (.032-.065)	.945	.937	-	3, 6, 7, 16
Pre-test model 4 plus MPB	270.504	202	.001	.050 (.033-.066)	.943	.935	-	3, 6, 7, 16
Difference pre-test mod 4 with and without MPB	7.061	3	.07					3, 6, 7, 16
Post-test model 4	288.720	199	.000	.060 (.044-.074)	.949	.941	-	3, 6, 7, 16
Post-test model 4 plus MPB	290.722	202	.000	.059 (.043-.073)	.950	.942	-	3, 6, 7, 16
Difference post-test mod with and without MPB	3.513	3	.319					3, 6, 7, 16

**Table 6.2: Alphas of the sums of attitude-items in pre- and post-test (complete and based on the CFA (Rit=item-test correlation: items can be found in appendix 4, 'r' indicates items that are reverse scored)**

Variable	Items	Alpha	Range rit's	Deleted items
<b>Pre-test (N=133)</b>				
Cognition pre-test CFA	1r, 2r, 4r, 5	.705	.394-.544	3, 6
Cognition pre-test	1r, 2r, 3r, 4r, 5, 6r	.747	.269-.577	-
Affect pre-test CFA	8r, 9, 10, 11	.734	.487-.640	7
Affect pre-test	7r, 8r, 9, 10, 11	.803	.456-.661	-
Subjective Norm pre-test	12r, 13r, 14r, 15r	.633	.350-.461	16r <sup>a</sup>
PBC pre-test	17r, 18r, 19r, 20r, 21r	.716	.340-.535	-
Intention pre-test	22r, 23r, 24, 25r, 26r	.678	.358-.509	-
<b>Post-test (N=127)</b>				
Cognition post-test CFA	1r, 2r, 4r, 5	.804	.452-.746	3, 6
Cognition post-test	1r, 2r, 3r, 4r, 5, 6r	.842	.421-.743	-
Affect post-test CFA	8r, 9, 10, 11	.736	.461-.647	7
Affect post-test	7r, 8r, 9, 10, 11	.784	.471-.640	-
Subjective Norm post-test	12r, 13r, 14r, 15r	.758	.397-.632	16r <sup>a</sup>
PBC post-test	17r, 18r, 19r, 20r, 21r	.741	.430-.664	-
Intention post-test	22r, 23r, 24, 25r, 26r	.632	.258-.480	-

a) item 16 showed no variance

**b) Do pre-test scores of experimental and control students differ at the start of the experiment?**

In table 6.3 the means of the pre-test scores (N=133; N-exp=70; N-control=63) are presented together with standard errors (se) and standard deviations (sd). The means of the pre-test scores, related to the Likert scale used (1 = I completely disagree, 2 = I partially disagree, 3 = I do not agree, nor disagree, 4 = I partially agree and 5 = I completely agree), show the lowest pre-test mean score (3.502 se=.080 sd=.917) is that of Affect for the whole sample as well as in the experimental group (3.557 se=.110 sd=.921) and control group (3.441 se=.115 sd=.916). In general, this means students tend to like the subject of English, as they agree with statements like: ‘learning English is great.’ and ‘learning English is fun.’. Keeping the standard deviation in mind, even the students’ lowest mean score leaned more towards ‘I do not agree, nor disagree’, than disagreement.

The highest pre-test mean score (4.211, se=.068, sd=.786) is that of Cognition for the whole sample as well as in the experimental group (4.218 se=.105 sd=.877) and control group (4.202 se=.085 sd=.676). In general, this means students see English as useful, as they agree with the statements like: ‘You need to learn the English language to understand the world.’ and ‘English is important for daily life.’. Keeping the standard deviation in mind, the students’ lowest mean scores leaned towards at least partial agreement, while some completely agreed.

**Table 6.3.: Means, standard errors (se) and standard deviations (sd) of pre-test scores on MPB constructs. Means on original Likert scale (1 = I completely disagree, 2 = I partially disagree, 3 = I do not agree, nor disagree, 4 = I partially agree and 5 = I completely agree)**

MPB construct	Pre-test mean (se)	Sd pre-test mean	Pre-test mean experimental group (se)	Sd pre-test mean experimental group	Pre-test mean control group (se)	Sd pre-test mean control group
Cognition CFA	4.211 (.068)	.786	4.218 (.105)	.877	4.202 (.085)	.676
Cognition all items	4.085 (.064)	.737	4.112 (.099)	.824	4.056 (.080)	.631
Affect CFA	3.579 (.081)	.932	3.629 (.110)	.921	3.524 (.119)	.948
Affect all items	3.502 (.080)	.917	3.557 (.110)	.921	3.441 (.115)	.916
Subjective norm	3.795 (.068)	.778	3.846 (.099)	.825	3.738 (.091)	.726
PBC	3.923 (.065)	.749	4.014 (.101)	.841	3.822 (.078)	.622
Intention	3.708 (.075)	.866	3.726 (.107)	.895	3.689 (.106)	.838

Multi-level regression analyses show that the pre-test scores for all MPB constructs do not differ significantly between experimental and control group students (see appendix; Tables AP10B1 to AP10B5).

The same can almost be said for the pre-test scores of language skills measurement. Multi-level regression analyses also show that the pre-test scores for correctly underlined superfluous words and incorrectly underlined words of the reading test do not differ significantly between experimental and control group students (see appendix; tables AP10B6 to AP10B8). The multi-level regression analyses of pre-test scores of the dictation test, as reported in table AP10B9 (see appendix), shows the experimental group scores significantly higher than the control group (3 points  $p < .05$ ) at the start of the study. The significance shows when model 4 is compared to model 5 (2.479% explained variance on student level and .361% explained total variance). The conclusion is that, apart from 1 variable (pre-test-dictation scores), the students in the experimental group did not significantly differ from the students in the control group at the start of this study.

The answer to research question: *c) 'Which constructs of the MPB best predict the intention to study the English language?'* has been given in the structural models (Figure 6.4-pre and Figure 6.5-post). The assumed causal pathways of the MPB in the pre-test show that Cognition has a significant path on Affect, which in turn has a significant path on Intention. The model fitted to the post-test data is very much like the model based on pre-test data. A difference between both models is that the PBC path to intention is not significant in the pre-test model and is significant, albeit modestly, in the post-test model. Also, as mentioned in the above, the pre-test model shows a correlation between Subjective Norm and Cognition larger than 1 while in the post-test model this correlation is only .844.

Changes between pre- and post-test, like the shift in the path from PBC to Intention, could be due to the treatment offered to half of the students, but this might also be caused by maturation.

It is clear that Cognition, Subjective Norm and PBC strongly correlate in the above models.

It is important to keep in mind the structural models used are regression models which implies that suppressor-effects<sup>39</sup> may occur, therefore in table 6.6 and table 6.7 (appendix 10.C) the zero-order

---

<sup>39</sup> When correlating variables are used in one analysis to predict one and the same dependent variable, the regression coefficients of the predictors can no longer be seen as indicative of how high a predictor correlates

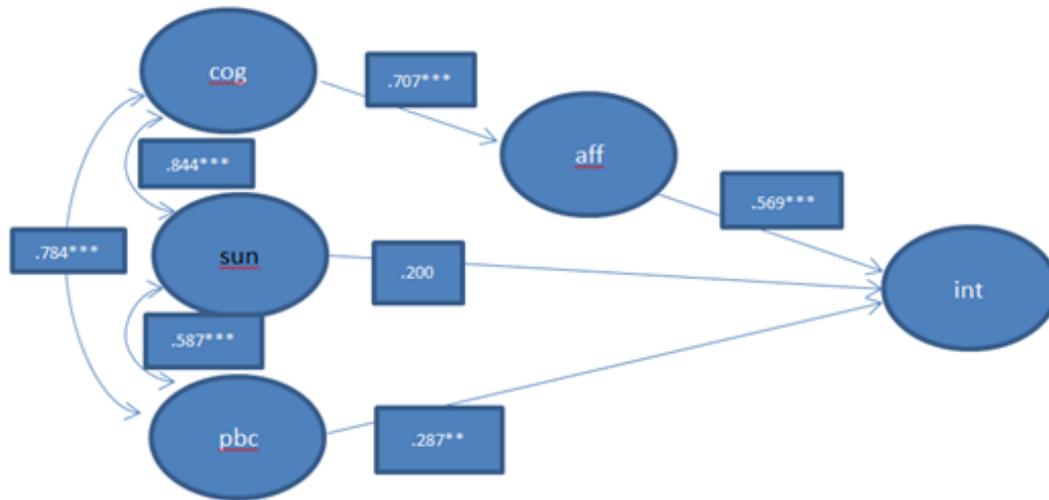
correlations between MPB-scores are given. These tables show that all of the MPB scores correlate relatively highly with the scores on Intention: pre-test from .462-.577 (appendix 10.C: table 6.7) and post-test from .491-.616 (appendix 10.C: Table 6.6). Therefore, the conclusion based on the results of the confirmatory factor analyses is that the intention to make an effort for studying English is probably influenced most by whether students like to study English (Affect), which in turn is influenced by the perceived usefulness (Cognition). However, taking the zero-order correlations into consideration, effects of Subjective Norm and PBC on the intention to make an effort to study English, cannot be excluded. So the answer to research question b) is that there is a clear indication that the pathway from Cognition to Affect to Intention is the most important for English language learning in the sample of students in secondary education in this study, but that correlations between Intention on the one hand and Subjective Norm and PBC on the other are relatively high and thus that a causal effect of both of these variables on Intention cannot be excluded.

**Figure 6.4: Pre-test model MPB estimated with option ‘categorical var.’: standardised factor loadings (\*= $p < .05$ ; \*\*= $p < .01$ ; \*\*\*= $p < .001$ )**



with the dependent variable. This is caused by the fact that when, in a regression model with more than one predictor, variance is explained by one variable, it can no longer be explained by another variable.

Figure 6.5: Post-test model MPB estimated with option ‘categorical var.’: standardised factor loadings (\*=p<.05; \*\*=p<.01; \*\*\*=p<.001)



### Regression Analyses of the effect of the treatment on MPB scores

In order to answer research question *d) 'If the MPB has been reliably and validly measured, what are the effects of the language teaching programme based on differentiated instruction on changes in the components of the MPB: Cognition, Affect, Subjective Norm, Perceived Behavioural Control and Intention?'* regression analyses have been performed as explained in the analysis section, 5.4. Multi-Level analyses have been used to ascertain the effect of the treatment on the attitude scores with background variables as well as the effect of the treatment on the attitude scores without background variables. For Cognition and Affect some items were removed in the CFAs to improve model fit. Therefore, for these two constructs, two possible sum scores are conceivable: sum scores based on only the items used in the final model resulting from both the CFAs and sum scores based on all items intended to measure the construct. In tables 6.6 and 6.7 (appendix 10.C) the correlations between both possible sums are reported. Given that these correlations are very high (for Cognition and Affect respectively .971 and .977 on pre-test and .945 and .981 on post-test) and that the reliability of the sums based on

all items are higher than those of the sums based on the items used in the final CFA models, sums based on all items are used in the multi-level regression analyses.

The step by step analyses of the effect of the treatment on changes in MPB scores (Cognition, Affect, Subjective Norm, Perceived Behavioural Control and Intention), with and without controlling for significant covariates, is presented in tables and accompanied by explanatory text in the appendix (10C).

### **Overall conclusion of the effect of the treatment on MPB constructs**

Generally speaking, a positive effect of the treatment on MPB constructs has been found. The students' view on the usefulness of learning English has improved, as the treatment shows a positive effect on the change in Cognition and this positive effect remains after controlling for the significant covariates 'country of birth mother' and 'country of birth father'.

The degree to which students 'liked' English (Affect) is also positively influenced by the treatment, whether controlled for the significant covariates or not. Affect for English as a subject is, however, also influenced by time spent watching TV (in the English language) and time playing games (in the English language), as both show to have significant positive correlations with more growth of Affect.

As far as the amount of Perceived Behavioural Control (PBC) is concerned, the treatment is found to have a significant positive effect on both when not correcting for covariates. After correcting for the significant covariate, age, the variable condition is no longer significant.

When looking at the Subjective Norm of students, not correcting for covariates means the students in the treatment group show a one-sided significantly more favourable change in Subjective Norm scores than the control group. This positive effect can be caused by the treatment, although it could also be an effect of the amount of English language teaching received in primary education (Eibo), as after controlling for Eibo, the variable 'condition' is no longer significant.

The development of the scores for Intention are also significantly more favourable for the students in the experimental group than for the students in the control group. This effect is significant with and without controlling the results for covariates.

All in all, the treatment has been found to have positive effect on the MPB constructs of the students in the sample of this study towards learning English as a subject in secondary education.

### **Reliability of language skills measurement**

In order to answer research question *e) 'Has the measurement of the language skills been reliable?'*

first Cronbach's alpha was calculated for the sums for the vocabulary and the listening/spelling tests for both pre- and post-test, as explained in the 'Analyses' chapter 6.3. In table 6.13 (appendix 10.C) the alphas are reported. As can be seen in the table (appendix 10.C: 6.13), the alpha's for vocabulary and dictation for both pre- and post-test are sufficient for research on group level (above .6). Due to the absence of items in the writing test, no alpha's have been calculated for writing scores. The reading test has been offered in two versions with completely different scores which have been turned into z-scores, as described in the analyses chapter, and do therefore not allow the calculation of alphas. Holistic assessors were used for the assessment of the writing products. These pairs of assessors only reported their scores as pairs, after having discussed and decided upon scores, which prevented the assessment of inter-rater reliability.

In order to answer research question *f) If the measurement of the language skills has been reliable, what are the effects of the language teaching programme based on differentiated instruction on changes in*

*student's English I) vocabulary, II) spelling skills, III) reading skills and IV) writing skills?* multi-level regression analyses were performed, as described in the analyses chapter. For each of the language aspects measured (vocabulary, spelling skills, reading skills and writing skills), the results of the analyses will be discussed below. The step by step analyses of the effect of the treatment on changes in language skills test scores (vocabulary, spelling skills, reading skills and writing skills), with and without controlling for significant covariates, is presented in tables and accompanied by explanatory text in the appendix (10C).

### **Results regression analyses of the effect of the treatment on language skills**

The regression analyses concerning the effect of the treatment on vocabulary (see Table 6.14 in appendix 10.C) shows that 7% of the variance in gains in vocabulary scores is school-related. The treatment does not

show a significant effect on changes in vocabulary. Furthermore, results show that girls' vocabulary scores progress slower than that of boys. Next to gender, the Cito-test score significantly predicts changes in vocabulary over time, as do time watching TV and time playing games. When these significant covariates are added to the regression model, the treatment is still not significant. The conclusion is therefore that the treatment, both analysed with and without correcting for covariates, does not show a significant effect on changes in vocabulary from pre- to post-test.

The results relating to the changes in the scores of the listening/spelling test (dictation) show a significant effect of the treatment on gains in dictation scores. After controlling for the pre-test score, the experimental group scores 2.7 points higher in the post-test than the control group. The treatment shows a positive effect on the development of listening/spelling skills without correcting for covariates.

Significant covariates are 'country of birth', 'native tongue mother' and 'language spoken with mother'. The 'country of birth father' variable is significant if the Wald test is taken as criterion but not when the difference in deviance between nested models is used. The effect of the treatment remains significant after correcting for significant covariates and even explains all class-related variance. So, the treatment shows a significant effect on the trend in dictation scores both with and without correcting for the significant covariates.

Quite noticeable are the covariates referring to having Dutch speaking parents vs. having parents that speak other languages. Students with a Dutch background (country of birth pupil, native tongue mother and language spoken with mother) consistently show a less positive trend on their scores for dictation. Perhaps the children with language backgrounds other than Dutch have developed more metalinguistic awareness or are more eager to add an international language to their language skills repertoire. It might be that the selection based on language background coincides with other traits not researched in this study. An explanation for these differences in trends cannot be based on the data gathered in this study and further research is needed to find out what might be causing this effect.

The effects of the language teaching programme based on differentiated instruction on Reading skills is investigated with two different variables: number of correctly underlined words and number of incorrectly

underlined words. This sub-division is based on the correlations calculated between the sum scores of the different tests as has been described in the analyses chapter above. The correlations (see appendix 10.C.: table 6.28.a), clearly show two different traits of the reading skill have been measured.

Without controlling for covariates, no significant difference in growth is found for the scores for correctly underlined words in the reading test between the experimental and control group. After correcting for the significant covariates 'Cito-test', 'Time reading English books for leisure' and 'Language spoken with father', a difference is seen in the scores for correctly underlined words in the reading test.

Analyses for the variable, incorrectly underlined words, show no effect of the treatment on growth in scores of incorrect underlined words in the editing test. Furthermore, results show that, apart from the Cito-test, none of the covariates is a significant predictor of the trend in incorrectly underlined words in the editing test. After controlling for the pre-test and the Cito score, the treatment shows a significant positive effect on the score for incorrectly underlined words.

The treatment turns out to be a significant predictor for the development in writing scores. 'Time watching television' appears to be the only (one-sided) significant covariate. After controlling for this variable the effect of the treatment remains significant and explains 6.6% of the variance in the development in writing scores.

### **6.5 Conclusion of the quasi experiment used to assess the effect of a language teaching programme based on differentiated instruction on the attitude towards learning English and the English language skills development of students in the first year of secondary education.**

To conclude this chapter on the quasi experiment used to assess the effect of the differentiated programme on the English language learning gains of first year students in secondary education in The Netherlands, first the additional research questions will be answered.

#### **a) *Has the measurement of the MPB been reliable and valid?***

The CFA showed the MPB to be more parsimonious than the measurement model without specifying a structural model. The measurement of the MPB has shown to be reliable with sufficient alphas and, apart

from three variables, a 'satisfactory' (.7) alpha. According to the RMSEA, CFI and TLI that indicated a 'close fit' and two 'fair fit' respectively, the MPB measurement was also valid.

**b) Do pre-test scores of experimental and control students differ at the start of the experiment?**

The heights of the MPB pre-test scores showed that the students in the sample of this study generally already tended to have a positive attitude towards the subject of English the first time the attitude test was offered. Even the lowest mean pre-test scores indicated there was at least a trend to agree with statements on Affect. The other constructs of the MPB had even higher pre-test mean scores.

For all MPB constructs it applies that no significant difference between the pre-test scores of the students in the experimental and control group were found. Apart from the pre-test-dictation scores, the same applies for the language skills instruments. This means that, generally speaking, the experimental group did not significantly differ from the students in the control group at the start of this study and both groups are comparable.

**c) Which constructs of the MPB best predict the intention to study the English language?**

There is a clear indication that the pathway from Cognition to Affect, to Intention is the most important for English language learning in the sample in this study of students in secondary education; the intention to make an effort to study English is probably influenced most by whether students like to study English (Affect), which in turn is influenced by the perceived usefulness (Cognition).

**d) If the MPB has been reliably and validly measured, what are the effects of the language teaching programme based on differentiated instruction on changes in the components of the MPB: Cognition, Affect, Subjective Norm, Perceived Behavioural Control and Intention?**

Overall, the treatment was found to have a positive effect on the attitude towards learning English as a subject in secondary education. The students' view on the usefulness of learning English improved, as the treatment shows a positive effect on the change in Cognition and this positive effect remains after controlling for the significant covariates 'country of birth mother' and 'country of birth father'.

How much students 'liked' English (Affect) is also positively influenced by the treatment, whether controlled for the significant covariates or not. Affect for English as a subject is, however, also influenced by time spent watching TV (in the English language) and time playing games (in the English language), as both show to have significant positive correlations with more growth of Affect.

When it comes to the amount of behavioural control students perceive they have over learning English, as shown in PBC-scores, the treatment is found to have a significant positive effect when not correcting for covariates. Covariates 'age' and the amount of English language teaching received in primary education (Eibo), on the other hand, are such important factors that when correcting for the significant covariate 'age', the variable condition is no longer significant.

***e) Has the measurement of the language skills been reliable?***

The alphas needed to be sufficient for research on group level (>.6). With .705 (vocabulary) and .948 (dictation) for the pre-test as well as .691 (vocabulary) and .922 (dictation) for the post-test, the alphas clearly show that the measurement of language skills were reliable for both the vocabulary and dictation test.

***f) If the measurement of the language skills has been reliable, what are the effects of the language teaching programme based on differentiated instruction on changes in student's English I) vocabulary, II) spelling skills, III) reading skills and IV) writing skills?***

The effect of the treatment was analysed separately for each of the skills: Vocabulary, Dictation, Reading (editing test) and Writing. The only instrument that did not show a significant effect on changes from pre- to post-test was the vocabulary test.

The treatment did show a significant effect on the trend in dictation scores both with and without correcting for the significant covariates. A positive trend is found for the scores for correctly underlined words in the reading test after correction for significant covariates. The incorrectly underlined words score in the reading (editing) test the treatment shows a positive effect after controlling for Cito-scores. The effect of the treatment on the Writing scores is significant and explains the 6.6% variance.

The above means that, apart from the vocabulary test, the results in scores from the other language skills as well as the scores from the attitude questionnaire, based on the MPB, show measurable positive effects and significant positive effects of the treatment offered in this experiment.

## **7. Delphi study on the factors impeding differentiated language teaching**

### **7.1. Introduction**

The positive effects of the language teaching programme based on differentiated instruction, found and reported on in the previous chapter, are quite satisfactory, especially when the implementation of the treatment is taken into account.

The number of schools and classes reported on in the above study is not the whole group of participants involved at the start of this study. Despite a number of preliminary steps taken as condition to ensure full partnership in this research and its experiment, i.e.:

- the involvement of English teachers from 9 interested secondary schools in The Netherlands in the preliminary discussions on the possibilities of differentiating language teaching;
- their involvement in the development of the instruments used in the experiment;
- pre-experimental phase training sessions;
- a digital platform offered with all instructions, explanations and materials;
- monthly visits for coaching on the implementation of the treatment as intended; as well as
- the possibility to have more visits, feedback or suggestions on demand,

only three schools remained in the experiment long enough to be reported on in the above chapters.

The three schools reported on made every effort to implement the language teaching programme based on differentiated instruction as intended, but as already reported in chapter 6.2.5, the school visits and intermediate contacts showed the teachers found the implementation difficult and only managed to maintain the programme as intended in the first months, after which they let go of a number of requirements and effectively offered half the treatment in the second part of the academic year.

The schools and teachers who abandoned the experiment all had their reasons and explanations for doing so. To just describe or list the reasons and explanations given at the time would not lead to the insights needed to help secondary education in general in the effective implementation of differentiated language teaching, as intended by this research.

To clarify which factors led to a number of teachers leaving the experiment and which factors kept teachers from implementing the language teaching programme based on differentiated instruction as intended, more research was needed.

In order to fully answer research question 5, "***Which skills do (future) teachers of English in secondary schools need to acquire or possess to offer a language teaching programme based on differentiated instruction as mentioned above?***" as well as fill the additional need for research on factors impeding differentiated language teaching, a Delphi study was set up; this will be reported on in the following chapter.

The English teachers in the Dutch secondary schools, who participated in the experiment of the research described above, were asked to participate in a Delphi study.

A Delphi study, which will be described more elaborately below, works with a panel of experts that are offered a set of questions or propositions iteratively in order to reach consensus.

The questions posed to the panel, consisting of all teachers of English involved in the experimental phase of this research (experimental and control groups), partly go beyond sub-question 5. The attrition of participants in the experiment, as well as the laborious implementation of the treatment offered, led to an extended focus. Besides the focus on teacher-skills needed to differentiate language teaching, as posed in sub-question 5, the Delphi study also focussed on the challenges met during the implementation of the treatment used in the experiment described in chapter 6. The Delphi study is expected to clarify the factors impeding differentiated language teaching and to offer possible solutions for these factors in future implementation.

## 7.2. Set up of the Delphi study

The original Delphi study was a research technique developed by the RAND Corporation, in a US Air Force Cold War study<sup>40</sup> in the 1940s (Gupta & Clarke, 1996) and 1950s. In this, at the time classified experiment (Dalkey & Helmer, 1951), the Delphi procedure was used "... to obtain the most reliable consensus of opinion of a group of experts ... by a series of intensive questionnaires interspersed with controlled feedback." (Dalkey & Helmer, 1963, p. 458). The three features of the original Delphi procedures are:

- a) Anonymous response to statements and questions, to reduce the effect of dominant individuals.
- b) Iteration and controlled feedback (giving a summary of the results to the participants between rounds), to reduce noise.
- c) Statistical group response. The opinion of each group member is represented in the final response, to reduce group pressure for conformity.

(Dalkey, 1969, p. 16).

The procedure contains the following steps:

- 1) Research question development through experience, Literature Review and/or pilot studies.
- 2) Selection of heterogeneous experts for the panel (between 5 and 20).
- 3) Delphi round 1: the participants anonymously react to the research question(s).
- 4) Results of round 1 are statistically analysed and together with offered rationales returned to the experts as feedback.
- 5) Delphi round 2: participants are asked to react to the feedback by giving arguments on points where they differ in view.
- 6) Results of round 2 are statistically analysed and together with offered rationales returned to the experts as feedback.

---

<sup>40</sup> At the time the consensus of the experts was desired because the US Air force required an accurate estimation of A-bombs the Soviet Union would need to deliver on target per industry in the USA in order to diminish US munitions output by 66%.

- 7) Delphi round 3: participants are asked to react to the feedback by giving arguments on points where they differ in view.
- 8) Results of round 3 are statistically analysed and together with offered rationales returned to the experts as feedback on the whole procedure.

(Rowe & Wright, 2001; Skulmoski, Hartman & Krahn, 2007)

Through a series of experiments to evaluate the procedures in the spring of 1968, Dalkey (1969) found that anonymous controlled feedback made group estimates more accurate than face-to-face discussions. These findings have been corroborated over time by several researchers (Rowe & Wright, 2001; Graefe, 2011).

In the past decades the method has been widely used in health and social care research (Baker, Lovell & Campbell, 2007) as well as in language teaching research. In the Netherlands, for instance, the Dutch 'National Action Programme' on Foreign Languages (NAP) used the Delphi procedure. Through the Delphi procedure a group of 41 Dutch experts from the field of education and the economic world were asked questions on supply and demand of foreign languages in The Netherlands. The procedure yielded thirty-four policy recommendations to the Dutch government to improve the situation in foreign language teaching (Van Els, 1990).

Although the term 'Delphi' is consistently used, Delphi studies can differ greatly in design. These differences arise from the situational nature of the design, in that it is guided by the research problem rather than by the requirements of a method (Hasson & Keeney, 2011). As Skulmoski, Hartman & Krahn (2007) noted: "The Delphi method is a mature and a very adaptable research method used in many research arenas by researchers across the globe." Its flexibility is shown through its use as a tool for judgment, decision-aiding, forecasting, programme planning and administration; to work with incomplete knowledge about a problem or phenomena; to solve problems that could benefit from the subjective judgments of individuals on a collective basis; to focus collective human intelligence on the problem at hand, or to investigate what does not yet exist (Skulmoski, Hartman & Krahn, 2007).

When the design of the Delphi method is altered because of differing aims, situations, problems or administration requirements, this is usually indicated by the use of the term 'modified Delphi research', although Hasson & Keeney (2011) list ten different Delphi design types.

Modified Delphi research in different designs has been widely discussed (e.g.: Gupta & Clarke, 1996; Skulmoski, Hartman & Krahn, 2007; Keeney, 2009) but a general aggregative definition can still be given:

*“Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem.”*

(Linstone & Turoff, 2002, p. 3).

The Delphi design used in this research is a modified version of the original Delphi design, as can be seen in table 7.1 below. Although iterativity and transparency have been retained, a number of aspects were modified; these will be clarified below.

The first modification concerns the selection of experts. Instead of a selection of experts in their field, the panel consisted of the teachers involved in the experiment described in chapter 6, together with their colleagues who were in the control group. All involved worked with the same (type of) students and had practical expertise in English teaching in secondary education.

The second modification is the use of a so called “exploration phase”, which, according to Ziglio (1996, p. 9), should offer broad or open-ended questions to explore the topic fully. The questions used in this study, as given below in the chapter on Delphi round 1, have been offered in one-on-one semi-structured interviews.

The third modification was the use of interviews instead of questionnaires, which is consistent with the other modifications of the features: ideal, philosophy, goal and feedback. The goal of this modified Delphi was a concise analysis of challenges and opportunities regarding the implementation of differentiated language teaching. Therefore, it was imperative to give all responses and key arguments back to the panel in the second round without too much interference from the researcher. Because of the goal and the small scale of this Delphi, a statistical analysis and representation of the won information would not have contributed to the process. The same arguments were applied to the ideal and philosophy of the traditional Delphi. In striving for a concise review of challenges and opportunities, there is no need for a ‘winning’ argument, or for complete consensus.

The features of the traditional Delphi and the modified Delphi as used in this research are summarised in table 7.1.

**Table 7.1 Comparison of traditional Delphi and modified Delphi used in this study.**

<b>Feature</b>	<b>Traditional Delphi</b>	<b>Modified Delphi (this study)</b>
<b>Similarities</b>		
Transparency	Anonymity of arguments	
Iterativity	Multiple rounds	
<b>Differences</b>		
Ideal	The best argument wins	The best arguments are widely recognized by the panel.
Philosophy	Consensus	Consensus, although dissension is possible and allowed.
Panel	Wide selection of experts	Selection of experts is based on experience with the treatment used, or familiarity with the schools participating in the experiment.
Goal	Accurate prediction	Concise review of challenges and opportunities.
Feedback	Median and interquartile	Non-statistical feedback due to group-size: all responses and key arguments are given back to the participants.
<b>Form of data</b>	Questionnaire	Interviews

### **Delphi round 1**

The first Delphi round consisted of 7 questions:

1. Can you describe in your own words what the connection problems are between primary and secondary schools for English teaching?
2. Can you describe in your own words what solutions you see for this problem?
3. Can you describe in your own words what your opinion is of the intervention offered the last year to help differentiate English teaching?
4. Can you describe in your own words which parts of the proposed interventions you see as useful or doable?
5. Can you describe in your own words what is the most effective introduction or implementation of a treatment like the one offered last year?
6. Or another differentiated approach?

7. Is there anything you would like to add concerning the connection problem, possible solutions, the treatment and/or differentiated language teaching?

These questions were offered to all 6 members of the panel in a round of one-on-one, semi-structured interviews in January and February 2015. In preparation for round 2 all answers, arguments and propositions were listed and offered to the panel in a document called: "Delphi Round 1 -Rephrased Complete" (appendix 11). This complete listing of all that was offered by all participants served as background information to the second document: "Delphi Round 1 -Rephrased and summarised for round 2"(appendix 12). In this second document the lists of answers, arguments and propositions were reorganised into propositions when the panel seemed to agree. When the panel was rather divided, juxtaposed positions were given.

In round 2 the panel was asked to react to each other's propositions and restate, change or add arguments to help clarify their position, or challenge the validity of what others claimed. In case of juxtaposed positions found in round one, the panel was asked to weigh and restate their point of view based on the complete list of arguments gathered in round one. These questions were offered to all members of the panel in a round of one-on-one interviews in April 2015.

In preparation of round 3 all answers, arguments and propositions were listed and offered to the panel in a document called: "Delphi Round 2-Rephrased and summarised for round 3 divided in italics (round 1) and italics as well as bold (round 2)" (appendix 13). This complete listing of all that was offered by all participants served as background information to the second document: "Round 3 proposed text"(appendix 14). In this second document the lists of answers, arguments and propositions was reorganised into a text describing the findings, positions and discussion that arose in the first two rounds.

In round 3 the panel was asked to ratify the arguments presented in the text by affirming the story adequately represents each panel member's position and ideas: the results are found below.

### **7.3. Results of the Delphi study**

The panel agrees the articulation between primary and secondary education levels in general and for English teaching in particular is problematic. The students' levels of proficiency in English on entering the

first year of secondary education differ enormously. These differences between students, which one of the panel-members described as 'unexposed to the English language' to 'language skills wise, years ahead of their peers', in English language skills levels have grown in the past decade due to VVTO. Different starting moments (at ages 4, 8 or 10) and the related differences in quantity and quality of English language teaching in primary schools has widened the gap between primary and secondary education. The view on students' English language skills is further obscured by out-of-school exposure (like gaming and watching TV). For instance, students sometimes acquired words like *flint* and *armour* because of their involvement in online role-playing games.

The above excludes simply discriminating between students who have had English and those who have not. The example was given of abler students who were not sufficiently challenged in primary education when acquiring the English language and who enter secondary education ill-prepared for the challenges it brings, because they believed they would be well able to perform. On the other hand, sometimes there are students who are painfully aware of the lack of English language education in their primary school career. They feel their deficiency is an unbridgeable ability gap between them and their peers and have, because of that feeling, even developed inhibitions towards foreign language learning in general.

On a more general note, the panel observed that, generally speaking, in their view, all students coming from primary education lack language the learning strategies and skills needed in secondary education. According to the panel the 'Cito' scores, used to indicate the expected success students would have in a certain stream, does not correctly reflect student aptitude and abilities, nor does it offer information on students' English language skills levels.

The above does not mean the panel feels that problems with the articulation between levels can be attributed solely to the primary school side of the divide. Secondary education also plays a part in this. The teaching approaches used in secondary school do not relate to what the students have experienced in their two to eight years of language education in primary school. Students are not offered the extended instruction or additional help they have grown accustomed to in primary education, and the additional materials needed to help students deal with a deficiency, are not readily available. The panel is aware of the fact that students who previously received none, or little, English language teaching, really need to

work hard in the first year of secondary school. The same goes for the teachers, who find it difficult to serve and challenge around 30 students with different language skills levels. In that sense, the panel respects the primary school approach and would like to apply some of its methodology. The general feeling is that more time could be spent on productive skills, like speaking, and students with deficiencies should be given more structured help. Ideally to have students work more creatively with their knowledge and skills they should be offered more differentiated, theme based, cross-curricular education.

A differentiated approach, like the treatment used in the experiment, was considered to possibly be part of an improvement of articulation between levels. The treatment was generally very well received, with some criticism and notes on the different parts that made up the treatment; these will be dealt with below (highlighted in bold) in separate paragraphs.

One recurrently referred to activity stands out, especially because it was not part of the treatment. All participants, from both the experimental and control group, reacted positively to the pre- and post-**test** used for the experiment. It triggered reactions like: “Very interesting to be informed of initial language skills levels and end-of-year year abilities” (Delphi participant A – Jan. 2015).

Their positive reactions were twofold. On the one hand the teachers were enthusiastic about being informed on language skills levels and background information on previous language teaching received and out-of-school exposure. On the other hand, their idea is to use tests formatively, as part of the educational process, not just as a threshold for secondary education. In that sense the panel would very much welcome and appreciate a clear standard for English language teaching used in primary education. The information of students having had two, four, or eight years of English language teaching does not give any guarantee of acquired language skills.

**Rubrics** can only be used when the students can relate to them. This means that the language used needs to be more on their level and the steps need to be small, clear and concrete (e.g. through instructional videos).

**Chapter-wide assignments** were well-received. The personal involvement of the students in the assignments, process and product helps them to be more engaged. Chapter-wide assignments can be

introduced when students are trained to take their time and revise their work so they can produce the best they have to offer. Revision work is made effective by the cyclical nature of the assignments.

**Cooperative learning** yielded several different reactions. During the interviews it turned out the panel had different views on what cooperative learning entailed and the experiences with these type of activities differed. All agreed it was hard work, required quality teaching skills and it would help to limit the number of students in a classroom to approximately 24.

The '**giving feedback**' part of the treatment was discussed in two ways. The first was giving feedback to peers in the process of learning and cooperation. This was viewed as a real life skill, which needs to be offered and trained in school. Students need to acquire the words with which they can effectively give feedback and to have a safe classroom environment to do so.

The second version of giving feedback referred to the feedback given to peers on a product, like a presentation. The inconsistent harshness, fairness, or even favouritism of peers giving feedback means teachers need to invest in the pedagogical climate in the classroom for it to be effective.

**Differentiating the assessments of products** was something the panel came to think of as more positive between the second and third round, although some reservations remained. The whole panel agreed this would work best if continued right down to the final and central exams. On top of that, the marks for English on the diploma should be linked to the CEFR, for it to have a real international effect.

Differentiation was also seen as a step away from the *culture of Cs* in which the highest possible achievement (an 'A') is not viewed as desirable, because a 'C' is good enough (Dutch: 'zesjes-cultuur'). The expressed reservations mainly referred to the idea that differentiated instruction would be far more effective when all teachers of all subjects in school were to invest together in methodology, structures, activities and materials needed to differentiate education.

**Offering choice in process/product** was unanimously seen as being good practice, with the prerequisite that you need to coach the students into having their individual choices reflect the most effective learning process. Choice might also be offered in the instruction by referring to online explanations (flip the classroom).

According to the panel **the implementation of a differentiated approach** relies heavily on the school environment. The panel was clear on the need to have the whole team involved from the start of the implementation process. All teachers need to have complete agreement on, and ownership of, the differentiated teaching approach with which the school is going to work and they all need to join in the effort to bring the approach about. Only then there is the safe environment needed in which teachers are willing to share their successes and failures and join wholeheartedly in work on a new development. To prevent the pitfalls of misconceptions, teachers need intensive training in differentiated teaching before implementation and should collaboratively develop the needed materials, assignments and activities, as this is viewed to be highly effective.

On the practical side of the implementation process, the panel indicated the need for frequent and consistent help, guidance and coaching. Structurally planned time for development, weekly progress meetings with the team and coaching on demand were seen as desirable support. Class size, classes timetabled together per subject and class formation, in which ability levels and percentage of students with behavioural problems are taken into account, would help. Just like digitally available materials, or even in class tablets/iPads.

The panel felt the list of **knowledge, skills and attitude** (future) teachers would need to have to effectively differentiate their teaching was concise and credible because it was co-constructed with English teaching colleagues. The lists were seen as mainly the task for teacher training and an arduous part of the life-long learning part of the job. Teachers need to have in-depth **knowledge** of:

- teaching strategies and methodological possibilities, including differentiation;
- the expectations of what the English language teaching curriculum is supposed to offer on the different secondary levels, per year as well as the ultimate goal and how these relate to the CEFR.

Teachers need to be to have the **skills** that allow them to:

- develop, change or find (online) additional assignments, materials, planners and activities;
- plan, organise, teach, instruct, coach and give feedback on different levels in one classroom;

- teach language learning strategies, study skills and how students can take responsibility for their language learning process;
- control classrooms (discipline and order) to offer the best learning experience possible in the current situation, while being aware of the moments they have to let go of planning because of the situation (students, colleagues, out-of-class incidents, current events, etc..) and take care of the group and its individuals before they can think about the subject again;
- train students in taking time to produce work and revise it to produce the best they can achieve;
- frequently and consistently offer theme-based, cross-curricular activities and speaking and writing activities;
- relate individual achievements and the quality of products to expected progress and possible achievements.

Teachers need to have a positive **attitude** towards:

- the different language skills levels in a classroom;
- changing assignments and activities to meet student needs and abilities;
- constructive communication with primary education;
- constant development and expansion of their methodological repertoire;
- keeping up with educational developments in general and language teaching in particular.

During the first two rounds a **discussion** arose around the implementation of differentiated language teaching. The divergence in arguments and experience appeared to be based on a general view on education. The 'culture of Cs' was said to be very present in Dutch education and society. Only a small minority of students is genuinely interested in personal development; the majority endures education and is more interested in the social interaction of school life beside the educational process.

The way English is taught at this moment does not help to move away from this culture as it has students grow accustomed to achieving through short cycle vocabulary and grammar tests which are not directly related to language skills. Within the current teaching approach our students are expected to perform within the bandwidth of the stream they are in and a lot of effort is given to underachievers who are

helped to climb towards the minimum required standard, leaving those that need no extra help to their own devices.

Most are of the opinion that the ideal situation would be to challenge students on their individual level and growth according to ability . The majority of secondary school teachers, however, use the current approach to (language) teaching and offer full frontal, whole class teaching which leads to the same course book based tests for everyone at the same time.

Although some panel members are slightly cynical about 'new' developments as offered in the described treatment, the general feeling is that it is important to let go of the course book. There is a desire to take responsibility for the educative process in the English language teaching classrooms instead of adhering to the teaching prescribed by a publisher. The individual student's needs and abilities are viewed as important enough to move away from current practice. In that sense the panel feels a drive for change. The publications, ideas and examples for a differentiated approach were known before schools joined in this experiment. Applying these principles and ideas to teaching activities, however, costs a lot of extra time and effort that the teachers just seem unable to put in. The current workload keeps them from wholeheartedly starting, joining or continuing experiments like the one at hand. They are heavily weighed down by the number of classes, number of students in classes, preparation, correction and other paperwork.

Apart from the pressure from the present workload, there is also a feeling that the investments needed for the treatment to succeed, i.e.: professional development, training in methodology, development of materials, structures and the time required to consistently implement the methodology, needs to be shared, at least school wide.

The more pedagogical part of teaching, such as giving feedback, cooperative learning, rubrics and differentiation, needs to be common ground for all teachers and students throughout the school. The shared responsibility, combined with students who are accustomed to the activities, would make a differentiated approach feasible.

Some of the panel members said that perhaps the time was right for differentiation. Apart from pressure from the ministry and inspectorate, the recent suggestion from the Dutch secondary education council<sup>41</sup> to offer students the opportunity to take final exams at different levels per subject was warmly welcomed by the panel. Differentiation of teaching in the first year of secondary education might then continue towards another level exam (VO-raad, 2015).

To conclude, two results from the final Delphi round need to be shared, although they are beyond the direct scope of this study. The first concerns the feedback received from the participating teachers concerning their experience with this research. All teachers unquestionably appreciated having had a voice in the studies they were involved in. The preliminary discussions on differentiation, as reported on in chapter 5, the contacts and support received during the experiment as reported on in chapter 6, as well as this Delphi study on the implementation of the treatment meant a lot to them. The collaboration in the treatment, as well as the participatory endeavour to determine the feasibility of the proposed educational reform, gave them the feeling of being partners in the research process.

The final result, which did not answer one of the research questions at hand, but is noticeable in itself, was a point made by one of the panel members: ***“Let’s keep in mind the children in the Dutch educational system are the happiest of the developed world.”***<sup>42</sup>

#### **7.4. Conclusion of the Delphi study**

When the contributions by the teachers in this Delphi study are viewed as part of their professional development (PD) and related to the structure of the six core features according to van Driel, et al. (2012), it is clear from the above that the treatment was consistent with their knowledge, beliefs and daily experienced problems (5). The teachers clearly appreciated the focus on classroom practice and research-based knowledge of teaching and learning (1) as well as being actively engaged in the treatment and/or its evaluation (2). The desire to collaborate with (more) colleagues and experts in a whole school movement

---

<sup>41</sup> Dutch: VO-raad.

<sup>42</sup> <http://www.unicef-irc.org/Report-Card-13>

towards differentiated instruction was very present (3). Apart from the remark that ‘perhaps the time was right for differentiation’ (p. 153) the panel also had something to say about the time needed for effective implementation. Although a substantial amount of time had been invested and the teachers had coaching meetings monthly (4), the panel unanimously agreed more time should be invested. The teachers needed more training before the initial implementation of the treatment and more time to cooperate and consult during the implementation of the treatment. To round off, the last core feature, which states that ‘school organisational conditions’ must be taken into account (6), was actually reversed by the teachers. The example given was the classroom with so many children and furniture that reorganising to allow cooperative activities was impossible. So, if differentiated instruction is imperative, something needs to happen to the classroom and preferably the number of children.

The final quote from one of the panel members in the previous chapter (7.3), refers to the UNICEF report card on the assessment of child well-being in 29 nations of the industrialized world (UNICEF, 2013; 2016).

The rankings in this report card are based on five dimensions of children’s lives: material well-being, health and safety, education, behaviours and risks, and housing and environment. The report card not only shows The Netherlands retains a position as the clear leader in these rankings, it also shows it is the only country ranked among the five countries in all dimensions of child well-being, as well as being the clear leader when the children themselves evaluate their well-being (95% of the children rate their own lives above the midpoint of the Life Satisfaction Scale (UNICEF, 2013, p. 39; 2016).

This positive finding by UNICEF was used as final quote for the chapter reporting on the results of the final Delphi round on the factors impeding differentiated language teaching. The quote was not intended as unanswerable remark, but is seen as a starting point. The findings of this study can be viewed as cautiously positive towards the possibilities and opportunities of differentiating language teaching. The teachers reported the need for more, wider and sustained support, but did not dismiss the treatment, its intentions and consequences. Based on the experiences of (as well as with) the Delphi panel and their input in the three rounds, as described above, this study might be viewed as a stepping stone. When the factors, considerations and circumstances listed in chapter 7.3 are taken into consideration, the positive results of the treatment as reported in chapter 6.4 might be built upon and expanded.

## **8. Conclusion and discussion**

This research was conducted to answer the main question: *'How and to what extent do English teaching in the final year of primary school and first year of secondary education differ from each other and what is the effect of differentiated English language teaching on the attitude towards learning English and the English language skills development of students in the first year of secondary education?'* For a concise answer to the main research question, six sub-questions were formulated and researched separately, as described in the chapters above. The answers to all these sub-questions will be discussed, as they accumulate into the final conclusion, below.

### **8.1. Sub-questions answered**

#### **1. *How has English language teaching in primary and secondary education in The Netherlands been introduced and how has it developed historically***

The brief historical contexts of English language teaching and differentiation in The Netherlands, as described in chapter 2 and 4 respectively, shows that ideas and approaches have not developed along clear lines. Some of the current research on approaches in education are even recurrences of publications between approximately forty years and almost a century old. When compared to changes in society in general, the change in English language teaching in secondary education in The Netherlands is slow and repetitive in the move towards ideal teaching practice. Especially the step from Grammar-Translation teaching towards the communicative approach has yet to be completed. The final (written) exams of secondary education in the Netherlands have been labelled by the government according to the CEFR communicative levels, and legislative requirements imposed on all students for passing the English exams, have made the position of the subject stronger, but that is where governmental influence ends. The English teaching offered on a practical level in The Netherlands does not (yet) reflect the can-do attitude towards language skills the CEFR is aiming for, as has been shown in the questionnaire research in chapter 3 on classroom activities.

Mandatory English teaching to children between ten and twelve years old in primary education in The Netherlands has been around for about 30 years and was founded on the communicative approach, with

an emphasis on speaking skills. Other activities besides listening, reading and writing, such as singing, (language) games and role-playing games are used to support language acquisition and speaking practice (Van Toorenburg & Bodde-Alderlieste, 2003; Bodde-Alderlieste, 2015).

These separate histories of English language teaching in two parts of one educational system, clarify the difference in approach. When primary education started to offer English, the aims, training and materials were developed on a well-defined approach. Secondary education was working toward the same approach, albeit with a history of centuries-old practice, principles and activities that sometimes withstood newly introduced educational approaches. Both primary and secondary education use the 'communicative approach', although the aims and content differ. Primary education has no standards or exams to work towards and aims to give all students in their mixed ability classrooms a positive attitude towards foreign language learning in general and some communicative skills in English. Secondary education assumes a shared general cognitive ability level in the streamed classrooms and has the task of offering students a language skill level for *all* skills in order for them to pass the final, school leaving, exams. Based on these differences *within* the approach to English language teaching, the level underneath the approach<sup>43</sup>, the design, practices and procedure logically differ greatly (Richards & Rodgers, 1986).

The concept of differentiating education has, as of yet, had no lasting effect on mainstream secondary education in The Netherlands despite available publications, research and good practice. In primary education in The Netherlands the teachers are accustomed to mixed ability classrooms for which teaching needs to be differentiated. English language teaching in primary schools capitalizes on the general pedagogical attitude of the teachers, the teaching methodology and materials based on the communicative approach.

In addition, the number of primary schools offering eight years of English has grown tremendously (European Platform, 2016).

---

<sup>43</sup> For more information see the paragraph on the difference between approach and method in chapter 1.

As of yet, differentiated instruction in secondary school English language teaching has not been part of mainstream practice and teacher training. Teachers of English in the first year of secondary education are confronted with an enormous variety in language skills levels, from 'unexposed to the English language' to 'years ahead' (Chapter 7), but do not have an answer to the situation. The combination of historical starting point within approach differences and the changes in English language teaching in primary education, explains the problematic articulation between levels.

## ***2. How does the teaching of English in the Netherlands differ between the final year of primary education and first year of secondary education?***

This study used a questionnaire survey to answer this question. Not unlike other questionnaire surveys, this study was hampered by a low response rate. In chapter three an extensive description is given to show what attempts have been made to compare the sample in the survey with the total population in The Netherlands. The following answer should thus be read bearing this in mind: the results only seemingly reflect the applied language teaching methodology in The Netherlands, due to the low response rate. The questionnaire survey, as described in chapter 3, showed that grammar as language teaching activity seems to play a pivotal role in English language teaching in secondary education in The Netherlands, as can also be found in (Hulshof, Kwakernaak, & Wilhelm, 2015). The importance is evident from the reported time spent on grammar as well as the choice of teaching activities, which does not relate to the developments language teaching has gone through in the past century, as discussed in chapter 2. Generally speaking, it seems clear that in primary education considerably more time is spent on speaking than in secondary education where, conversely, considerably more time is spent on grammar. Singing songs and playing language games are almost exclusively used in primary education. Conversely, teaching grammar, reading and testing play a dominant role in secondary education (cf. Chapter 3). These differences, as described above, indicate that more needs to be done to come to effective and constructive articulation between levels. To prevent the waste of capital invested in the language skills of students and to offer an English education that gives *all* students the opportunity to progress, certain things need to happen. First

and foremost, language teaching must be responsive to the mixed ability situation and offer differentiated instruction. In order to effectively offer differentiation, the articulation between levels needs to improve.

**3. Which aspects should a language teaching programme based on the principles of Differentiated Instruction for the first year of secondary school include, to ensure knowledge and skills acquired in primary school are used to their full potential and enable execution of activities on different levels?**

The literature found in the second literature review, as described in chapter 4, provided the foundation for the language teaching programme based on differentiated instruction developed for the treatments in the experiment of this research.

The main points found were summarised in a list of 33 elements (cf. Chapter 4), which generally speaking concerned the need to take individual differences of students into consideration when planning for the most effective language teaching possible. In order to challenge more students to full participation and offer more success experiences, students must be offered choice and influence in both the educational process and the product to help assess learning.

The list contained required teacher qualities, expectations of the students and elements for the programme, when considered part of the professional development of the teachers, it needed to have six core features (Van Driel et al., 2012). According to Van Driel et al., the programme should:

- focus on classroom practice and research-based knowledge of teaching and learning (1);
- actively engage teachers into their professional development related to daily practice (2);
- have teachers learn in collaboration with colleagues and experts (3);
- invest a substantial amount of time (timespan and actual hours), sustained by coaching (4);
- be in coherence with teachers' knowledge, beliefs and daily experienced problems as well as consistent with reforms and policies on school and national levels (5); and
- take 'school organisational conditions' into account (6).

For students to effectively become partners of the educational process, they must be (made) aware of the relationship between the day-to-day teaching of English they receive and the final goal for English of the

educational stream they are in. The dialogue between what 'needs' to happen according to the curriculum and what students need to have, do, experience and produce to reach the final goal or the intermediate steps of that curriculum, is how this partnership shapes the content and products of differentiation.

In that sense, student self-assessment of progress and achievement needed to play a role for learning gains to become relative to individual student capacity and commitment, for which challenge and scaffolding can be measured (Tomlinson & Allan, 2000).

For teachers it is important to know that students' feelings of safety and being appreciated as individuals are a prerequisite for effective learning. This can be made explicit by investing in what interests students, offering choice and the organisation of cooperative learning activities in flexible groups to allow structured moments for individual feedback on the work in progress.

Generally speaking, it is logical to build on the relationship between students' everyday life and the language learning activities as well as on the relationship between language learning activities and the skills and abilities students are working towards (Moskovsky et al., 2013).

The treatment should allow time for revision and discussions on progress with peers and the teacher, which implies the need for assignments that extend well beyond the lesson level. Products in the treatment, such as writing pieces and oral presentations, should be related to the ability level of the individual student and previous achievements. In that way the appreciation of the products can be fully differentiated (Tomlinson, 2014).

The above implies that grading should be avoided if possible, or that the grade should reflect the relationship between achievement and expected personal growth (as opposed to related to in-class performance). The emphasis needs to be on what students 'can do' with their English language skills.

Last but not least, research showed that success in this change of methodology towards differentiation, required a positive attitude towards the aims of differentiating instruction (Beecher & Sweeny, 2008), student-teacher relationships and high expectations for all students. In that sense, it is crucial for schools that would like to differentiate their (language) teaching to understand that differentiated instruction is not a technique or procedure. The whole package of theory and principles that have helped build this language teaching programme based on differentiated instruction clearly shows that differentiated instruction is an

approach (Richards and Rogers, 1986). The programme was made to fit within a functioning and running secondary school system and to that effect was made as non-intrusive, or undemanding, as possible. Schools that indisputably want to differentiate instruction, can use the treatment as a first step in the right direction. These schools should, however, be fully aware that many more steps need to follow to effectively differentiate their (language) teaching.

**4. To what extent does a language teaching programme based on differentiated instruction effectively increase the attitude towards learning English and increase the learning gains of first year students in secondary education?**

The instruments used to measure the Model of Planned Behaviour (Ajzen & Fishbein, 1980) in this quasi/experiment were reliable and valid. The treatment shows a positive effect on the change in Cognition, Affect and Intention, with and without controlling for covariates. The positive effect found from the treatment for Subjective Norm, however, can both be caused by the treatment as well as by the amount of Eibo received in primary education. The effect found from the treatment for Perceived Behavioural Control approaches significance. When correcting for the significant covariate 'age' and for Intention, the only significant covariate is 'the amount of time spent watching television'. After correcting for this covariate, the variable 'condition' is no longer significant.

The above means that, generally speaking, students who received the treatment had a more positive attitude towards learning English than those in the control group.

Although the treatment does not have a significant effect on changes in the scores of the vocabulary test, a significant effect is found of the treatment on the trend of dictation scores between pre- and post-test (with or without corrections for the significant covariates). After correction for significant covariates, a difference is found for the scores for correctly underlined words in the reading test, when only correcting for Cito-scores. In the case of the incorrectly underlined words score in the reading (editing) test, the treatment shows a positive effect after controlling for Cito-scores (with or without controlling for gender). The effect of the treatment on the Writing scores is again significant.

The above means that apart from the vocabulary test, the listening/spelling, reading and writing tests used in the pre- and post-test show measurable positive effects of the treatment offered in this experiment. All the information given above actually shows that, generally speaking, students who received the treatment had a more positive attitude towards learning English than those in the control group. In addition, students who received the treatment also had significantly better scores for the language skills tests than those in the control group. The interesting thing about these results is that all these positive effects were found as effect of the language teaching programme based on differentiated instruction offered, while offering the programme as intended was difficult and did not happen completely as intended.

**5. Which skills do (future) teachers of English in secondary schools need to possess or acquire to offer a language teaching programme based on differentiated instruction, as mentioned above?**

The direct answer to sub-question 5 is found in the following list produced by the panel of the Delphi panel, as described in chapter 6. The skills were mainly seen as the task for teacher training as well as an arduous part of the life-long learning part of the job. It should be noted that this list is the product of a panel, consisting of a small group of teachers (6) of whom 3 had participated in the experimental phase of this study. The list below should be read bearing in mind that it does not represent a summary of literature, nor does it represent the view of a sizable randomly selected group of teachers. It is, on the other hand, the feedback acquired from teachers involved in the experiment or working in the same schools.

**1. Teachers need to have in-depth knowledge of:**

- a. teaching strategies and methodological possibilities, including differentiation;
- b. the expectations of what the English language teaching curriculum is supposed to offer on the different secondary levels and who this relates to the CEFR per year.

**2. Teachers need to have the skills that allow them to:**

- a. develop, change or find (online) additional assignments, materials, planners and activities;
- b. plan, organise, teach, instruct, coach and give feedback on different levels in one classroom;

- c. teach language learning strategies, study skills and how students can take responsibility for their language learning process;
- d. control classrooms (discipline and order) to offer the best learning experience possible in the current situation, while being aware of the moments they have to let go of planning because of the situation (students, colleagues, out-of-class incidents, current events, etc..) and take care of the group and its individuals before they can think about the subject again;
- e. train students to take their time to produce work and revise it to produce the best they can achieve;
- f. frequently and consistently offer theme based, cross-curricular activities and speaking and writing activities;
- g. relate individual achievements and the quality of products to expected progress and possible achievements.

**3. Teachers need to have a positive attitude towards:**

- a. the different language skills levels in a classroom;
- b. changing assignments and activities to meet student needs and abilities;
- c. constructive communication with primary education;
- d. constant development and expansion of their methodological repertoire;
- e. keeping up with educational developments in general and language teaching in particular.

Apart from the above list to answer sub-question 5, the Delphi study was also used to discuss the articulation between levels, solutions and the treatment offered. After three rounds, the panel agreed on the problematic articulation between Dutch primary education and secondary education levels in general and for English language teaching in The Netherlands in particular.

The English language skills levels of students entering the first year of Dutch secondary education differ enormously (cf. Chapter 7), from 'unexposed' to 'years ahead of their peers' and these differences have

grown in the past decade, due to both Early English teaching as well as out-of-school exposure (Edelenbos, 1993; Herder & De Bot, 2005; Europees Platform, 2016).

Students cannot be labelled or streamed merely based on previously received English language teaching because of the wide variety in education of English language skills and language learning strategies students can obtain due to difference in approach and quality of teaching offered in primary schools. Secondary education, however, does not relate its English language teaching to what the students have experienced in their first eight years of education and does not differentiate accordingly. In that sense, some of the primary school methodology is viewed as desirable practice.

For teachers in secondary education it is important to let go of the course book and take responsibility for the educative process in the English language teaching classrooms.

Ideally more time could be spent on productive skills, like speaking; students with deficiencies should be helped more systematically and students should be offered more differentiated, theme-based, cross-curricular education so they can work more creatively with their knowledge and skills.

**A differentiated approach**, like the treatment as used in the experiment, was generally very well received. All teachers involved certainly appreciated the collaborative approach used in the implementation of the treatment, the subsequent interviews on the implementation and the attempt to discuss the feasibility of educational reform (Tomlinson, 1998; Robison, 2004).

Differentiation in general is viewed as an opportunity to step away from the 'culture of Cs', said to be prevalent in Dutch education and society. According to most, challenging students on their individual level and growth corresponding to ability would be the ideal situation.

Along those lines all participants, from both the experimental and control group, reacted very positively to the information gathered through the pre- and post-test used for the experiment. Apart from a need for a clear standard for English language teaching primary education, the idea is to use wider scoped tests formatively, for education and as part of the educational process (Tomlinson, 2005; Panadero, Tapia & Huertas, 2012).

**The components of the treatment** were each discussed separately:

**Rubrics** were problematic as these can only be used when the students can relate to them.

**Chapter-wide assignments** were well received, as involving the students personally in the assignments, process and product helps them to be more engaged.

**Cooperative learning** received mixed reactions based on varying experiences. In this experiment, the number of students in a classroom seemed to be a critical factor for success or failure.

The '**giving feedback**' requirement needed more in-school training.

**Differentiating the assessments** of products was something the panel became more positive about between the second and third round, although some reservations remained.

Offering **choice** in process/product was unanimously agreed upon as being good practice.

In general, although the panel holds positive views towards differentiation, they believe there are two recurrent issues that form a threshold.

In the first place, the current workload of teachers, caused by the number of classes taught, number of students in classes, preparation, correction and other paperwork is reported to weigh down heavily on them and prevents them experimenting. In one of the classrooms, for instance, the number of students made it literally impossible to reorganise the furniture for cooperative learning activities; this is not only problematic to cooperative learning activities, but was even found to negatively influence student achievement (Brühwiler & Blatchford, 2011).

Secondly, the teachers were very clear on the need for shared responsibility. They believe changes to the educational process can only be successful if the endeavour is shared by the school and its consequences accepted by the authorities.

At school level, activities such as the use of feedback, cooperative learning, rubrics and differentiation, need to be common ground for all teachers and students throughout the school, not just for the subject of English. The whole cycle of tests, assessments and (final) grades for English should be linked to the CEFR and related to individual abilities.

On a broader level, students should be offered the possibility to take final exams at different levels per subject, because the investment in differentiating education should not rest solely on the shoulders of the individual subject teacher. For a change in education, like the one offered as treatment in an experiment in

this research, teachers require a lot of support and time, both of which are unavailable at this moment and are rather costly to organise for schools.

## **8.2. General conclusion**

The first part of the general research question asked *how English teaching in the final year of primary school and first year of secondary education differed from each other*. The research presented in this dissertation shows that English teaching in the final year of primary school differs greatly from the English teaching offered in the first year of secondary education.

The difference in approach to English teaching is, to a certain extent, understandable from a historical point of view. Apparently it is impossible to really go beyond well-established principles and activities that were already challenged more than a century ago. Rooted in the tradition of grammar-translation, English language teaching has withstood the influence from individual reformers, cooperating teachers, theorists in 'movements' and even supranational bodies such as the Council of Europe.

Primary education had a fresh start some thirty years ago and has, from the start, based the applied English teaching methodology and activities on the communicative approach. Grammar only plays a role in the English teaching in primary schools when primary school teachers have not been properly trained. As they are generalists trained to teach all subjects, with pedagogical principles as guiding strand through the whole primary curriculum, the required safe environment and expected endeavour to offer all students a success experience is also part of their English language teaching.

Secondary school teachers are trained differently and work in a completely different situation. English teachers in secondary schools are specialists for whom knowledge of the language, phonology, syntax and the history of the English speaking world are central to their training. Grammar (teaching) is part and parcel of practically all course books used on all levels. Their classes are supposed to be to some extent homogeneous as far as intelligence and study aptitude is concerned. Success or failure would indicate the student was in the wrong stream of secondary education. The current growth of Early English language teaching in primary education is leading to more mixed ability groups in secondary schools and the need for differentiation is quite evident. The secondary school English teachers who participated in this research all

thought it was well worth investing time and effort in differentiation and they did so for more than a year. Even the teachers in the control groups, represented in the Delphi study, reacted positively to the described activities and principles.

The second part of the main research question dealt with **the effect of the treatment on the development of student attitude towards learning English and the English language skills of students in the first year of secondary education**. The treatment itself was influenced by two factors: fidelity and size. Overall, the privilege of cooperating with English teaching colleagues was great and a positive experience, despite the reported inability to completely implement the treatment as proposed, discussed and trained. The number of teachers involved was far greater at the start than reported in the research above, but due to unfortunate situations and circumstance the actual experiment was limited to three schools and six classes. Despite the small numbers and the partial implementation of the treatment, overall significant positive effects were found, as reported above, for language skills as well as attitude towards learning the language. These positive effects of the treatment are an indication of how powerful differentiated instruction really is. The treatment, as used in this study, does not offer fully differentiated language teaching, as it had been developed to be used within the already running secondary school system. So, when even partial implementation of (watered down) differentiated language teaching gives these kinds of positive results, fully and completely offering differentiated instructions might improve general educational achievements. The changes in the classrooms, as generated by this research, the cooperative approach to the implementation of differentiated language teaching, enhance the position and role of the student in the educational process. Through the pre-test information, the teachers learned more about their students' language skills levels, interest and disposition towards English language learning than ever before. The teachers used all the personal, individual information in the initial introduction of the rubric, through which they asked students to self-assess their language skills. These two views on what students were capable of came together in the choices offered with the first chapter-wide assignments, the subsequent teacher-student discussions on what the options really were and especially the assessment of end-of-chapter products.

For the students in the experimental group the experience of the treatment in which they, their abilities and their preferences play an important role, must have been the positive factor that generated the significant positive effects reported: an experience that should be offered to *all* students.

### **8.3 Discussion**

The study, as reported upon above, resulted in an experiment with a collaboratively designed teaching programme, for which positive results have been found. Despite the use of core features of a professional development programme, as given by Driel, et al. (2012), and the involvement of the teachers in the development of the treatment, the role of professional development in this study is below the mark in the opinion of some researchers. Desimone (2009) proposes a unified conceptual framework for the study of teachers' professional development. While that may enhance the opportunities for setting up comparative and complementary programmes, it seems to be a bit optimistic about the feasibility of such an approach. The variation in perspectives of different teachers, even in a restricted study like the one reported on here, makes the implementation of a unified framework at least problematic.

The fact that this study is performed by one person limits the possibilities, at least on a physical scale.

Research offers a myriad of possibilities and alternatives. Bereiter (2014), for instance, proposes the use of 'principled practical knowledge' (principled know-how and know-why) and to limit the search for knowledge to the point where it solves problems. Jansen et al (2015) reacted to Bereiter's ideas on PPK. In their analysis they conclude that PPK, as portrayed by Bereiter, "does not offer much practical guidance for 2 potential users: professional designers and teachers.... Even an enriched form of PPK still does not suffice to address the challenging issues of practicality teachers face." The authors explain the magnitude and dimensions that underlie practicality in the everyday work of teachers and suggest how recent work on fast and frugal heuristics can contribute to helping teachers to make instructional innovations practical. These suggestions are certainly relevant for the present study and its further development, since the challenges of the development and implementation of even a limited programme proved to be considerable.

Along the same line lies the use of implementation research in a small scale study like the one at hand. In their overview of the literature on the implementation of evidence-based programmes in education, Fixsen

et al. (2005) point out the need to make the implementation of an innovation a topic of research by itself. They call for 'applied research to better understand service delivery processes and contextual factors to improve the efficiency and effectiveness of program implementation at local, state, and national levels.' In their conclusion they quote Petersilia (1990) who concludes that, "The ideas embodied in innovative social programs are not self-executing." (126) Instead, what is needed is an "implementation perspective on innovation--an approach that views post adoption events as crucial and focuses on the actions of those who convert it into practice as the key to success or failure". In the present dissertation, the development of the new approach has been the focus, but it is obvious that future research on the type of innovation presented in this dissertation should take this embedding into account.

## 9. English and Dutch Summary

### Summary

The teaching of English in primary education, compulsory in the Netherlands since 1986, varies greatly in how it is offered, ranging from a few lessons in group 8 to consistent English language teaching throughout the primary school (group 1 to 8). In addition, there is substantial difference in the quantity of teaching time offered, the use of teaching materials and the use of specialists such as native speakers. Secondary education does not take into account the significantly different English language skills levels of learners who enter from primary education. To examine whether the articulation between levels can be improved by providing differentiated language education in the first year of secondary education, a quasi-experimental study was carried out. After the introduction in the first chapter, the second chapter of the thesis describes the history of teaching English in the Dutch context. In chapter three a cross-sectional survey among teachers of primary and secondary education is used to describe the difference in approach to teaching English between primary and secondary education. The results of this survey, as described in Chapter 3, appear to confirm the expectations regarding the differences between teaching English in primary and secondary education, as found in earlier studies. In chapter four a literature review on the various forms of differentiated education is presented to be used as the foundation for the development of an intervention that provides a differentiated approach to English language teaching in the first year of secondary education. The developed intervention was tested through a quasi-experimental study among 68 students in the experimental group and 59 in the control group. The pre-test took place at the beginning of the first year of secondary education and the post-test was held at the end of this first year. The studied dependent variables concern the attitude towards learning for the subject of English and various aspects of English language skills. The attitude measurement is based on the Model of Planned Behaviour of Ajzen and Fishbein (1980) and contains, besides the study-behaviour itself, all constructs included in the model: Cognition, Affect, Subjective Norm, Perceived Behavioural Control and Intent. The measured aspects of English proficiency related vocabulary and the skills of listening, spelling, reading and writing. The results show that, for almost all components, the treatment of the MPB shows a significant positive effect, both with and without

controlling for significant covariates. The only times the effect of the treatment was not significant, were for Perceived Behavioural Control and Subjective Norm, when controlling for significant covariates. Without controlling for covariates, these two constructs also showed significant positive effects. On growth in the measured aspects of language skills, if not adjusted for significant covariates, a positive significant effect of the treatment was found for all the measured aspects of language skills, except for vocabulary. After adjusting for covariates, the effect of the treatment turned out to be no longer significant for Perceived Behavioural Control and Subjective Norm. The conclusion is that the developed treatment for providing differentiated teaching of English has positive effects on both the development of language skills and the attitude components of the Model of Planned Behaviour. In conclusion, a Delphi study was carried out among the participating teachers to ascertain which aspects of the treatment can be improved in order to optimize future implementation.

#### Samenvatting

Het onderwijs Engels in de basisschool, verplicht in Nederland sinds 1986, krijgt erg diverse vorm, variërend van enkele lessen in het laatste jaar (groep 8) tot consistent onderwijs Engels in de gehele basisschool (groep 1 tot en met 8). Daarnaast is er veel verschil in hoeveelheid geboden lestijd, het gebruik van lesmaterialen en de inzet van specialisten zoals native speakers. In het voortgezet onderwijs wordt geen rekening gehouden met het zeer diverse niveau Engels van de leerlingen die uit het basisonderwijs instromen. Om na te gaan of door het geven van gedifferentieerd taalonderwijs in het eerste jaar van het voortgezet onderwijs de aansluiting verbeterd kan worden, is een quasi experimentele studie verricht. Na de introductie in het eerste hoofdstuk, wordt in het tweede hoofdstuk van het proefschrift de geschiedenis van het onderwijs Engels in de Nederlandse context beschreven. In hoofdstuk drie wordt aan de hand van een cross-sectionele peiling onder leerkrachten van basis- en voortgezet onderwijs het verschil in aanpak van het onderwijs Engels tussen basis- en voortgezet onderwijs in kaart gebracht. Uit de resultaten van deze peiling, die worden beschreven in hoofdstuk drie, blijken de verwachtingen aangaande de verschillen tussen het onderwijs Engels in basis- en voortgezet onderwijs, zoals die naar voren komen uit eerdere studies, bevestigd te worden. In

hoofdstuk vier wordt een literatuurstudie gepresenteerd naar de onderscheiden vormen van gedifferentieerd onderwijs om op basis hiervan een interventie te ontwikkelen die voorziet in een gedifferentieerde aanpak van het onderwijs Engels in het eerste jaar van het voortgezet onderwijs. De ontwikkelde interventie is getoetst aan de hand van een quasi-experimentele studie onder 68 leerlingen in de experimentele en 59 in de controlegroep. De voormeting vond plaats aan het begin van het eerste jaar van het voortgezet onderwijs en de nameting is aan het eind van dit eerste jaar gehouden. De onderzochte afhankelijke variabelen betreffen de attitude ten aanzien van het leren voor het vak Engels en onderscheiden aspecten van de Engelse taalvaardigheid. De attitudemeting is gebaseerd op het Model of Planned Behaviour van Ajzen en Fishbein (1980) en bevatte, behalve het studeergedrag zelf, alle in het model opgenomen constructen: de cognitie, het affect, de subjectieve norm, de Perceived Behavioral Control en de intentie. De gemeten aspecten van taalvaardigheid Engels betroffen vocabulaire en de luister-, spelling-, lees- en schrijfvaardigheid. De uitkomsten laten zien dat de treatment voor bijna alle componenten van het MPB een significant positief effect laat zien, zowel zonder als met controle voor significante covariaten. De enige keren dat het effect van de treatment niet significant bleek, was voor Perceived Behavioral Control en voor Subjectieve Norm indien gecontroleerd werd voor significante covariaten. Zonder controle waren ook voor deze twee constructen de effecten significant positief. Op groei in de gemeten aspecten van taalvaardigheid werden indien niet gecorrigeerd voor significante covariaten voor alle gemeten taalaspecten een positief significant effect van de treatment gevonden, behalve voor vocabulaire. Na correctie voor covariaten bleek het effect van de treatment niet meer significant voor Perceived Behavioral Control en voor Subjectieve Norm. De conclusie luidt dat de ontwikkelde treatment voor het geven van gedifferentieerd onderwijs Engels positieve effecten heeft op zowel de taalontwikkeling als op de attitudecomponenten van het Model of Planned Behavior. Tot slot is nog aan de hand van een Delphi procedure onder de deelnemende docenten nagegaan welke aspecten van de treatment nog verbeterd kunnen worden om toekomstige implementatie te kunnen optimaliseren.

## 10. References

- Adamson, B. (2008). 24 Fashions in Language Teaching Methodology. in Davies, A., & Elder, C. (Eds.) *Handbook of applied linguistics*. (pp. 604-622). Oxford: Blackwell.
- Adler, M., & Ziglio, E. (Eds.) (1996). *Gazing into the oracle: The Delphi method and its application to social policy and public health*. London: Jessica Kingsley Publishers.
- Affholder, L. P. (2003). *Differentiated instruction in inclusive elementary classrooms*. (Unpublished EdD thesis). Kansas: University of Kansas.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behaviour. *Journal of Applied Social Psychology, 32* (4), 665-683.
- Ajzen, I. (1988). *Attitudes, personality and behaviour*. Milton Keynes, UK: Open University Press.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes, 50*(2), 179-211.
- Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behaviour*. Englewood Cliffs, NJ: Prentice Hall.
- Alexander, P. A., Kulikowich, J. M., & Jetton, T. L. (1994). The role of subject-matter knowledge and interest in the processing of linear and nonlinear texts. *Review of Educational Research, 64*(2), 201-252.
- Algozzine, B., & Anderson, K. M. (2007). Tips for teaching: Differentiating instruction to include all students. *Preventing School Failure: Alternative Education for Children and Youth, 51*(3), 49-54.
- Amabile, T. M. (1996). *Creativity in context: Update to "the social psychology of creativity."*. Westview press. Boulder, Colorado.
- Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York, NY: W. H. Freeman,.
- Anderson, K. M., (2007). Differentiating instruction to include all students. *Preventing School Failure, 51*(3), 49-54
- Ayres, L. P. (1912). Measuring educational processes through educational results. *The School Review, 300-309*.

- Baker, J. A., Lovell, K. & Campbell, M. (2007) Multidisciplinary consensus of best practice for pro re nata (PRN) psychotropic medications within acute mental health settings: a Delphi study. *Journal of Psychiatric and Mental Health Nursing*, 14, 478–484.
- Barsalou, L. (1992). *Cognitive psychology: An overview for cognitive scientists*. Hillsdale, NJ: Erlbaum.
- Beecher, M., & Sweeny, S. M. (2008). Closing the achievement gap with curriculum enrichment and differentiation: One school's story. *Journal of Advanced Academics*, 19(3), 502-530.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological bulletin*, 112(3), 400.
- Berliner, D. C. (1993). The 100-year journey of educational psychology. In Fagan, T. K., & VandenBos, G. R. (Eds.). *Exploring applied psychology: Origins and critical analyses*. (pp. 37-78). Washington, DC, APA
- Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: School of Education, King's College.
- Blok, H. (2004). Adaptief onderwijs: Betekenis en effectiviteit. *Pedagogische Studiën*, 81(1), 5-27.
- Bodde-Alderlieste, M. (2015). Ambities voor vroeg Engels Niet het lespakket bepaalt de leerlijn, maar het beoogde ambitieniveau. *Levende Talen Magazine*, 102(6), 22-26.
- Bodde-Alderlieste, M., & Schokkenbroek, J. (2013). Scenario's voor Engels in het primair onderwijs: Gevolgen voor de brugklas. *Levende Talen Magazine*, 100 (7), 20-24.
- Boekholt, P. T. F., & de Booy, E. P. (1987). *Geschiedenis van de school in Nederland vanaf de middeleeuwen tot aan de huidige tijd*. Assen: Van Gorcum.
- Boer, M. de (2003). *English... a matter of concern*. Utrecht, Retrieved 29-05-2016 from: [home.planet.nl/~bodde009/pages/publipab.htm](http://home.planet.nl/~bodde009/pages/publipab.htm)
- Boer, G.C. de, Minnaert, A.E., & Kamphof, G. (2013). Gifted Education in the Netherlands. *Journal for the Education of the Gifted*, 36(1), 133-150.
- Bond, R., & Castagnera, E. (2006). Peer supports and inclusive education. *Theory Into Practice*, 45, 224-229.
- Bonnet, G. (2004). *The assessment of pupils' skills in English in eight European countries 2002*. Paris: European Network of policy makers for the evaluation of education systems.

- Bosker, R.J.,(2005) *De grenzen van gedifferentieerd onderwijs*, Rede uitgesproken bij de aanvaarding van het ambt van hoogleraar theoretische en vergelijkende onderwijskunde en het onderwijsonderzoek aan de Faculteit der Psychologische, Pedagogische en Sociologische Wetenschappen van de Rijksuniversiteit Groningen op 11 januari 2005. Retrieved 29-05-2016 from: [redes.eldoc.ub.rug.nl/FILES/root/2005/r.j.bosker/bosker.pdf](http://redes.eldoc.ub.rug.nl/FILES/root/2005/r.j.bosker/bosker.pdf)
- Bowen, J. D. (1978). The identification of irrelevant lexical distraction. An editing task. *TESL Reporter*, 12(1), 1-3.
- Branden, K. van den (1997). Effects of negotiation on language learners' output. *Language learning*, 47(4), 589-636.
- Bransford, J.D., Brown, A.L., & Cocking, R. (1999) *How people learn, Brain, Mind, Experience, and School*, National Research Council, Washington, DC: National Academy Press.
- Brinton, D. M., Snow, M. A., & Wesche, M. B. (1989). *Content Based Second Language Instruction*. New York: Newbury House.
- Brosvic, G. M., Epstein, M. L., Dihoff, R. E., & Cook, M. L. (2006). Retention of esperanto is affected by delay-interval task and item closure: A partial resolution of the delay-retention effect. *The Psychological Record*, 56(4), 597-615.
- Brown, H. D. (2002). English language teaching in the "post-method" Era: toward Better Diagnosis, Treatment, and Assessment. In J.C. Richards & W.A. Renandya (Eds.), *Methodology in Language Teaching: An Anthology of Current Practice*. (pp. 9-19). Cambridge: CUP.
- Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, 21(1), 95-108.
- Bryman, A. (2012), *Social Research Methods*, Fourth edition, Oxford University Press, New York.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT journal*, 10(1), 15-42.
- Bulterman-Bos, J. (2007). " Bij de groep horen": hoe de imitatietendens het leraren moeilijk kan maken om adaptief onderwijs vorm te geven. *Pedagogische Studiën*, 84(3), 159.

- Burmeister, P. & Piske, T. (2008). Schriftlichkeit im fremdsprachlichen Sachfachunterricht an der Grundschule. In Böttger, H. (Hrsg.), *Fortschritte im frühen Fremdsprachenlernen. Ausgewählte Tagungsbeiträge Nürnberg 2007* (pp. 183-193). München: Domino Verlag.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied linguistics*, 1(1), 1-47.
- Canale, M. (1983) From communicative competence to communicative language pedagogy. In Richards, J.C. & Schmidt, R.W. (Eds.) *Language and Communication*. (pp. 2–27). London, Longman.
- Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. In Soler, E. A., & Jordà, P. S. (Eds.). *Intercultural language use and language learning*. (pp. 41-57). Dordrecht: Springer.
- Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code mixing. Age and the acquisition of English as a foreign language, 77-93. In Mayo, M. D. P. G., & Lecumberri, M. L. G. (Eds.) *Age and the acquisition of English as a foreign language 4* (pp. 77-93). Clevedon: Multilingual Matters.
- Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of learning disabilities*, 35(5), 386-406.
- Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language*, 35(1), 26-58.
- Conda, A., de Kraay, T., & Feuerstake, M. (2012). *And yet all different: handreiking voor een curriculum Engels op de pabo*. Leiden: Expertisecentrum MVT.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American psychologist*, 12(11), 671.
- Cronbach, L. J., & Snow, R. E. (1969). Individual Differences in Learning Ability as a Function of Instructional Variables. Stanford, CA: Stanford University Press.
- Csikszentmihalyi, M. Rathunde, K., & Whalen, S. (1993). *Talented teenagers: The roots of success and failure*. Cambridge, UK: CUP.
- Dalkey, N. & Helmer, O. (1963) Experimental Application of the Delphi Method to the Use of Experts. *Management Science*, Vol. 9, (3) 458-467.

- Damasio, A. R. (2000). *The feeling of what happens: Body, emotion and the making of consciousness*. London: Random House.
- Dasberg, L., & Jansing, J. W. G. (1978). *Meer kennis, meer kans: het Nederlandse onderwijs 1843-1914*. Haarlem: Fibula-Van Dishoeck.
- Davies, A. (1975) Two Tests of Speeded Readings. In Jones, R. L., & Spolsky, B. (pp.119-130) *Testing Language Proficiency*. Arlington, VA: Center for Applied Linguistics.
- Davies, J., & Brember, I. (2001). The closing gap in attitudes between boys and girls: A 5-year longitudinal study. *Educational Psychology, 21*(1), 103-114.
- DeKeyser, R. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *Modern Language Journal, 77*, 501-514.
- Deunk, M. I., Doolaard, S., Smalle-Jacobse, A., & Bosker, R. J. (2015). *Differentiation within and across classrooms: A systematic review of studies into the cognitive effects of differentiation practices*. GION onderwijs/onderzoek, Rijksuniversiteit Groningen.
- Dewaele, J.-M. (2012) Learner internal psychological factors. In Herschensohn, J. & Young-Scholten, M. (Eds.) *The Cambridge Handbook of Second Language Acquisition*, 159-179. Cambridge, UK: Cambridge University
- Dewaele, J. M., & Al-Saraj, T. M. (2015). Foreign Language Classroom Anxiety of Arab learners of English The effect of personality, linguistic and sociobiographical variables. *Studies in Second Language Learning and Teaching 5* (2). 205-228
- Dickinson, L. (1995). Autonomy and motivation a literature review. *System, 23*(2), 165-174.
- Dion, E., Morgan, P. L., Fuchs, D., & Fuchs, L. S. (2004). The promise and limitations of reading instruction in the mainstream: The need for a multilevel approach. *Exceptionality, 12*(3), 163-173.
- Dörnyei, Z. (2005). *The psychology of the language learner*. London, UK: Erlbaum.
- 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.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. NY, NY: Routledge.

- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test* (third edition). Circle Pines, MN: American Guidance Service.
- EarlyBird (2016) Retrieved 29-05-2016 from: [www.EarlyBirdie.nl](http://www.EarlyBirdie.nl)
- Echevarria, J. & Short, D. J. (2000). *Using multiple perspectives in observations of diverse classrooms: The Sheltered Instruction Observation Protocol (SIOP)*. Paper presented at the annual meeting of the American educational research association. ERIC (ED 441 334).
- Edelenbos, P. (1993). *De aansluiting tussen Engels in het basisonderwijs en Engels in het voortgezet onderwijs. Eindrapport* (SVO-project nr. 0015). Groningen: RION
- Edelenbos, P., Schoot, F. & Verstralen, H. (2000). *Balans van het Engels aan het einde van de basisschool 2*. Periodieke Peiling van het Onderwijs Niveau, Rapport 17. Arnhem: Cito.
- Education (2014) In Encyclopædia Britannica. Retrieved 29-05-2016 from <http://www.britannica.com/EBchecked/topic/179408/education/47649/Progressive-education>
- Eijck, K. van, & Graaf, P.M. de (2004). The big five at school. *The Netherlands Journal of Social Sciences*, 40
- Ek, J. A. van (1975). Systems Development in Adult Language Learning: The Threshold Level in a European Unit/Credit System for Modern Language Learning by Adults. Strasbourg: Council of Europe
- Ek, J.A. van & Alexander L.(1980) *Threshold level English in a European Unit/Credit System for Modern Language Learning by Adults*. Oxford: Pergamon.
- Ellis, R. (1993). The structural syllabus and second language acquisition. *Tesol Quarterly*, 27(1), 91-113.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 525–551). Oxford: Blackwell.
- Ellis, E. S. E., & Worthington, L. A. (1994). *Research synthesis on effective teaching principles and the design of quality tools for educators (Technical Report No. 5)*. Eugene: University of Oregon. NCITE.
- Els, T.J.M. van (1990). *Horizon Taal. Nationaal actieprogramma moderne vreemde talen*. Zoetermeer: Ministerie van OC&W.

- Els, T.J.M. van, (1992) Foreign language teaching in the Netherlands, 1880-1940: an outline of methodological developments, in Anick Giroud (éd.). *Aspects de l'histoire de l'enseignement des langues: 1880-1940*. Neuchâtel: Bulletin CILA 56, 35-47.
- Engel, G., Trimbos, B., Drew, I., Groot-Wilken, B. (2007). CIDREE REPORT – *English in Primary Education in The Netherlands, North-Rhine Westphalia(NRW) and Norway*. Aurau: CIDREE.
- Europees Platform. (2016). Overzicht basisscholen met vroeg vreemdetalenonderwijs (VVTO)  
[http://www.talenopdebasisschool.nl/downloads/adressenlijst\\_VVTO.pdf](http://www.talenopdebasisschool.nl/downloads/adressenlijst_VVTO.pdf)
- Europees Platform (2010). *Nieuwsbrief Vroeg Vreemdetalenonderwijs*, 7, nr. 6. Retrieved 10-09-2010 from  
<https://www.epnuffic.nl/primair-onderwijs/talenonderwijs/vroeg-vreemdetalenonderwijs-vvto,.>
- Fallan, L. (2006). Quality reform: Personality type, preferred learning style and majors in a business school. *Quality in Higher Education*, 12(2), 193-206.
- Field, A. (2009). *Discovering statistics using SPSS*. Thousand Oaks, CA: Sage publications.
- Figueras, N. (2012). The impact of the CEFR. *ELT journal*, 66(4), 477-485.
- Filene, P.G. (1970) An Obituary for "The Progressive Movement", *American Quarterly* , 22 (1) (Spring, 1970), 20-34
- Fretschner, M. (2014). Ajzen's Theory of Planned Behavior in Entrepreneurship Education Research: An Introduction and Review of Impact Studies. In Weber, S., Oser, F., Achtenhagen, F., Fretschner, M., & Trost, S. (Eds.). *Becoming an Entrepreneur* (pp. 249-277). Dordrecht: Sense Publishers, 249-277.
- Fuchs, D., Fuchs, L. S., Al Otaiba, S., Thompson, A., Yen, L., McMaster, K. N., Svenson, E., & Yang, N. J. (2001). K-PALS: Helping kindergartners with reading readiness: Teachers and researchers in partnerships. *Teaching Exceptional Children*, 33, 78-80.
- Fulk, B., & Montgomery-Grymes, D. J. (1994). Strategies to improve student motivation: Description of the causes of poor students' motivation and practical strategies for improvement. *Intervention in School and Clinic*, 30(1), 28-33.
- Gabrielatos, C. (2002) Shopping at the ELT Supermarket: Principled Decisions and Practices. *English Language Teaching News* 144.

- Garcia Mayo, M. P., & Pica, T. (2000). Is the EFL environment a language learning environment? *Working Papers in Educational Linguistics*, 16(1), 1-24.
- Gelpi, E. & Bowen, J. (2014). European education in the 17th and 18th centuries, in *Encyclopædia Britannica*. Retrieved 29-05-2016 from:  
<http://www.britannica.com/EBchecked/topic/179408/education/47552/The-Calvinist-Reformation>
- George, P.S.(2005) A Rationale for Differentiating Instruction in the Regular Classroom. *Theory Into Practice*, 44: 3, 185 — 193.
- Geurts, B., & Hemker, B. T. (2013). *Balans van het Engels aan het einde van de basisschool 4: uitkomsten van de vierde peiling in 2012*. Arnhem: Cito.
- Gibson, E.L.(2005) *Addressing the needs of diverse learners through differentiated instruction*. Humboldt State University. Retrieved 29-05-2016 from: <http://humboldt-dspace.calstate.edu/handle/2148/527>
- Gras, R. R. (1967). *Studietoetsen voor moderne talen: een onderzoek naar hun bruikbaarheid ter bepaling van receptief-schriftelijke taalbeheersing van het Frans, Duits en Engels aan de HBS-B*, JB Groningen : Wolters-Noordhoff
- Graefe, A., & Armstrong, J. S. (2011). Comparing face-to-face meetings, nominal groups, Delphi and prediction markets on an estimation task. *International Journal of Forecasting*, 27(1), 183-195.
- Gremmo, M. J., & Riley, P. (1995). Autonomy, self-direction and self-access in language teaching and learning: The history of an idea. *System*, 23(2), 151-164.
- Grittner, F.M. (1975) Individualized Instruction: An Historical Perspective, *The Modern Language Journal* Volume 59, Issue 7, 323–406.
- Guild, P.B., & Garger, S. (1998). *What Is Differentiated Instruction? Marching to Different Drummers*. 2nd Ed. Alexandria VA: ASCD
- Gupta & Clarke 1996 Theory and Applications of the Delphi Technique - a bibliography 1975-1994. *Technological Forecasting and Social Change* 53, 185-211

- Guskey, T. R. (2005). *Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications*. Paper presented the annual meeting of the American Educational Research Association, April 11 – 15, Montreal: AERA
- Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks: Corwin Press.
- Hall, T., Vue, G., Strangman, N. & Meyer, A. (2002). *Differentiated instruction and Implications for UDL Implementation - Effective Classroom Practices Report*, (NCAC), Washington, DC: CAST & OSEP,
- Halpern, D. F., & Hakel, M. D. (2002). Learning That Lasts a Lifetime: Teaching for Long-Term Retention and Transfer. *New Directions for Teaching and Learning*, 2002(89), 3-7.
- Handschin, C. (1919). Individual differences and supervised study: Adapting work to individual differences. *Modern Language Journal*, 3, 158-173.
- Hasselgreen, A. (Ed.). (2012). *Assessment of young learner literacy linked to the Common European Framework of Reference for Languages*. Strasbourg: Council of Europe.
- Hasson, F. & Keeney, S. (2011) Enhancing rigour in the Delphi technique research. *Technological Forecasting & Social Change*, 78. 1695–1704.
- Hattie, J. (2009). *Visible Learning*. Abingdon: Routledge.
- Heesters, K., Feddema, M., Schoot, F. van der, Hemker, B. (2008) *Balans van het Engels aan het einde van de basisschool 3*. Uitkomsten van de peiling in 2006. Arnhem: Cito
- Herder, A. & Bot, K. de, (2005). *Vroeg vreemdetalenonderwijs in internationaal perspectief*. Literatuurstudie in opdracht van het Europees platform. Rijksuniversiteit Groningen
- Hermans-Nymark, L.D. (2006). *English in the EFL classroom: Why not? Classroom discourse patterns and teachers' beliefs*. Dissertatie. Nijmegen: Radboud Universiteit.
- History of Education (1992) In the New Encyclopædia Britannica, Vol. 18, pp. 44-57, Chicago: Encyclopedia Britannica.
- Holdinga, L. (2007) *Van Engels in het Basisonderwijs naar Engels in het Voorgezet Onderwijs: Een onderzoek naar mogelijke verbetering van de aansluiting*, Unpublished Master's Thesis, Utrecht: Utrecht University.

- Holec, H. (1981). *Autonomy in foreign language learning*. Oxford: Pergamon. (First published 1979, Strasbourg: Council of Europe).
- Horwitz, E. (2000) Teachers and students, students and teachers- an ever-evolving partnership. *The Modern Language Journal*, 84, 523–35.
- Howatt, A. P. R. (1982). Language teaching must start afresh!. *ELT Journal*, 36(4), 263-268.
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. Second Edition. New York, NY: Routledge.
- Huebner, T.A. (2010) Differentiated Learning. *Educational Leadership Magazine*, Volume 67, (5) 79-81, Alexandria, VA: ACSD.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.
- Hughes, A. (2003). *Testing for language teachers* (second edition). Cambridge: Cambridge University Press.
- Huibregtse, I. (2001). *Effecten en didactiek van tweetalig voortgezet onderwijs in Nederland*. PhD Dissertation, University of Utrecht.
- Hulshof, H., Kwakernaak, E., & Wilhelm, F. (2015). *Geschiedenis van het talenonderwijs in Nederland: Onderwijs in de moderne talen van 1500 tot heden*. Passage: Groningen
- Hymes, D. (1972). Models of the Interaction of Language and Social Life. In Gumperz J. J. & Hymes, D. (Eds.) *Directions in Sociolinguistics: The Ethnography of Communication* (pp. 35-71). New York: Holt, Rinehart and Winston.
- Imelman, J. D., & Meijer, W. A. J. (1986). *De nieuwe school, gisteren & vandaag*. Amsterdam/Brussel: Elsevier.
- Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, brain, and education*, 1(1), 3-10.
- Inspectie van het Onderwijs (1991). *Engels in basisonderwijs en brugklas*. Zoetermeer: OC&W.
- Inspectie van het Onderwijs (1999). *Engels in de basisvorming: Evaluatie van de eerste vijf jaar*, Den Haag: SDU.
- Inspectie van het Onderwijs (2006), *Onderwijsverslag 2004/2005*, De Meern: IvO.

- Johnsen, S. (2003). Adapting instruction with heterogenous groups. *Gifted Child Today*, 26(3), 5-6.
- Kasper, L. F. (1997). The impact of content-based instructional programs on the academic progress of ESL students. *English for Specific Purposes*, 16(4), 309-320.
- Kedde, J. S. (2004). The CEF and the secondary school syllabus. In Morrow, K. (Ed.) *Insights from the common European framework*. (pp. 43-54). Oxford, UK: OUP.
- Knoll, M. (2009). From Kidd to Dewey: the origin and meaning of 'Social efficiency'. *Journal of Curriculum Studies*, 41(3), 361-391.
- Koeze, P. A., (2007) *Differentiated Instruction: The Effect on Student Achievement in an Elementary School*. Paper 31. (Retrieved 29-05-2016 from: <http://commons.emich.edu/theses/31University>)
- Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, 51(4), 472-477.
- Koning, P. de (1973). *Interne Differentiatie*. Amsterdam: APS/RITP.
- Kormos, J., Kiddle, T., & Csizér, K. (2011). Systems of goals, attitudes, and self-related beliefs in second-language-learning motivation. *Applied Linguistics*, 32(5), 495-516.
- Koster, C. J. (1986). English FLES in The Netherlands: How Good Must a Teacher Be?. *The modern language journal*, 70(1), 8-12.
- Kramsch, C. (2009). *The Multilingual subject*. Oxford: Oxford University Press.
- Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning, and development. In Renninger, K. A. Hidi, S. & Krapp A. (Eds.), *The role of interest in learning and development* (pp. 3-25). Hillsdale, NJ: Erlbaum.
- Krol, K., Janssen, J., Veenman, S., & van der Linden, J. (2004). Effects of a cooperative learning program on the elaborations of students working in dyads. *Educational Research and Evaluation*, 10(3), 205-237.
- Kumaravadivelu, B. (2001). Toward a postmethod pedagogy. *Tesol Quarterly*, 35(4), 537-560.
- Fiske, E. B., & Ladd, H. F. (2010). The Dutch Experience with Weighted Student Funding. *Phi Delta Kappan*, 92(1), 49-53.
- Lai, E. R. (2011). Motivation: A literature review. *Pearson's Research Report*, 2, 4-5.

- Landrum, T.J., & McDuffie, K.A. (2010): Learning Styles in the Age of Differentiated Instruction, *Exceptionality: A Special Education Journal*, 18:1, 2010, 6-17, NY, NY: Routledge.
- Lantolf, J., & Pavlenko, A. (1995). Sociocultural theory and second language acquisition. *Annual Review of Applied Linguistics*, 15, 108-124.
- Larsen-Freeman, D., & Freeman, D. (2008). Language moves: The place of “foreign” languages in classroom teaching and learning. *Review of Research in Education*, 32(1), 147-186.
- Lasagabaster, D., & Sierra, J. M. (2009). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal*, 1(2), 4-17.
- Latz, A.O., Speirs Neumeister, K. L., Adams, C. M., & Pierce, R. L. (2008). Peer coaching to improve classroom differentiation: Perspectives from Project CLUE. *Roeper Review*, 31(1), 27-39.
- Lawrence-Brown, D. (2004). Differentiated instruction: Inclusive strategies for standards-based learning that benefit the whole class. *American secondary education*, 34-62.
- LeDoux, J. (2008). Remembrance of emotions past. In K. Fischer & M. H. Immordino-Yang (Eds.), *The Jossey-Bass reader on the brain and learning* (pp. 151–179). San Francisco: Jossey-Bass.
- Lightbown, P. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In Seliger, H. & Long M. (Eds.), *Classroom oriented research in second language acquisition* (pp. 217-245). Rowley, MA: Newbury House.
- Lightbown, P. M., Spada, N., & Wallace, R. (1980). Some effects of instruction on child and adolescent ESL learners. In Scarcella, R. C. & Krashen S. D. (Eds.), *Research in second language acquisition: Selected papers of the Los Angeles Second Language Acquisition Forum* (pp. 162-172). Rowley, MA: Newbury House.
- Lightbown, P., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: effects on second language learning. *Studies in second language acquisition*, 2(4), 429-448.
- Link, A.S. (1959) What Happened to the Progressive Movement in the 1920's, *The American Historical Review*, Vol. 64, No. 4 (Jul., 1959), pp. 833-851.

- Linstone, H. A., & Turoff, M. (2002). *The Delphi Method: Techniques and Applications*. Newark, NJ: New Jersey Institute of Technology.
- Little, D. (2006). The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. *Language Teaching*, 39(03), 167-190.
- Llera, J. A. B., & Sánchez, L. P. (2011). More than a century of educational psychology: overview and future prospects. *Papeles del Psicólogo*, 32(3), 204-231.
- Lochtman, K. (2002). Oral corrective feedback in the foreign language classroom: how it affects interaction in analytic foreign language teaching. *International Journal of Educational Research*, 37, 271-283.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In Ritchie, W. C. & Bhatia T. K. (Eds.), *Handbook of language acquisition, vol. 2: second language acquisition* (pp. 413-468). New York: Academic Press.
- Luhrman, G.J. (1993) *Erasmus and Foreign Language Acquisition*, in Noordegraaf, J. & Vonk, F. (Eds.). *Five hundred years of foreign language teaching in the Netherlands 1450-1950* (pp. 9-17). Amsterdam: Stichting Neerlandistiek VU.
- Lyster, R. (1994). The effects of functional-analytical teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics*, 15(3), 263-288.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37-66.
- Lyster, R., Lightbown, P. M., & Spada, N. (1999). A response to Truscott's 'What's wrong with oral grammar correction'. *The Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 55(4), 457-467.
- 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.
- Mackey, A. (2002). Beyond production: Learners' perceptions about interactional processes. *International Journal of Educational Research*, 37, 379-394.
- Manning, W. H. (1987). Development of Cloze-elide tests of English as a second language. *ETS Research Report Series*, 1987(1), i-136.

- Marzano, R. (2000). *Transforming classroom grading*. Alexandria, VA: ASCD.
- McTighe, J. & Brown, J.L.(2005) 'Differentiated Instruction and Educational Standards: Is Detente Possible?', *Theory Into Practice*, 44: 3, 234 — 244.
- McQuarrie, L. M., & McRae, P. (2010). A provincial perspective on differentiated instruction: The Alberta Initiative for School Improvement (AISI). *Journal of Applied Research on Learning*, 3(4), 1-18.
- McQuarrie, L., McRae, P., & Stack-Cutler, H. (2008). *Differentiated instruction, Provincial Research Review*. Edmonton: Alberta Initiative for School Improvement.
- Miller, A. (1991). Personality types, learning styles and educational goals. *Educational Psychology*, 11(3-4), 217–238.
- Miller, M. A., & Stoeckel, P. R. (2015). *Client education: Theory and practice*. Burlington, MA: Jones & Bartlett Publishers.
- Ministerie van OC&W (2011). *Actieplan vo 'Beter Presteren': opbrengstgericht en ambitieus. Het beste uit leerlingen halen.* 's-Gravenhage, OCW.
- Mohan, B. & Beckett, G. H., (2001). A functional approach to research on content-based language learning: Recasts in causal explanations. *The Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 58(1), 133-155.
- Moon, T. R.(2005) The Role of Assessment in Differentiation, *Theory Into Practice*, 44: 3, 226-233.
- Moskovsky, C., Alrabai, F., Paolini, S., & Ratcheva, S. (2013). The effects of teachers' motivational strategies on learners' motivation: A controlled investigation of second language acquisition. *Language Learning*, 63(1), 34-62.
- Muthén LK, Muthén BO. *Mplus: The Comprehensive Modeling Program for Applied Researchers*. User's Guide. Los Angeles, CA: Muthén & Muthén; 1998–2006.
- Nekoda, K. (2003). The connection problem of English language teaching in primary and secondary education in the Netherlands : a comprehensive study on the measures for improvement . *Bulletin of the Graduate School of education, Hiroshima University*, 51, 187-195.
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47, 203-211.

- Noordegraaf, J., & Vonk, F. (1993). *Five hundred years of foreign language teaching in the Netherlands 1450-1950*. Stichting Neerlandistiek VU, Amsterdam.
- Oakes, J., Gamoran, A., and Page, R.N.(1992) Curriculum differentiation: opportunities, outcomes and meanings. In Jackson, P.W.(Ed.) (1992)*Handbook of research on curriculum*, American Educational Research Association, ( 570-608) NY: Macmillan publishing
- Oadlin, T. (1986). Another look at passage correction tests. *TESOL Quarterly*, 20(1), 123-130.
- Onderwijsraad (2008) *Advies Vreemde talen in het onderwijs*, Nr. 20080102/915, juni 2008. Den Haag: Onderwijsraad,
- Oostdam, R.J., (2009),*Tijd voor dikke leerkrachten - Over maatwerk als kern van goed onderwijs, Openbare les uitgesproken op 18 november 2009*, Hogeschool van Amsterdam.
- Oostdam, R. en Toorenborg, H. van, (2002). Leuk is not enough. *Levende Talen Tijdschrift*, 3 (4), 16-19.
- Op 't Eynde, P,(2004) Leren doe je nooit alleen, differentiatie als sociaal gebeuren. *Impuls voor onderwijsbegeleiding* 1, 4-13, Leuven: Acco.
- Opdenakker, M. C., & Van Damme, J. (2006). Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice. *Teaching and teacher education*, 22(1), 1-21.
- Ostrowska, K. (2014). English in Adult Education (Master thesis) Reykjavík: Háskóla Íslands University.
- Panadero, E., Tapia, J. A., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and individual differences*, 22(6), 806-813.
- Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions. In Schutz, P.A. & Pekrun, R. (Eds.). *Emotion in education* (pp. 13-36). Amsterdam : Academic Press.
- Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In Pekrun, R., & Linnenbrink-Garcia, L. *International handbook of emotions in education* (120-141). NY, NY: Routledge.
- Pica, T., Holliday, L., Lewis, N. & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, 11, 63-90.

- Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Learner-learner interaction as a resource for language learning. *TESOL Quarterly*, 30(1), 59-84.
- Pica, T., & Washburn, G. N. (2003). Negative evidence in language classroom activities: A study of its availability and accessibility to language learners. *I.T.L. Review of Applied Linguistics*, 301-344.
- Pligt, J. V. D., & De Vries, N. K. (1998). Expectancy-value models of health behaviour: The role of salience and anticipated affect. *Psychology and Health*, 13(2), 289-305.
- Polderman, T. J. C., Boomsma, D. I., Bartels, M., Verhulst, F. C., & Huizink, A. C. (2010). A systematic review of prospective studies on attention problems and academic achievement. *Acta Psychiatrica Scandinavica*, 122(4), 271-284.
- Popham, W. J. (2011). *Transformative assessment in action: An inside look at applying the process*. Alexandria, VA: ASCD.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, 135(2), 322.
- Rabley, S. (2008) *The Fireboy*. Amsterdam: Pearson Education Ltd.
- Rayman, R. (1981) Joseph Lancaster's Monitorial System of Instruction and American Indian Education, 1815-1838. *History of Education Quarterly*, 395-409.
- Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, 98(1), 209.
- Reezigt, G. J. (1993). *Effecten van differentiatie op de basisschool*. Groningen: RION.
- Reigeluth, C. (2012). Instructional theory and technology for the new paradigm of education. *Revista de Educación a Distancia*, 32, 1-18.
- Richards, J.C. & Rodgers, T.S. (1986). *Approaches and methods in language teaching*. Cambridge: CUP (5<sup>th</sup> print 1999).
- Robison, E. M. (2004). *Teacher decision-making in utilising differentiated instruction*. Scranton, Pa. : Marywood University.
- Rock, M. L., Gregg, M., Ellis, E., & Gable, R. A. (2008). REACH: A framework for differentiating classroom instruction. *Preventing School Failure*. 52(2), 31-47.

- Rollins, R.C. (2014) Origins of Progressivism. In *Encyclopædia Britannica online*. Retrieved 29-05-2016 from <http://www.britannica.com/EBchecked/topic/616563/United-States/77842/Origins-of-progressivism>
- Runco, M. A., & Chand, I. (1995). Cognition and creativity. *Educational psychology review*, 7(3), 243-267.
- Rutherford, W., & Sharwood Smith, M., (1985). Consciousness-raising and universal grammar. *Applied Linguistics*, 6, 274-282.
- Rutherford, W., & Sharwood Smith, M., (1988). *Grammar and second language teaching*. New York: Newbury House.
- Rowe, G., & Wright, G. (2001). Expert opinions in forecasting: the role of the Delphi technique In Armstrong, J. S. (Ed.). *Principles of forecasting: a handbook for researchers and practitioners (Vol. 30)* (125-144). New York, NY: Springer Science+ Business.
- Sattarpour, S., & Ajideh, P. (2014). Investigating the Relatedness of Cloze-Elide Test, Multiple-Choice Cloze Test, and C-test as Measures of Reading Comprehension. *Journal of English Language Teaching and Learning*, 6(13), 89-112.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese. In R. Day (Ed.), *"Talking to Learn": Conversation in second language acquisition* (237-326). Rowley, MA: Newbury House.
- Schooten, E. van, & Emmelot, Y., (2004). *De integratie van taal- en vakonderwijs*. Amsterdam: SCO-Kohnstamm Instituut. (SCO-rapport 717).
- Schooten, E. J. van, (2005). *Literary response and attitude toward reading fiction* (Doctoral dissertation), University of Groningen.
- Scudellari, M. (2015). The science myths that will not die. *Nature*, 528(7582), 322-325.
- Seidel, T. & Shavelson, R.J., (2007) Teaching Effectiveness Research in the Past Decade: The Role of Theory and Research Design in Disentangling Meta-Analysis Results, *Review of Educational Research*, 77 (4) 454–499.
- Sitton, R. (1995). *Improve student spelling achievement*. Spokane, WA: Egger Publishing.

- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2008). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement* 2009, 69: 493.
- Slavin, R. E. (1995). *Cooperative learning* (second edition). Boston: Allyn and Bacon.
- Sligte, H., Bulterman-Bos, J., & Huizinga, M. (2009). *Maatwerk Voor Latente Talenten? Uitblinken op alle niveaus; Eindrapport* Amsterdam, Netherlands: SCO-Kohnstamm Instituut, University of Amsterdam.
- Smith, R. C. (2005). *An Investigation into the Roots of ELT: With a Particular Focus on the Career and Legacy of Harold E. Palmer (1877-1949)* (Unpublished Doctoral dissertation, University of Edinburgh).
- Snow, R. E. (1992). Aptitude theory Yesterday, today, and tomorrow. *Educational psychologist*, 27(1), 5-32.
- Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of information technology education*, 6, 1.
- Spada, N., & Lightbown, P. M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, 15, 205-224.
- Sprenger, M. (2003). *Differentiation through learning styles and memory*. Thousand Oaks, CA: Corwin Press Incorporated.
- Staatsen, F., Heebing, S., & van Renselaar, E. (2009). *Moderne vreemde talen in de onderbouw*. Bussum: Coutinho.
- Stauffer, R. (1959). Individualizing Reading Instruction - A Backward Look. *Elementary English*, 36(5), 335-341. Retrieved 29-05-2016 from <http://www.jstor.org/stable/41384896>
- Stiggins, R. J., Arter, A.J., Chappius, J. & Chappius S. (2004). *Classroom assessment for student learning: doing it right--using it well*. Portland, Oregon: Assessment Training Institute, Inc.,
- Stillings, N., Feinstein, M., Garfield, J., Risland, E., Rosenbaum, D., Weisler, S. and Barker-Ward, L. (1987). *Cognitive science: an introduction*. Cambridge, MA: The MIT Press.
- Stoks, G.L.M.M. & Voortman, M. (1982) *Wegwijs in Engels op de basisschool*. Enschede: SLO
- Stoll Lillard, A. (2005) *Montessori – The Science behind the genius*, New York: Oxford University Press.

- Stronge, J. H., Ward, T. J., Tucker, P. D., & Hindman, J. L. (2007). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education*, 20(3-4), 165-184.
- Stronsnider, R. & Lyon, C. (1997). Including students with disabilities into the regular classroom. *Education*, 117, 611-622.
- Subban, P. (2006) *Differentiated instruction: A research basis*, International Education Journal, 2006, 7(7), 935-947, Perth: Shannon Research Press.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass, S. & Madden C. (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review/La revue canadienne des langues vivantes*, 50, 158-164.
- Swain, M. (1995). Three functions of output in second language learning: In Cook G. & Seidlhoffer B. (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M. (2001). Integrating Language and Content Teaching through Collaborative Tasks. *The Canadian Modern Language Review/La Revue Canadienne des langues vivantes*, 58, 1, 44-63.
- Swan, M. (1985). A critical look at the communicative approach (2). *ELT journal*, 39(2), 76-87.
- Thiergen, O. (1900) Reform in the German Higher Schools and the Reform Movement in *Modern Language Teaching, The School Review*, Vol. 8, No. 4 (Apr., 1900), pp. 230-233
- Thijs, A. M., Tuin, D., & Trimbos, B. (2011). *Engels in het basisonderwijs: verkenning van de stand van zaken*. Enschede: Netherlands Institute for Curriculum Development, SLO.
- Thorndike, E. L. (1906). *The principles of teaching - Based on psychology.*, Syracuse, NY: The Mason Press
- Tieso, C. (2005). The Effects of Grouping Practices and Curricular Adjustments on Achievement. *Journal for the Education of the Gifted*, 29(1), 60-89.
- Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, 64(1), 37-54.

- Tokuhama-Espinosa, T. N. (2008). *The Scientifically Substantiated Art of Teaching: A Study in the Development of Standards in the New Academic Field of Neuroeducation (mind, Brain, and Education Science)*. Ann Arbor, MI: ProQuest.
- Tomlinson, C. A. (1999). Leadership for Differentiated Classrooms. *School Administrator*, 56(9), 6-11.
- Tomlinson, C. A. (2000). Reconcilable differences: Standards-based teaching and differentiation. *Educational leadership*, 58(1), 6-13.
- Tomlinson, C. A., (2001). *How to differentiate instruction in mixed-ability classrooms*. (2nd Ed.) Alexandria, VA: ASCD.
- Tomlinson, C. A. (2003). *Fulfilling the promise of the differentiated classroom: Tools and strategies for responsive teaching*. Alexandria: ASCD.
- Tomlinson, C. A. (2004) The Möbius Effect: Addressing Learner Variance in Schools, *Journal of Learning Disabilities*, 37, 516
- Tomlinson, C. A. (2005) Grading and Differentiation: Paradox or Good Practice?, *Theory Into Practice*, 44 (3), 262 — 269
- Tomlinson, C. A. (2014). *Differentiated classroom: Responding to the needs of all learners.*, Alexandria, VA: ASCD
- Tomlinson, S. (1997). Edward Lee Thorndike and John Dewey on the science of education. *Oxford Review of Education*, 23(3), 365-383.
- Tomlinson, C. A., & Allan, S. D. (2000) *Leadership in differentiating schools and classrooms*. ASCD, Alexandria, VA.
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, 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, 119-145.
- Tomlinson, C. A. & Imbeau, M. (2010) *Leading and Managing a Differentiated Classroom*, 2010, Alexandria, VA: ASCD.

- Tomlinson, C. A., & Kalbfleisch, M. L. (1998). Teach Me, Teach My Brain: A Call for Differentiated Classrooms. *Educational Leadership*, 56(3), 52-55.
- Tomlinson, C. A., & McTighe, J. (2006). Integrating differentiated instruction and understanding by design, *Theory Into Practice*, 44 (3), 262 — 269
- Tomlinson, C. A. ,Moon, T. R., & Callahan, C., M. (1998). How Well Are We Addressing Academic Diversity in the Middle School. *Middle School Journal*, 29, 3-11.
- Toorenburg van, H., & Bodde-Alderlieste, M. (2003) *Vakdossier Engels basisschool*. Enschede: SLO
- Torrance, E. (1995). Insights about creativity: Questioned, rejected, ridiculed, ignored. *Educational Psychology Review*, 7, 313–322.
- Tosun, B. C. (2009). A New Challenge in the Methodology of the Post-Method Era. *Journal of Language and Linguistic Studies*, 5(2), 1-8.
- Turner, J. C. (1993). A motivational perspective on literacy instruction. In Leu, D. J. & Kinzer, C. K. (Eds.), *Examining central issues in literacy research, Theory, and Practice. Forty-Second Yearbook of the National Reading Conference*. (pp. 153-161). Chicago: NRC.
- Tyler, J. A. (2009). Moving beyond scholar-practitioner binaries: Exploring the liminal possibilities of the borderlands. *Advances in Developing Human Resources*, 11(4), 523-535.
- UNICEF (The United Nations Children's Fund) (2013). *Report Card 11: Child wellbeing in rich countries*. Florence, Italy: Unicef Innocenti Research Centre.
- Ushioda, E., & Dörnyei, Z. (2009). *Motivation, language identities and the L2 self*. Bristol, UK: Multilingual Matters.
- Vallerand, R., Gagné, F., Senecal, G., & Pelletier, L. (1994). A comparison of the school intrinsic motivation and perceived competence of gifted and regular students. *Gifted Child Quarterly*, 36, 68–72.
- Valette, R. M. (1967). *Modern language testing: A handbook*. New York: HBW Inc.
- Van Driel, J. H., Meirink, J. A., Van Veen, K., & Zwart, R. C. (2012). Current trends and missing links in studies on teacher professional development in science education: a review of design features and quality of research. *Studies in Science Education*, 48(2), 129-160.

- Van Veen, K. (2008). Analysing teachers' working conditions from the perspective of teachers as professionals: The case of Dutch high school teachers. In J. Ax & P. Ponte (Eds.), *Critiquing praxis: Conceptual and empirical trends in the teaching profession* (pp. 91-112). Rotterdam: Sense Publishers.
- van Veen, K., Zwart, R. C., Meirink, J. A., & Verloop, N. (2010). *Professionele ontwikkeling van leraren: een reviewstudie naar effectieve kenmerken van professionaliseringsinterventies van leraren*. ICLON/Expertisecentrum Leren van Docenten.
- Vatz, K., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Aptitude-treatment interaction studies in second language acquisition. Findings and methodology. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 273- 292). Amsterdam: John Benjamins.
- Verhelst, N., Van Avermaet, P., Takala, S., Figueras, N., & North, B. (2009). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: CUP.
- Verspoor, M. H., Schuitemaker-King, J. Van Rein, E. M. J., Bot, C. J. L. de & Edelenbos, P. (2010). *Tweetalig onderwijs: vormgeving en prestaties. Onderzoeksrapport*. Groningen: Rijksuniversiteit Groningen
- Vinjé, M. (Ed.) (1993). *Balans van het Engels aan het einde van de basisschool. Uitkomsten van de eerste peiling Engels einde basisonderwijs*. (PPON-reeks nr. 4). Arnhem: Cito
- VO-raad (2015). *Diploma op maat: ruimte voor talent in het voortgezet onderwijs*. Retrieved 29-05-2016 from <http://www.vo-raad.nl/userfiles/bestanden/Gepersonaliseerd%20leren/Notitie-Diploma-op-maat-ruimte-voor-talent-in-het-vo-maatwerkdiploma.pdf>.
- Vonk, F (1993) Between School and University The study of German in Utrecht, 1876-1921, in Noordegraaf, J., & Vonk, F. (1993). *Five hundred years of foreign language teaching in the Netherlands 1450-1950*. Amsterdam: Stichting Neerlandistiek VU: 9-17.
- Washburne, C. W. (1922). Educational measurement as a key to individual instruction and promotions. *The Journal of Educational Research*, 5(3), 195-206.
- Washburne, C. W. (1953). Adjusting the Program to the Child. *Educational Leadership*, 11(3), 139-140.

- White, L. (1987). Against comprehensible input: The input hypothesis and the development of L2 competence. *Applied Linguistics*, 8, 95-110.
- White, L. (1990). Implications of learnability theories for second language learning and teaching. In Halliday, M. A. K., Gibbons, J. & Nicholas, H. (Eds.), *Learning, keeping and using language*, Vol. 1. Amsterdam: Benjamins.
- White, L. (1991). Adverb placement in second language acquisition: some effects of positive and negative evidence in the classroom. *Second Language Research*, 7(2), 122-161.
- White, L., Spada, N., Lightbown, P. M., & Ranta, L. (1991). Input enhancement and L2 question formation. *Applied Linguistics*, 12, 416-432.
- Wilhelm, F. A. (2005). *English in the Netherlands: a history of foreign language teaching 1800-1920: with a bibliography of textbooks*. Utrecht: Gopher Publishers.
- Wilhelm, F. A. (2009) *Breeze or storm? The European Reform Movement and Dutch Foreign Language Teaching, c1880-1920*, Documents pour l'histoire du français langue étrangère ou seconde (online), 43, 2009: <http://dhfles.revues.org/881>
- Wilkins, D. A. (1972). *Grammatical, Situational and Notional Syllabuses*. Paper presented at the AILA (3, Copenhagen, August 1972). Heidelberg: Julius Gross
- Wiggins, G. (1993). *Assessing student performance*. San Francisco: Jossey-Bass.
- Wiggins, G. P., & McTighe, J. (1998). *Understanding by design*. Alexandria, Va: ASCD
- Williams, J. (1997). *Effectiveness of learner and teacher-initiated focus on form*. Paper presented to Second Language Research Forum, Michigan: Michigan State University.
- Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In Doughty C. & Williams J. (Eds.), *Focus on form in second language acquisition* (pp. 139-155). New York: CUP.
- Yeo, R. (2003). *Defining science: William Whewell, natural knowledge and public debate in early Victorian Britain* (Ideas in Context, Vol. 27). Cambridge: CUP
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of educational Psychology*, 82(1), 51.

## 11. Appendix

### Table of contents

1. Questionnaire on the methodology used for English language teaching (Ch. 3).....	201
2. Structure matrix from the PCA (Chapter 3).....	203
3.1. The 37 specific types of implementation barriers .....	205
3.2. Rubrics as used during the experimental phase (Chapter 5):	
3.2.A. Rubrics used in the AYLITT project,.....	206
3.2.B. Rubrics as developed for the language teaching programme based on differentiated instruction.....	207
4. Attitude test used during the experimental phase (Chapter 6).....	209
5. Vocabulary MC test (Chapter 6).....	210
6. Dictation (Chapter 6).....	225
7. Editing test texts (Chapter 6).....	226
8. Writing test assignments (Chapter 6).....	228
9. Background information form for the students in the experiment (Chapter 6).....	229
10. Additional tables Chapter 6:	
10.a. Confirmatory factor analyses MPB pre- and post-test groups,.....	231
10.b. MI analyses to verify whether experimental and control group differ,.....	233
10.c. Correlations of sums of the MPB scores and Results of multi-level analyses of the effects of the treatment on language skills scores after controlling for covariates.....	236
10.d. Correlations of Reading skills scores. ....	275
11. Delphi Round 1 -Rephrased Complete (Chapter 7). ....	276
12. Delphi Round 1 -Rephrased and summarised for round 2 (Chapter 7). ....	282
13. Delphi Round 2-Rephrased and summarised for round 3 (Chapter 7). ....	287
14. Delphi Round 3 proposed text (Chapter 7).....	295

## **1. Questionnaire on the methodology used for English language teaching in primary- and secondary education.**

### **1. Personal details**

- 1.1 School BRIN – number
- 1.2 Gender (M/F)
- 1.3 Age (in years)
- 1.4 Country of birth
- 1.5 Teacher training: Bachelor / Master / PhD.
- 1.6 How long have you been teaching: (in years)
- 1.7 Do you work in primary / secondary education

### **Primary school teaching details**

2. How much English do you offer on average throughout the school year?
  - 2.1 How often, on average throughout the school year, is your English teaching based on a coursebook?
  - 2.2 How often, on average throughout the school year, is your English teaching not based on a coursebook?
  - 2.3 How often, on average throughout the school year, do you discuss the meaning of English words in other subjects taught?
  - 2.4 How often, on average throughout the school year, do you discuss the English culture in other subjects taught?
  - 2.5 How often, on average throughout the school year, do you sing English songs with your learners?
  - 2.6 How often, on average throughout the school year, do you play English (language) games with your learners?
  - 2.7 How often, on average throughout the school year, do you offer other subjects in English songs?  
(scale used for 2.1 – 2.7: Never – <10times/year – 1-3 times/month – less than once a week – once a week)
  - 2.8 Does your school offer English language teaching to 4- to 12-year-olds? (Yes/No)
  - 2.9 In which year of your school do you start to offer English? (1 = 4-year-olds...8=12-year-olds)
  - 2.10 When learners spent their school career in your school, with how many years of English teaching do they leave your school?
  - 2.11 Do you cooperate with a (near) Native Speaker? (Yes/No)
  - 2.12 Are you a Native Speaker? (Yes/No)
  - 2.13 When you taught English this past academic year, how long did an English lesson last?  
(scale used: < 10 min. – 10-30min. – 30-45min. – 45-60min. – >60min.)
  - 2.14 When English was offered by a (near) Native Speaker this past academic year, how long did an English lesson last?  
(scale used: N/A – < 10 min. – 10-30min. – 30-45min. – 45-60min. – >60min.)
  - 2.15 In what way do you offer information about English language learning gains of your learners to the secondary school they go to?  
(scale used: I don't – I hand over the marks obtained – I hand over personally written reports – I hand over language portfolio's – I discuss the language skills of individual students – another way:....)
  - 2.16 Is there any other information about your English teaching you would like to share? (open)

### **Generic part of the questionnaire**

3. What does your English teaching look like?  
How much time do you spend on average on:
  - 3.1 Organisation of the school, classroom, materials and such?
  - 3.2 Listening activities?
  - 3.3 Reading activities?
  - 3.4 Speaking activities?
  - 3.5 Writing activities?

3.6 Grammar

3.7 History, culture and current affairs from the English speaking world

(scale used for 3.1 – 3.7: no time – a few minutes – 5 minutes – 10 minutes – 15 minutes – 20 minutes – 25 minutes – 30 minutes – 35 minutes – 40 minutes – 45 minutes – 50 minutes – 55 minutes or more.)

**Which activities are used in your teaching?**

3.8 I use the English as language of instruction

3.9 Listening to English stories

3.10 Listening for information in English

3.11 Listening to authentic English listening materials

3.12 Watching Video / YouTube / TV

3.13 Reading English stories

3.14 Reading informative texts in English

3.15 Speaking - dialogues/conversations

3.16 Speaking - presenting/monologues – show and tell

3.17 Singing

3.18 Writing - copying

3.19 Writing – practice (match/puzzle/connect/cloze)

3.20 Writing – creative writing

3.21 Writing – filling in forms

3.22 Making a Poster or brochure

3.23 Grammar – explaining rules/constructs

3.24 Grammar – practice

3.25 Total Physical Response - say and act (with drama and singing as well)

3.26 Drama – acting out meaning, situations or stories

3.27 Playing (language) games

3.28 Drills and choral work

3.29 Cooperative learning activities

3.30 Group work

3.31 Project based activities

3.32 Theme based activities

3.33 Activities based on course books

3.34 Activities adapted to the learners' perception of the world

3.35 Activities adapted to the learners' language skills

3.36 Activities adapted to the learners' wishes or suggestions

3.37 Activities adapted to the learners' English teaching experiences in previous years

3.38 Diagnostic tests

3.39 Formative tests

3.40 Summative tests

3.41 Other:.....

(scale used for 3.8 – 3.41: almost never – half of the lessons taught – almost every lesson)

## 2. Structure matrix

Structure Matrix

	Component							
	1	2	3	4	5	6	7	8
21 Using spoken English for all interactions during the lesson	,090	,206	,277	,311	-,550	,333	-,304	-,378
22 Listening to stories in English	,300	,194	,322	,230	-,810	,085	,130	-,092
23 Listening for information in English	,292	,056	,390	,184	-,820	,127	,180	-,232
24 Listening to authentic English listening materials	,169	,065	,196	,163	-,799	,050	-,034	,171
25 watching Videos / YouTube / TV(recordings)	,425	,112	,074	,496	-,469	,383	,114	,370
26 Reading stories in English	,468	-,210	,366	,460	-,581	,387	,279	,075
27 Reading informative texts in English	,370	-,462	,328	,449	-,478	,437	,248	,179
28 Speaking - dialogues/conversations	,350	,390	,539	,215	-,450	-,090	,049	-,321
29 Speaking - presenting/performing monologues – show and tell	,596	,102	,359	,493	-,321	,251	-,143	,198
Singing	,141	,780	,289	,108	-,250	-,257	,206	-,067
30 Writing – copying	,097	,315	,751	,020	-,252	,268	,014	-,139
31 Writing – processing (new) language (match/puzzle/connect/cloze)	,206	,205	,791	,150	-,318	,171	-,037	-,259
32 Writing – creative writing	,476	-,362	,545	,352	-,354	,240	-,110	,208
33 Writing – filling in forms	,374	-,256	,716	,205	-,280	,049	,197	,094
34 Creating posters or brochures	,558	-,039	,391	,410	-,284	,218	,082	,342
35 Grammar – explaining and/discovering rules	,237	-,164	,155	,333	-,094	,912	-,023	-,041
36 Grammar – practise	,246	-,198	,213	,317	-,146	,902	,022	-,040
37 Total Physical Response - say and act (also in drama and song)	,288	,683	,270	,277	-,276	,051	,434	,140
38 Drama – acting out meaning of words/expressions/stories	,369	,582	,202	,529	-,306	-,034	,312	,210
39 Playing language games	,432	,312	,286	,585	-,337	,194	,286	,157
40 Drills and choral work	,231	,339	,145	,328	-,280	,137	,700	-,028
41 Cooperative learning activities	,721	,176	,252	,326	-,260	,081	,379	-,302
Group work	,805	,037	,278	,361	-,217	,051	,217	-,060
42 Projects	,801	-,058	,210	,427	-,318	,338	-,027	,339
43 Thematic activities	,721	,021	,258	,513	-,377	,299	,088	,235
44 Course book based activities	,184	-,046	,270	,269	-,211	,238	,034	-,668
45 Activities geared to the pupils' perception of their environment	,406	,047	,059	,889	-,214	,302	,122	,013
46 Activities geared to the pupils' language skills levels	,301	-,025	,045	,871	-,219	,245	,054	-,097
47 Activities adapted to the pupils' wishes or suggestions	,445	,129	,176	,783	-,245	,266	,165	,115
48 Activities adapted to the pupils' ELT experiences of previous years	,420	,046	,250	,762	-,210	,361	,004	,083
49 Diagnostic tests (to tailor the programme to the findings)	,683	,040	,242	,443	-,264	,435	-,115	,116
50 Formative tests (to tailor the activities/programme to the findings)	,591	,055	,248	,451	-,421	,548	-,216	,141
51 Summative tests	,306	,191	,227	,343	-,365	,493	-,216	,153

Extraction Method: Principal Component Analysis.  
Rotation Method: Oblimin with Kaiser Normalization.

Component	1	2	3	4	5	6	7	8
1	1,000	,036	,313	,474	-,321	,223	,118	,124
2	,036	1,000	,084	,077	-,121	-,085	,081	-,055
3	,313	,084	1,000	,183	-,344	,151	,053	-,068
4	,474	,077	,183	1,000	-,315	,331	,105	,083
5	-,321	-,121	-,344	-,315	1,000	-,194	-,055	-,006
6	,223	-,085	,151	,331	-,194	1,000	-,095	,037
7	,118	,081	,053	,105	-,055	-,095	1,000	-,005
8	,124	-,055	-,068	,083	-,006	,037	-,005	1,000

Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization.

**3.1. The 37 specific types of implementation barriers According to Sanetti and Kratochwill (2009), in 4 categories.**

**Intervention**

1. Intervention compatibility
2. Time/duration
3. Materials/resources required
4. # of interventionists required
5. Actual effectiveness
6. Extent adaptable to context/needs
7. Ease of implementation
8. Complexity
9. Rate of behaviour change
10. Improvement over previous practices

**Organisation**

11. Leadership support
12. Adequate staffing
13. Time for planning/implementation
14. Facilitation strategies
15. Opportunities for communication
16. Access to needed materials
17. Adequate facilities
18. Integration of existing priorities
19. Mechanisms for communication
20. Adequate funding
21. Positive climate
22. Positive norms regarding change

**Implementer**

23. Perceptions of recipient
24. Skill proficiency
25. Self-efficacy
26. Perceived need
27. Willingness to try the intervention
28. Perceived intervention effectiveness
29. Motivation to implement
30. Shared decisions/buy-in/vision
31. Perceptions of relative advantage
32. Perceptions of role compatibility

**External environment**

33. Level of support from stakeholders
34. Bureaucratic/political barriers
35. Level of opposition
36. Consistency with policies
37. Coordination with other agencies

### 3.2. Rubrics as used during the experimental phase

#### A. Rubrics as used in the AYLLIT-project (The AYLLIT criteria (Hasselgreen et al., 2011; 23-24))

Levels	Overall structure and range of information	Sentence structure and grammatical accuracy	Vocabulary and choice of phrase	Misformed words and punctuation
<b>Above B1</b>	Is able to create quite complicated texts, using effects such as switching tense and interspersing dialogue with ease. The more common linking words are used quite skilfully.	Sentences can contain a wide variety of clause types, with frequent complex clauses. Errors in basic grammar only occur from time to time.	Vocabulary may be very wide, although the range is not generally sufficient to allow stylistic choices to be made.	Misformed words only occur from time to time.
<b>B1</b>	Is able to write texts on themes which do not necessarily draw only on personal experience and where the message has some complication. Common linking words are used.	Is able to create quite long and varied sentences with complex phrases, e.g. adverbials. Basic grammar is more often correct than not.	Vocabulary is generally made up of frequent words and phrases, but this does not seem to restrict the message. Some idiomatic phrases used appropriately.	Most sentences do not contain misformed words, even when the text contains a wide variety and quantity of words.
<b>A2/B1</b>	Is able to make reasonable attempts at texts on familiar themes that are not completely straightforward, including very simple narratives. Clauses are normally linked using connectors, such as "and", "then", "because", "but".	Sentences contain some longer clauses, and signs are shown of awareness of basic grammar, including a range of tenses.	Vocabulary is made up of very common words, but is able to combine words and phrases to add colour and interest to the message (e.g. using adjectives).	Clear evidence of awareness of some spelling and punctuation rules, but misformed words may occur in most sentences in more independent texts.

<b>A2</b>	Can write short straightforward coherent texts on very familiar themes. A variety of ideas are presented with some logical linking.	Is able to make simple independent sentences with a limited number of underlying structures.	Vocabulary is made up of very frequent words but has sufficient words and phrases to get across the essentials of the message aspired to.	Some evidence of knowledge of simple punctuation rules, and the independent spelling of very common words.
<b>A1/A2</b>	Can adapt and build on a few learnt patterns to make a series of short and simple sentences. This may be a short description or set of related facts on a very familiar personal theme.		Can use some words which may resemble L1, but on the whole the message is recognisable to a reader who does not know the L1. Spelling may be influenced by the sound of the word and mother tongue spelling conventions.	
<b>A1</b>	Can write a small number of very familiar or copied words and phrases and very simple (pre-learnt) sentence patterns, usually in an easily recognisable way. The spelling often reflects the sound of the word and mother tongue spelling conventions.			
<b>Approaching A1</b>	Makes an attempt to write some words and phrases, but needs support or a model to do this correctly.			

Table 1: AYLLIT writing descriptor grid

## B. Rubrics as used in the treatment during the experimental phase of this study

Niveauschaal Schrijven sept. 13.1				
niveau	Teksten	Zinnen	Woorden & Interpunctie	Originaliteit
13	Ik schrijf complexe teksten. Ik gebruik verschillende tijden. Ik kan citeren in een tekst.	Ik schrijf correcte zinnen in de tijd die het best past bij mijn verhaal of informatie.	Ik pas mijn woordgebruik aan waar nodig.	
+				
12				
+				
11	Ik schrijf teksten over dingen die ik zelf niet heb meegemaakt, ook als het ingewikkeld is. Ik kan stukjes tekst makkelijk en goed aan elkaar verbinden.	Ik kan langere zinnen schrijven die uit meerdere delen kunnen bestaan.	Ik kan één onderwerp met afwisselende woorden beschrijven.	
+				
10				
+				
9	Ik schrijf redelijk begrijpelijke stukken tekst over bekende onderwerpen die niet direct simpel zijn. Ik gebruik 'and, then, because, but om zinnen aan elkaar te plakken.	Ik schrijf langere zinnen die uit twee delen kunnen bestaan. De zinnen die ik schrijf zijn (grammaticaal) correct	Ik kan ook minder vaak gebruikte woorden goed spellen. Ik gebruik soms dezelfde woorden.	
+				
8				
+				
7	Ik schrijf korte duidelijke teksten over bekende onderwerpen. Ik schrijf in een logische volgorde, of in samenhang.	Ik schrijf in complete simpele zinnen.	Ik kan vaak gebruikte woorden goed spellen. Ik gebruik vaak dezelfde woorden	
+				
6				
+				
5	Ik gebruik stukken en delen (stones) van zinnen om een verhaal te schrijven. Ik gebruik korte beschrijvingen en opsommingen.	Ik gebruik woorden die op het Nederlands lijken. Ook iemand die geen Nederlands kent begrijpt het verhaal of de informatie.		Ik zoek woorden, uitdrukkingen en gegevens op en oefen vooraf om een interessante presentatie te leveren.
+				
4				
+				
3	Ik schrijf heel bekende woorden, of (gekopieerde) stukjes (stones) aan elkaar.	Ik schrijf woorden vaak zoals ze klinken.		Ik zoek woorden op die ik nodig heb voor de presentatie. Ik oefen wel eens. Ik probeer aan de opdracht te voldoen.
+				
2				
+				
1	Ik probeer woorden en zinsdelen in het Engels te schrijven. Ik heb veel hulp en voorbeelden nodig.			Ik gebruik alleen de woorden die ik al ken en probeer aan (een deel van) de opdracht te voldoen.

Niveauschaal Spreken sept. 13.1				
niveau	Woordgebruik	Communiceren	Vlotheid	Originaliteit
+			Ik praat aan één stuk door. Ik onderbreek mezelf niet.	
5	Ik gebruik simpele woorden en zinnen en kan praten over persoonlijke gegevens en 'echte situaties. Ik gebruik geen Nederlandse woorden of zinnen.	Ik kan in hele zinnen een gesprekje voeren over wat ik leuk vindt en meemaak.	Ik onderbreek mezelf minder dan de helft van de tijd.	Ik zoek woorden, uitdrukkingen en gegevens op en oefen vooraf om een interessante presentatie te leveren.
+				
4	Ik gebruik simpele woorden en zinnen. Ik gebruik bijna geen Nederlandse woorden of zinnen.	Ik kan met korte zinnen een gesprekje voeren over wat ik leuk vindt en meemaak.	Ik onderbreek mezelf ongeveer de helft van de tijd.	
+				
3	Ik gebruik simpele woorden en zinnen. Ik gebruik minder dan de helft van de tijd Nederlandse woorden of zinnen.	Ik kan met losse woorden, delen van zinnen en gebaren vertelb over wat ik leuk vindt en meemaak.	Ik onderbreek mezelf meer dan de helft van de tijd.	Ik zoek woorden op die ik nodig heb voor de presentatie. Ik oefen wel eens. Ik probeer aan de opdracht te voldoen.
+				
2	Ik probeer Engels te spreken. Ik gebruik ongeveer de helft van de tijd Nederlandse woorden. Ik heb gebaren nodig om mijzelf duidelijk te maken.	Ik kan met 'Yes' en 'No' en losse woorden' reageren op vragen over wat ik leuk vindt en meemaak.	Ik zeg met moeite en lang nadenken bijna wat ik wil zeggen.	
+				
1	Ik probeer Engels te spreken. Ik gebruik meer dan de helft van de tijd Nederlandse woorden. Ik heb gebaren nodig om mijzelf duidelijk te maken.	Ik kan met 'Yes' en 'No' reageren op vragen over wat ik leuk vindt en meemaak.	Ik heb moeite om iets te zeggen in het Engels	Ik gebruik alleen de woorden die ik al ken en probeer aan (een deel van) de opdracht te voldoen.

#### 4. Attitude test used during the experimental phase

##### Wat vind jij van Engels in het onderwijs?

Hieronder staan 26 zinnen die iets zeggen over hoe mensen tegen het leren van Engels op school aankijken, hoe belangrijk het is, wat zij er over denken, hoe zij zich erbij voelen. De opdracht is om bij iedere regel te kiezen hoeveel je het met ze eens bent en een kruisje te zetten in het vak wat hoort bij jouw ideeën over het leren van Engels. Denk niet te lang na over een antwoord, je eerste indruk is goed. Er zijn geen goede of foute antwoorden.

Een kruisje onder 1 = Ik ben het er **helemaal** mee eens

Een kruisje onder 2 = Ik ben het er een **beetje** mee eens

Een kruisje onder 3 = Ik ben het er **niet** mee eens, maar ook **niet** mee oneens

Een kruisje onder 4 = Ik ben het er een **beetje** mee oneens

Een kruisje onder 5 = Ik ben het er **helemaal** mee oneens

Een voorbeeld: Iemand wil weten of jij vindt dat je geld moet kunnen verdienen met het leren van Engels. Daar ben je het helemaal mee eens; dus zet je een kruisje onder 1.

		1	2	3	4	5
0.	Bij voorbeeld Ik zou geld moeten krijgen voor iedere voldoende die ik haal voor Engels	X				

Nu is het aan jou om per regel aan te kruisen of je het eens bent met de zinnen die er staan. Ze gaan allemaal over het leren van Engels op school.

		1	2	3	4	5
1.	Engels leren is goed voor je baan na school					
2.	Engels leren is nodig om de wereld te begrijpen					
3.	Engels leren is belangrijk voor het gewone dagelijks leven					
4.	Engels leren is belangrijk voor de school of opleiding na deze school					
5.	Engels leren is nergens goed voor					
6.	Engels leren is nodig voor het spreken met buitenlanders (bijv. op vakantie)					
7.	Engels leren is fantastisch					
8.	Engels leren is leuk					
9.	Engels leren betekent niet zo veel voor mij					
10.	Ik heb een beetje een hekel aan het leren van Engels					
11.	Ik haat Engels leren					
12.	Iedere Nederlandse jongere hoort Engels te leren					
13.	Scholen zouden meer tijd aan onderwijs Engels moeten besteden					
14.	Het leren van Engels hoort bij een goede opleiding					
15.	Een Nederlandse leerling moet zijn best doen om Engels te leren.					
16.	Iemand die niet probeert om Engels te leren moet zich schamen					
17.	Ik heb genoeg tijd om Engels te leren					
18.	Ik heb een rustige plek om mijn Engels te leren					
19.	Ik kan net zoveel Engelse TV/films/series/websites/etc. bekijken als ik wil					
20.	Ik krijg genoeg uitleg en materialen om Engels te leren					
21.	De lessen Engels die ik krijg sluiten goed aan bij wat ik al kan (niet te makkelijk of te moeilijk)					
22.	Mijn volgende huiswerk voor Engels ga ik zo snel mogelijk doen					
23.	Ik ga binnenkort een Engelse film/serie/documentaire kijken					
24.	Ik denk niet dat ik binnenkort mijn best ga doen voor Engels					
25.	In de komende vakantie ga ik ook wat voor Engels doen					
26.	Als ik binnenkort Engels sprekende mensen ontmoet, ga ik zeker Engels tegen ze spreken					

**5. Vocabulary MC test (based on the Peabody) used as pre- and post-test language skills measurement instrument during the experimental phase**

Deze toets meet hoeveel Engelse woorden je kent. Hij bestaat uit 40 keer de vraag:

**Kruis het vakje aan wat hoort bij het Engelse woord: ...**

Je ziet elke keer vier plaatjes waar je er één van moet kiezen die het best past bij het Engelse woord. Hieronder vind je een voorbeeld waarbij je het vakje moet aankruisen wat het best hoort bij het woord 'ball'. Het plaatje wat het best past is in dit geval het tweede plaatje, rechts bovenin. Je moet dan vakje twee aankruisen.

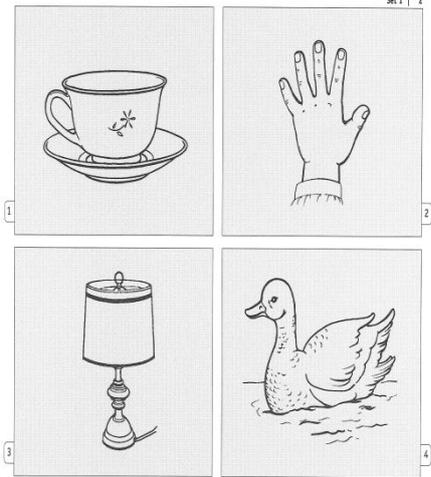
0. Kruis het vakje aan wat hoort bij het Engelse woord:		ball	
			
		1	X 2
		3	4

Wanneer je per ongeluk het verkeerde vakje hebt aangekruist, of wanneer je van gedachte verandert, dan zet je een cirkel om het kruis wat verkeerd is. Hieronder staat een voorbeeld:

0. Kruis het vakje aan wat hoort bij het Engelse woord:		ball	
			
		1	X 2
		3	⊗ 4

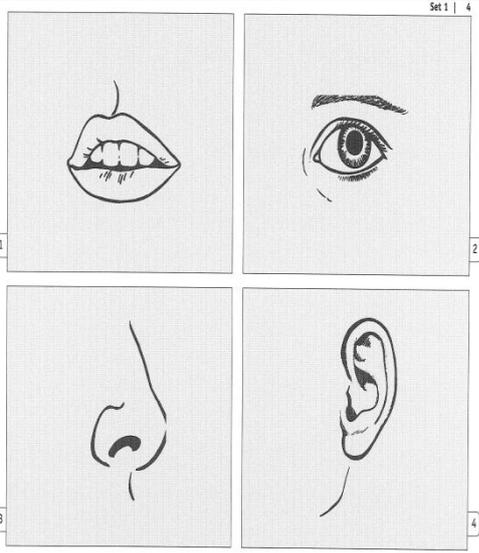
Je mag nu nog vragen stellen, wanneer we begonnen zijn mag je niets meer vragen aan medeleerlingen of je leerkracht.

1. Kruis het vakje aan wat hoort bij het Engelse woord: **hand**



1	2
3	4

2. Kruis het vakje aan wat hoort bij het Engelse woord: **eye**



1	2
3	4

3. Kruis het vakje aan wat hoort bij het Engelse woord: **present**

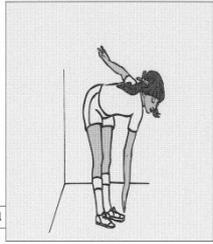


1	2
3	4

4. Kruis het vakje aan wat hoort bij het Engelse woord: **jumping**

↑ Set 17 | 204

Set 2 | 14



1



2



3



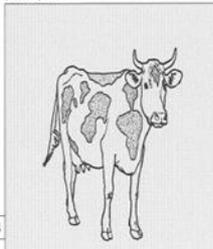
4

1	2
3	4

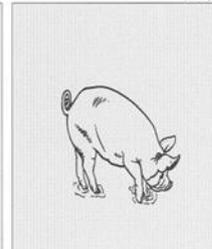
5. Kruis het vakje aan wat hoort bij het Engelse woord: **cow**

↑ Set 17 | 204

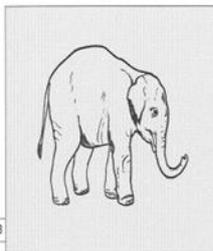
Set 2 | 14



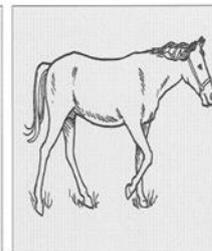
1



2



3



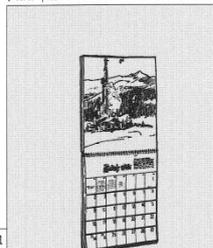
4

1	2
3	4

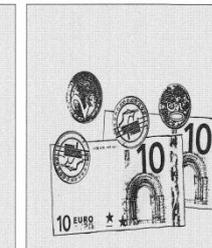
6. Kruis het vakje aan wat hoort bij het Engelse woord: **money**

↑ Set 16 | 187

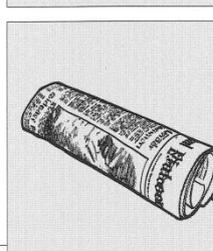
Set 3 | 31



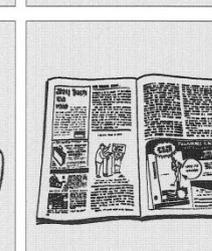
1



2



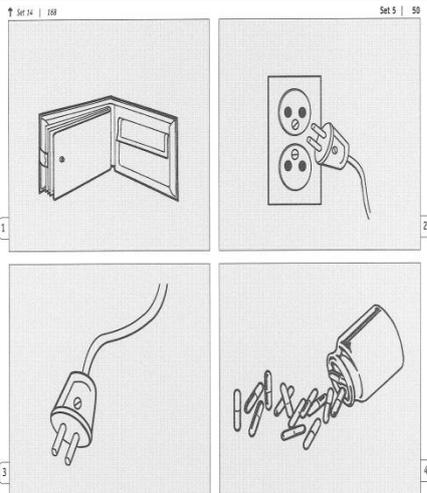
3



4

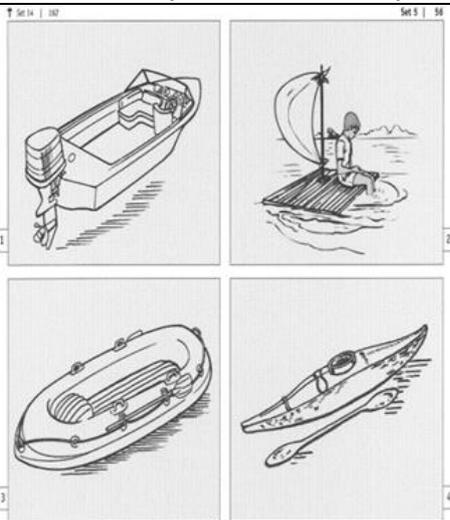
1	2
3	4

7. Kruis het vakje aan wat hoort bij het Engelse woord: **socket**



1	2
3	4

8. Kruis het vakje aan wat hoort bij het Engelse woord: **raft**



1	2
3	4

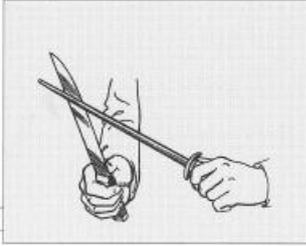
9. Kruis het vakje aan wat hoort bij het Engelse woord: **deliver**



1	2
3	4

10. Kruis het vakje aan wat hoort bij het Engelse woord: **pour**

Set 25 | 180



1

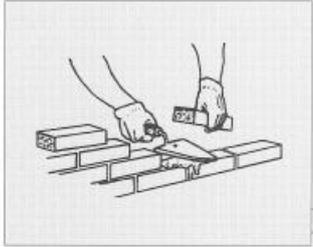
Set 3 | 32



2



3

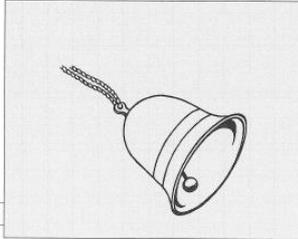


4

1	2
3	4

11. Kruis het vakje aan wat hoort bij het Engelse woord: **tin**

Set 25 | 180

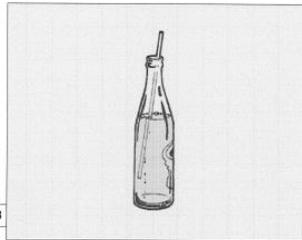


1

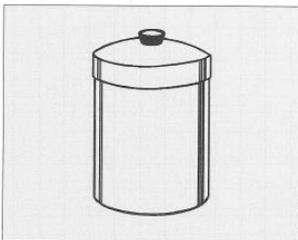
Set 4 | 38



2



3

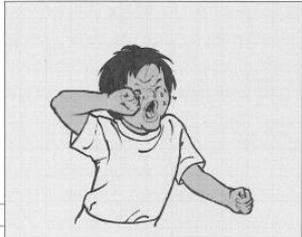


4

1	2
3	4

12. Kruis het vakje aan wat hoort bij het Engelse woord: **scared**

Set 25 | 178



1

Set 4 | 40



2



3



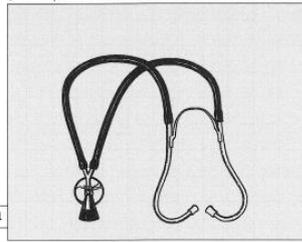
4

1	2
3	4

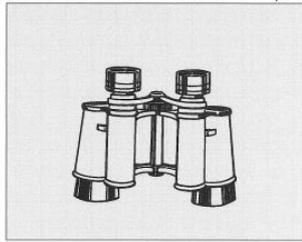
13. Kruis het vakje aan wat hoort bij het Engelse woord: **binoculars**

↑ Set 15 | 277

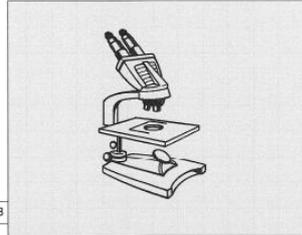
Set 4 | 41



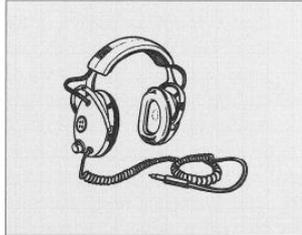
1



2



3



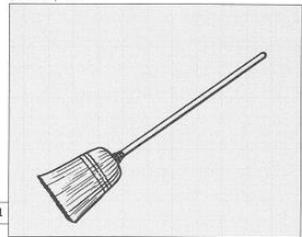
4

1	2
3	4

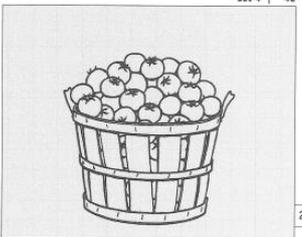
14. Kruis het vakje aan wat hoort bij het Engelse woord: **broom**

↑ Set 15 | 276

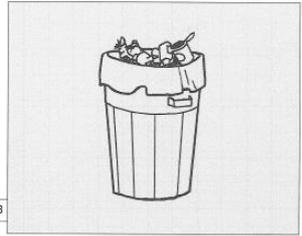
Set 4 | 42



1



2



3



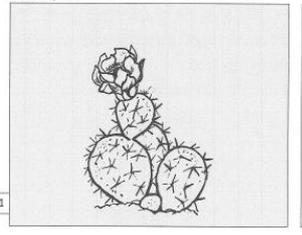
4

1	2
3	4

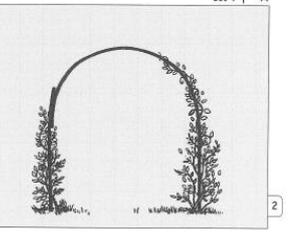
15. . Kruis het vakje aan wat hoort bij het Engelse woord: **bush**

↑ Set 15 | 274

Set 6 | 44



1



2



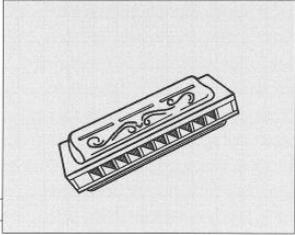
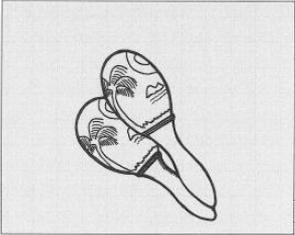
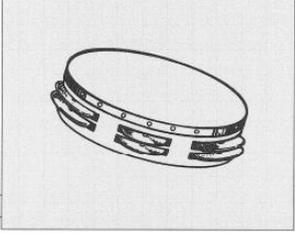
3



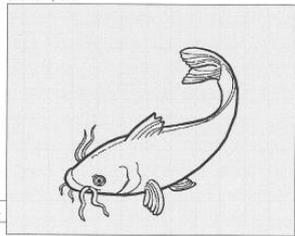
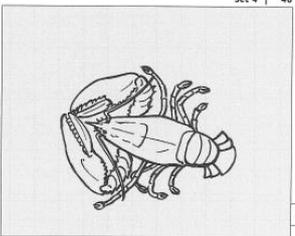
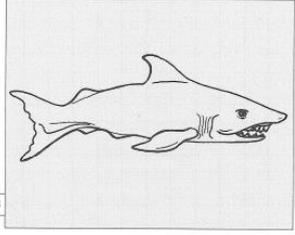
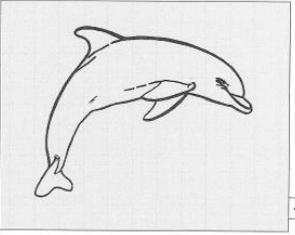
4

1	2
3	4

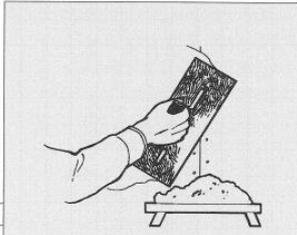
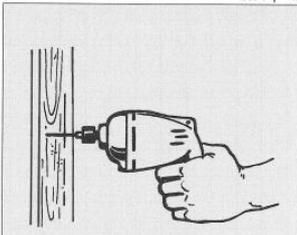
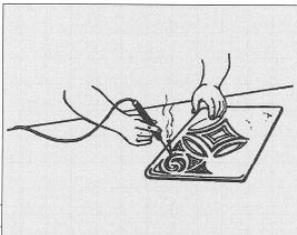
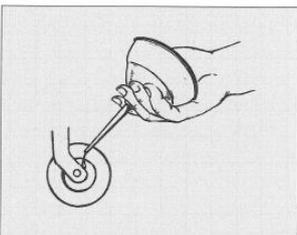
16. . Kruis het vakje aan wat hoort bij het Engelse woord: mouth organ

↑ Set 15   173  1	Set 4   45  2	<table border="1"> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">3</td> <td style="text-align: center;">4</td> </tr> </table>	1	2	3	4
1	2					
3	4					
 3	 4					

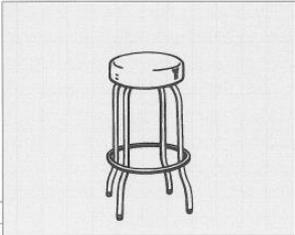
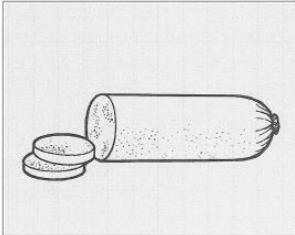
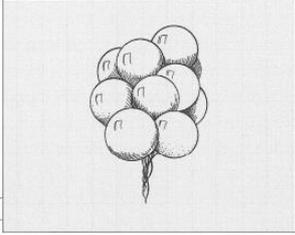
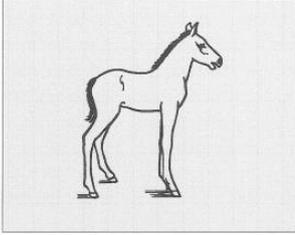
17. . Kruis het vakje aan wat hoort bij het Engelse woord: shark

↑ Set 15   172  1	Set 4   46  2	<table border="1"> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">3</td> <td style="text-align: center;">4</td> </tr> </table>	1	2	3	4
1	2					
3	4					
 3	 4					

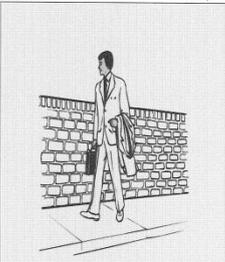
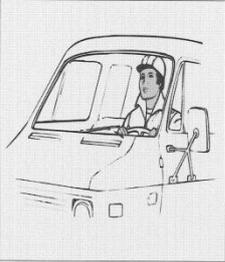
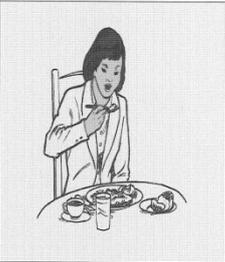
18. . Kruis het vakje aan wat hoort bij het Engelse woord: drill

↑ Set 15   172  1	Set 4   47  2	<table border="1"> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">3</td> <td style="text-align: center;">4</td> </tr> </table>	1	2	3	4
1	2					
3	4					
 3	 4					

19. Kruis het vakje aan wat hoort bij het Engelse woord: **stool**

		<table border="1"> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">3</td> <td style="text-align: center;">4</td> </tr> </table>	1	2	3	4
1	2					
3	4					
						

20. Kruis het vakje aan wat hoort bij het Engelse woord: **pedestrian**

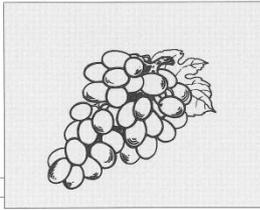
		<table border="1"> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">3</td> <td style="text-align: center;">4</td> </tr> </table>	1	2	3	4
1	2					
3	4					
						

21. Kruis het vakje aan wat hoort bij het Engelse woord: **repair**

		<table border="1"> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2</td> </tr> <tr> <td style="text-align: center;">3</td> <td style="text-align: center;">4</td> </tr> </table>	1	2	3	4
1	2					
3	4					
						

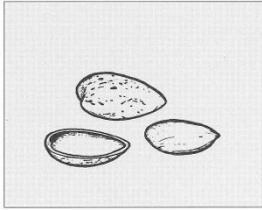
22. Kruis het vakje aan wat hoort bij het Engelse woord: **vegetable**

↑ Set 22 | 142

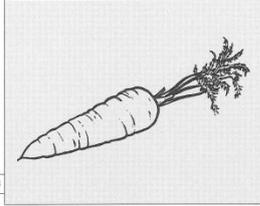


1

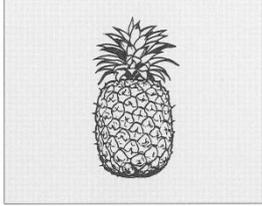
Set 7 | 76



2



3



4

1	2
3	4

23. Kruis het vakje aan wat hoort bij het Engelse woord: **marry**

↑ Set 23 | 139



1

Set 7 | 79



2



3



4

1	2
3	4

24. Kruis het vakje aan wat hoort bij het Engelse woord: **castle**

↑ Set 17 | 133



1

Set 9 | 99



2



3



4

1	2
3	4

25. Kruis het vakje aan wat hoort bij het Engelse woord:

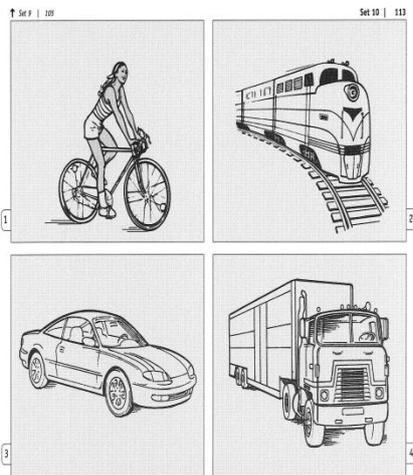
**heel**



1	2
3	4

26. Kruis het vakje aan wat hoort bij de Engelse woorden:

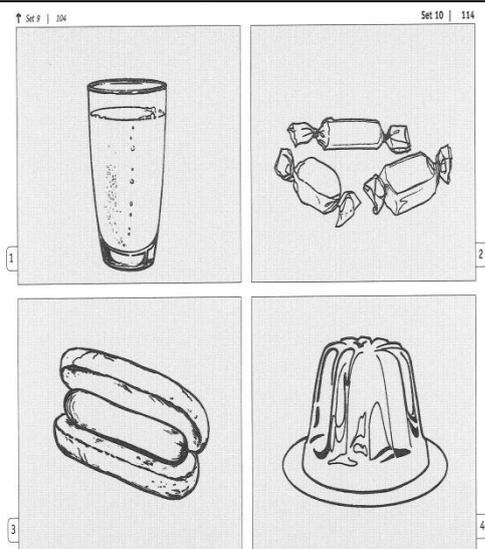
**public transport**



1	2
3	4

27. Kruis het vakje aan wat hoort bij het Engelse woord:

**liquid**



1	2
3	4

28. Kruis het vakje aan wat hoort bij het Engelse woord:

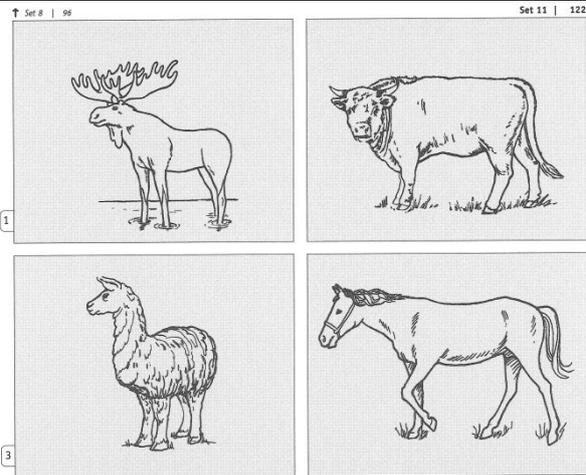
**competitive**



1	2
3	4

29. Kruis het vakje aan wat hoort bij het Engelse woord:

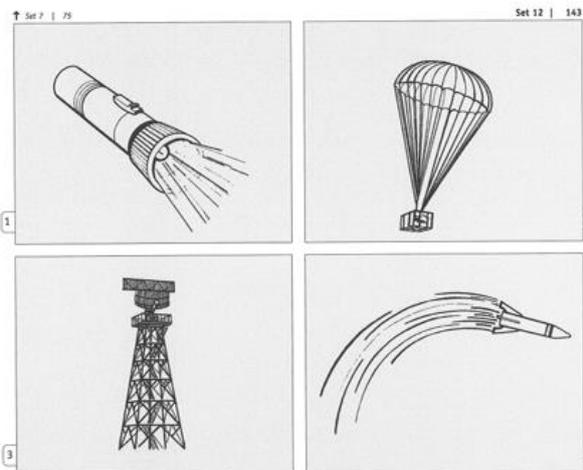
**ox**



1	2
3	4

30. Kruis het vakje aan wat hoort bij het Engelse woord:

**projectile**



1	2
3	4

31. Kruis het vakje aan wat hoort bij het Engelse woord:

**culinary**

↑ Set 7 | 74

Set 12 | 144



1



2



3



4

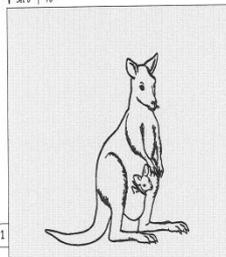
1	2
3	4

32. Kruis het vakje aan wat hoort bij het Engelse woord:

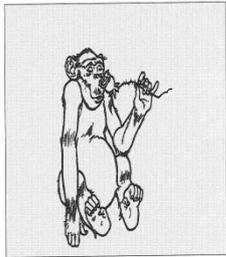
**primate**

↑ Set 6 | 70

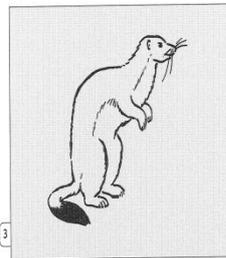
Set 13 | 148



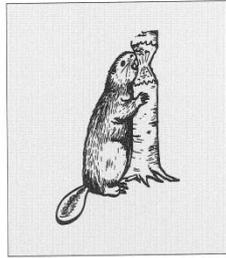
1



2



3



4

1	2
3	4

33. Kruis het vakje aan wat hoort bij het Engelse woord:

**property**

↑ Set 6 | 64

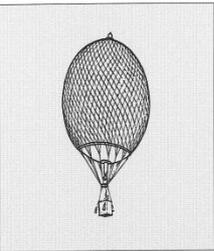
Set 13 | 154



1



2



3



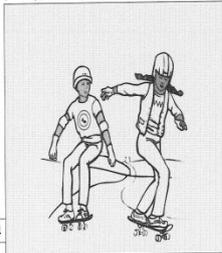
4

1	2
3	4

34. Kruis het vakje aan wat hoort bij het Engelse woord:

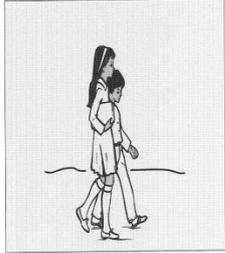
**stroll**

↑ Set 5 | 15

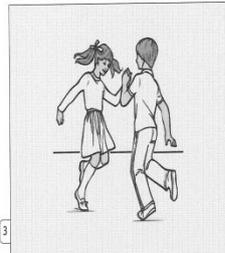


1

Set 14 | 163



2



3



4

1	2
3	4

35. Kruis het vakje aan wat hoort bij het Engelse woord:

**assembly**

↑ Set 9 | 53



1

Set 14 | 165



2



3



4

1	2
3	4

36. Kruis het vakje aan wat hoort bij het Engelse woord:

**melancholy**

↑ Set 4 | 46



1

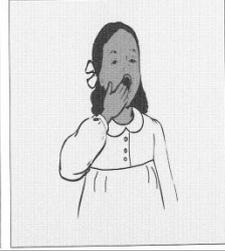
Set 15 | 172



2



3



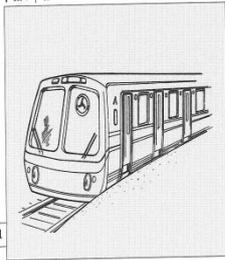
4

1	2
3	4

37. Kruis het vakje aan wat hoort bij het Engelse woord:

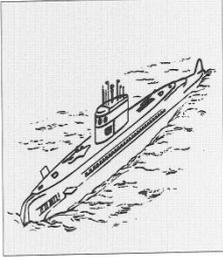
**aviator**

↑ Set 4 | 39

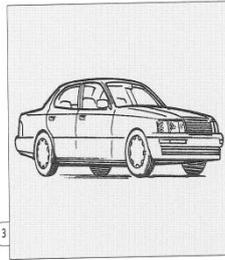


1

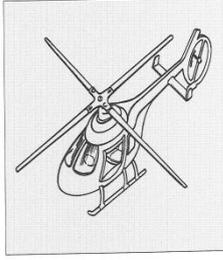
Set 15 | 179



2



3



4

1	2
3	4

38. Kruis het vakje aan wat hoort bij het Engelse woord:

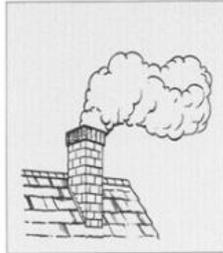
**blaze**

↑ Set 4 | 38



1

Set 15 | 180



2



3



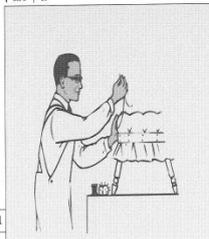
4

1	2
3	4

39. Kruis het vakje aan wat hoort bij het Engelse woord:

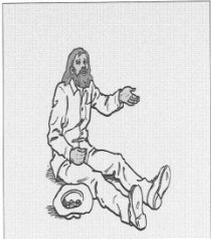
**padded**

↑ Set 2 | 29

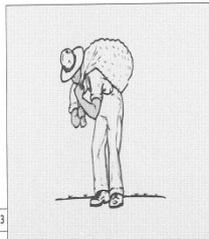


1

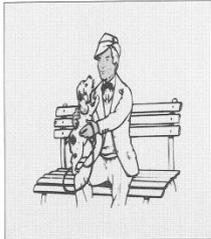
Set 17 | 198



2



3



4

1	2
3	4

40. Kruis het vakje aan wat hoort bij het Engelse woord:

**lament**

Set 17 | 17



1

Set 17 | 201



2



3



4

1	2
3	4

**6. Dictation used as pre- and post-test language skills measurement instrument during the experimental phase.**

1. Hello, my name is peter
2. I am 14 years old.
3. He lives in London.
  
4. His sister always walks to school.
5. She likes going to the zoo
  
6. Peter is playing games on the computer.
7. Janet is riding her bicycle outside.
  
8. Where were they staying when their mom went away?
9. Every day they eat two toast, fried eggs and sausages.
  
10. My teacher says I am good at Geography because I read a lot.

**7. Editing test texts used as pre- and post-test language skills measurement instrument during the experimental phase (added words in bold and italics).**

**Version A**

Peter is a young boy. He is seven years old. He lives in a **put** big house with his brothers and sisters. Peter has four brothers and five sisters. He **new** is the youngest of the ten children in the house **take**. His mother and father also have nine brothers and sisters. All of **right** them have ten children. When there is a party for the family there **between** are a lot of children.

Janet loves horses. Every day she goes **because** to her horse. She feeds her horse, cleans it and rides it. Her horse **until** is brown and white. It likes to walk to the beach. On the **don't** beach it likes to run very fast. Janet wears special **always** clothes when she rides. She has special trousers, boots and a helmet. At home Janet **set** has a lot of things with horses. On the door of **same** her room is an enormous poster with a horse.

Fred plays **went** soccer. Every day he goes to a field to play with his friends. They make **end** a goal of their jackets and kick the ball around. When **sound** they play together they forget the time. They are always late for dinner. On **take** Saturday he goes to his club to play. Every Saturday he plays a match. His **does** team is very good. Next week they play the last **while** match. If they win the match they will be champions.

Sheila goes shopping. She wants **give** new clothes and a bag. She takes a train to the city **school**. In ten minutes Sheila walks to the shop. In the shop she tries a **write** lot of clothes. The clothes are beautiful. The clothes **any** cost a lot of money. Sheila buys all the clothes. Now she wants a **too** bag. She sees beautiful bags. Sheila looks at her money. The bags cost **important** a lot of money. Sheila has no money for a bag **air**. She walks to the train. She takes the train home.

Jenny and Tom are going **until** to eat a special dinner. They start with small sandwiches. On the sandwiches are eggs **those**, mayonnaise and cheese. They drink fruit juice from big glasses. After the sandwiches they eat **think** soup. Then they have spaghetti with small meat balls and tomatoes. The last thing they **house** eat is an ice-cream. They have chocolate sauce and strawberries on their **say** ice-cream. Jenny is full she can eat no more. Tom wants another ice-cream **want**.

## 7. Editing test texts used as pre- and post-test language skills measurement instrument during the experimental phase (added words in bold and italics).

### Version B

#### The fireboy

This is the story of Hapu. He lives **world** in Egypt at the time **three** of Queen Cleopatra. His father has a **write** small factory. He makes gold **day** tables and chairs. Hapu helps him. He is a 'fireboy'. Every day Hapu keeps the **different** fire burning. The work is hot and very hard. But what can **big** he and his father do? They are poor **along** people. They need to make money.

Every evening they sit beside the Nile **under**. His father talks to Hapu about the work, the factory **know**, the chairs, tables - everything. He tells Hapu they will be rich one day. He doesn't **school** know when or how, but they will be rich. Then they stand up. It is **asked** late and time to go **large** home. Three days later Hapu and his father are making a table. Suddenly **might** his father stops working. He has a pain in his head. He is **until** ill and can no longer work. Hapu brings his father to his bed. The next **give** day Hapu goes to work alone. His **between** father cannot go to work. He is ill.

Hapu is working alone **children**. They need money, but what can he do? Then Hapu has an idea. He **write** has to make something for someone **think** very rich. Somebody like the Queen.

Hapu's father is ill **come** for two weeks. Hapu makes **why** chairs and tables. Hapu makes a gold necklace too **number**. Hapu thinks Queen Cleopatra is going to like the necklace. Queen Cleopatra **few** will pay a lot of money for the necklace look.

When the necklace is ready he puts it in his **sound** bag. He takes the necklace to the palace **any**. There are two guards in front of the palace. They both have a sword. They look dangerous. They ask what Hapu wants to **home** do. Hapu says he wants to see Queen Cleopatra. The guards say **keep** Queen Cleopatra does not want to talk to boys.

Hapu sits outside the Palace and **food** waits. Later in the afternoon suddenly the doors open. Queen Cleopatra comes out **important** of the Palace to go **want** to the Nile. She is not alone. There are a lot **going** of people with Queen Cleopatra. Hapu runs to the Queen and holds his hand up. **together** Stop! Hapu shouts. Please, stop! What do you want? Queen Cleopatra asks. Hapu **saw** gives her the necklace. The **thought** queen looks at the necklace. Did you make this? Queen Cleopatra **both** asks. Hapu tells her about the factory and his father. He tells her he made **under** the necklace for her. He tells her they need the money.

Queen Cleopatra **line** likes the necklace. She gives **went** Hapu a lot of money. Hapu and his father are rich.



**9. Background information of students in the sample**

**1. Wat is je naam:**

**2. Wat is je geboortedatum? Vul in.**

(dag)	(maand)	(jaar)
-------	---------	--------

**3. Ben je een jongen of een meisje? Kruis aan.**

- een jongen
- een meisje

**4. Waar ben je geboren? Kruis aan/vul in.**

- in Nederland
- in Suriname
- in Turkije
- op Aruba/de Nederlandse Antillen
- in Marokko
- ergens anders, namelijk \_\_\_\_\_

**5. Waar is je moeder geboren? Kruis aan/vul in.**

- in Nederland
- in Suriname
- in Turkije
- op Aruba/de Nederlandse Antillen
- in Marokko
- ergens anders, namelijk \_\_\_\_\_

**6. Waar is je vader geboren? Kruis aan/vul in.**

- in Nederland
- in Suriname
- in Turkije
- op Aruba/de Nederlandse Antillen
- in Marokko
- ergens anders, namelijk \_\_\_\_\_

**7. Welke taal spreek je meestal met je moeder? Kruis aan/vul in.**

- meestal Nederlands
- meestal een andere taal, namelijk \_\_\_\_\_
- evenveel Nederlands als een andere taal, namelijk \_\_\_\_\_
- niet van toepassing

**8. Welke taal spreek je meestal met je vader? Kruis aan/vul in.**

- meestal Nederlands
- meestal een andere taal, namelijk \_\_\_\_\_
- evenveel Nederlands als een andere taal, namelijk \_\_\_\_\_

niet van toepassing

**9. Op welke school zit je nu?:** \_\_\_\_\_

**10. In welke klas zit je nu?:** \_\_\_\_\_

**11. Van welke basisschool kom je:** \_\_\_\_\_

**12. Wat was je Cito-score in groep 8?:** \_\_\_\_\_

**13. Op mijn basisschool heb ik Engels gehad in groep(en)** Kruis aan/vul in.

alleen in groep 8	<input type="checkbox"/>	van groep 1 t/m 8	<input type="checkbox"/>
in groep 7 en 8	<input type="checkbox"/>	geen Engels gehad	<input type="checkbox"/>
in groep 5 t/m 8	<input type="checkbox"/>	Anders, namelijk	_____

**14. Hoeveel Engelstalige TV kijk je?** Kruis aan/vul in.

nooit	<input type="checkbox"/>	iedere dag	<input type="checkbox"/>
1 x per week	<input type="checkbox"/>	(meerdere) uren per dag	<input type="checkbox"/>
bijna iedere dag	<input type="checkbox"/>	Anders, namelijk	_____

**15. Hoeveel Engelstalige spellen speel je** (op een PC, iPad, telefoon, etc.)? Kruis aan/vul in.

nooit	<input type="checkbox"/>	iedere dag	<input type="checkbox"/>
1 x per week	<input type="checkbox"/>	(meerdere) uren per dag	<input type="checkbox"/>
bijna iedere dag	<input type="checkbox"/>	Anders, namelijk	_____

**16. Hoe vaak lees je een Engels boek?** Kruis aan/vul in.

nooit	<input type="checkbox"/>	maandelijks	<input type="checkbox"/>
1 x per jaar	<input type="checkbox"/>	wekelijks	<input type="checkbox"/>
iedere 3 maanden	<input type="checkbox"/>	Anders, namelijk	_____

**10. Tables Chapter 6:**

**10.A. Confirmatory factor analyses MPB pre- and post-test groups;**

**Table AP10A1: Results of the model fits: Original model: Cognition=1-6; Affect=7-11; Subj. Norm=12-15; PBC=17-21; Intention=22-26. (pre-test N=133; post-test N=127)**

model	$\chi^2$	df	p	RMSEA (90% C.I.)	CFI	TLI	SRMR	Deleted items
Pre-test model 1	490.809	265	.000	.080 (.069-.091)	.791	.764	.086	16
Pre-test model 2	422.493	242	.000	.075 (.063-.087)	.824	.799	.081	3, 16
Pre-test model 3	357.564	220	.000	.069 (.055-.081)	.845	.822	.068	3, 7, 16,
Pre-test model 4	308.808	199	.000	.064 (.050-.078)	.864	.842	.066	3, 6, 7, 16,
Pre-test model 4 plus MPB	312.236	202	.000	.064 (.050-.078)	.863	.843	.067	3, 6, 7, 16
Difference pre-test model 4 with and without MPB	3.428	3	n.s.					
Post-test model 1	601.093	265	.000	.100 (.089-.111)	.753	.720	.108	16
Post-test model 4	355.101	199	.000	.079 (.065-.092)	.843	.817	.074	3, 6, 7, 16
Post-test model 4 plus MPB	356.611	202	.000	.078 (.064-.091)	.844	.822	.075	3, 6, 7, 16
Difference post-test model 4 with and without MPB	1.51	3	n.s.					
<b>Categorical CFA</b>								
Pre-test model 4	265.036	199	.001	.050 (.032-.065)	.945	.937	-	3, 6, 7, 16
Pre-test model 4 plus MPB	270.504	202	.001	.050 (.033-.066)	.943	.935		3, 6, 7, 16
Difference pre-test model 4 with and without MPB	7.061	3	.07				-	
Post-test model 4	288.720	199	.000	.060 (.044-.074)	.949	.941	-	3, 6, 7, 16
Post-test model 4 plus MPB	290.722	202	.000	.059 (.043-.073)	.950	.942	-	3, 6, 7, 16
Difference post-test model 4 with and without MPB	3.513	3	.319					

---

**Structural model (MPB) pre-test:**

---

	est.	SE	est. SE	p-value
INT ON				
AFF	0.532	0.104	5.095	0.000
SUN	0.291	0.154	1.892	0.058
PBC	0.132	0.121	1.097	0.273
AFF ON				
COG	0.650	0.077	8.476	0.000
SUN WITH				
COG	1.016	0.052	19.561	0.000
PBC WITH				
COG	0.845	0.052	16.384	0.000
SUN	0.671	0.080	8.418	0.000

**Structural model (MPB) post-test:**

INT ON				
AFF	0.569	0.123	4.619	0.000
SUN	0.200	0.125	1.593	0.111
PBC	0.287	0.103	2.778	0.005
AFF ON				
COG	0.707	0.064	11.127	0.000
SUN WITH				
COG	0.844	0.041	20.550	0.000
PBC WITH				
COG	0.784	0.056	14.024	0.000
SUN	0.587	0.073	8.044	0.000

---

**10.B. MI analyses to verify whether pre-test scores for MPB and language skills between experimental and control group are comparable at the onset of the study.**

**Table AP10B1: Results multi-level analyses testing differences for ‘Cognition’ at pre-test for experimental and control pupils (student N=133; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3
<b>Fixed part</b>			
Intercept	4.056*** (.092)	4.045*** (.152)	4.025*** (.155)
Condition (exp.=1; contr.=0)	.056 (.127)	.046 (.212)	.068 (.131)
<b>Random part</b>			
Pupil variance	.538 (.066)	.493 (.062)	.494 (.062)
Class variance		.045 (.039)	.003 (.021)
School variance			.045 (.048)
Deviance	295.069	290.073	288.830
Ref. model		1	2
Fit improvement		X <sup>2</sup> =4.996 df=1	X <sup>2</sup> =1.243 df=1
		p<.05	p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table AP10B2: Results multi-level analyses testing differences for ‘Affect’ at pre-test for experimental and control pupils (student N=133; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3
<b>Fixed part</b>			
Intercept	3.441*** (.115)	3.410*** (.191)	3.402*** (.197)
Condition (exp.=1; contr.=0)	.116 (.158)	.117 (.267)	.129 (.155)
<b>Random part</b>			
Pupil variance	.831 (.102)	.759 (.095)	.757 (.095)
Class variance		.072 (.062)	.001 (.029)
School variance			.078 (.080)
Deviance	352.865	347.531	345.666
Ref. model		1	2
Fit improvement		X <sup>2</sup> =5.334 df=1	X <sup>2</sup> =1.865 df=1
		p<.05	p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table AP10B3: Results multi-level analyses testing differences for ‘Subjective Norm’ at pre-test for experimental and control pupils (student N=133; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2
<b>Fixed part</b>		
Intercept	3.738*** (.097)	3.738*** (.097)
Condition (exp.=1; contr.=0)	.108 (.134)	.108 (.134)
<b>Random part</b>		
Pupil variance	.598 (.073)	.598 (.073)
Class variance		.000 (.000)
School variance		
Deviance	309.148	309.148
Ref. model		
Fit improvement		X <sup>2</sup> =.000 df=1
		p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table AP10B4: Results multi-level analyses testing differences for ‘PBC’ at pre-test for experimental and control pupils (student N=133; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3
<b>Fixed part</b>			
Intercept	3.822*** (.093)	3.803*** (.160)	3.787*** (.170)
Condition (exp.=1; contr.=0)	.192 (.129)	.193 (.224)	.204# (.122)
<b>Random part</b>			
Pupil variance	.548 (.067)	.497 (.062)	.489 (.061)
Class variance		.053 (.044)	.000 (.000)
School variance			.063 (.060)
Deviance	297.351	291.687	287.891
Ref. model		1	2
Fit improvement		X <sup>2</sup> =5.664 df=1	X <sup>2</sup> =3.796 df=1
		p<.05	p<.10 <sup>a</sup>

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

a = Since we are testing whether adding a variance level improves model fit, the probability should be divided by 2, meaning a significant fit improvement at  $p < .05$ .

**Table AP10B5: Results multi-level analyses testing differences for ‘Intentions’ at pre-test for experimental and control pupils (student N=133; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3
<b>Fixed part</b>			
Intercept	3.689*** (.109)	3.663*** (.179)	3.656*** (.192)
Condition (exp.=1; contr.=0)	.037 (.150)	.043 (.251)	.042 (.143)
<b>Random part</b>			
Pupil variance	.743 (.091)	.684 (.086)	.669 (.083)
Class variance		.063 (.054)	.000 (.000)
School variance			.078 (.076)
Deviance	338.005	333.483	329.350
Ref. model		1	2
Fit improvement		X <sup>2</sup> =4.522	X <sup>2</sup> =4.133
		df=1	df=1
		p<.05	p<.05

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

### Language skills

**Table AP10B6: Results multi-level analyses testing differences for ‘EdCor’ at pre-test for experimental and control pupils (student N=145; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3
<b>Fixed part</b>			
Intercept	-.038 (.118)	-.159 (.381)	-.159 (.381)
Condition (exp=1; contr=0)	.074 (.165)	.160 (.131)	.160 (.131)
<b>Random part</b>			
Pupil variance	.985 (.116)	.615 (.073)	.615 (.073)
Class variance		.408 (.343)	.000 (.000)
School variance			.408 (.343)
Deviance	409.278	351.456	351.456
Ref. model		1	2
Fit improvement		X <sup>2</sup> =57.822	X <sup>2</sup> =.000
		df=1	df=1
		p<.001	p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table AP10B7: Results multi-level analyses testing differences for ‘EdInc’ at pre-test for experimental and control pupils (student N=145; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3
<b>Fixed part</b>			
Intercept	.186 (.116)	.329 (.334)	.324 (.333)
Condition (exp.=1; contr.=0)	-.364* (.162)	-.484 (.471)	-.475 (.298)
<b>Random part</b>			
Pupil variance	.953 (.112)	.699 (.084)	.699 (.084)
Class variance		.303 (.192)	.104 (.108)
School variance			.198 (.223)
Deviance	404.527	374.043	372.739
Ref. model		1	2
Fit improvement		X <sup>2</sup> =30.484	X <sup>2</sup> =1.304
		df=1	df=1
		p<.001	p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table AP10B8: Results multi-level analyses testing differences for ‘vocabulary’ at pre-test for experimental and control pupils (student N=145; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3
<b>Fixed part</b>			
Intercept	23.042*** (.532)	22.802*** (1.180)	22.686*** (1.210)
Condition (exp.=1; contr.=0)	-.704 (.745)	-.579 (1.664)	-.457 (.683)
<b>Random part</b>			
Pupil variance	20.120 (2.363)	16.957 (2.035)	16.804 (1.994)
Class variance		3.432 (2.383)	.000 (.000)
School variance			3.664 (3.268)
Deviance	846.742	832.514	827.934
Ref. model		1	2
Fit improvement		X <sup>2</sup> =14.228 df=1 p<.001	X <sup>2</sup> =4.580 df=1 p<.05

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table AP10B9: Results multi-level analyses testing differences for ‘dictation’ at pre-test for experimental and control pupils (student N=145; class N=6; school N=3) (all continuous predictors gm) (s.e. between brackets)**

model	1	2	3	4	5
<b>Fixed part</b>					
Intercept	47.986*** (1.522)	46.329*** (5.370)	46.147*** (5.452)	47.758*** (5.330)	46.147*** (5.452)
Condition (exp=1; contr=0)	2.028 (2.131)	2.868 (7.588)	3.134* (1.578)		3.134* (1.578)
<b>Random part</b>					
Pupil variance	164.510 (19.321)	90.901 (10.906)	89.666 (10.642)	92.201 (10.943)	89.666 (10.642)
Class variance		82.491 (49.680)	.000 (.000)		
School variance			85.298 (71.024)	83.283 (69.432)	85.298 (71.024)
Total variance			174.964	175.484	174.964
Deviance	1151.423	1084.116	1074.938	1078.828	1074.938
Ref. model		1	2		4
% expl. pupil var.					2.479%
% expl. school var.					-
% expl. total var.					.361%
Fit improvement		X <sup>2</sup> =67.307 df=1 p<.001	X <sup>2</sup> =9.178 df=1 p<.01		X <sup>2</sup> =3.890 df=1 p<.05

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

## 10.C

**Table 6.6: correlations between sums of post-test items**

<b>N=127</b>	<b>Sumcogpost-cfa</b>	<b>Sumaffpost-cfa</b>	<b>Sumsunpost-cfa</b>	<b>sumPBCpost-cfa</b>	<b>Sumintpost-cfa</b>
sumcogpost	,971**	,518**	,677**	,598**	,616**
sumaffpost	,444**	,977**	,426**	,442**	,587**
sumsunpost	,655**	,428**	1,000**	,477**	,499**
sumPBCpost	,615**	,448**	,477**	1,000**	,491**
sumintpost	,594**	,576**	,499**	,491**	1,000**

**Table 6.7: correlations between sums of pre-test items**

<b>N=133</b>	<b>sumcogpreca</b>	<b>sumaffpreca</b>	<b>sumsunpreca</b>	<b>sumPBCpreca</b>	<b>sumintpreca</b>
sumcogpre	,945**	,477**	,631**	,569**	,480**
sumaffpre	,404**	,981**	,391**	,458**	,577**
sumsunpre	,619**	,379**	1,000**	,459**	,467**
sumPBCpre	,575**	,468**	,459**	1,000**	,462**
sumintpre	,485**	,566**	,467**	,462**	1,000**

### Effect of the treatment on changes in scores for Cognition

In table-6.8 model 1 and 2 are used to determine the variance levels needed, model 3 is used to determine the effect of the treatment ('condition') without controlling for covariates. In model 4 to 7 covariates are added that have shown significant correlations with the post-test scores or with the change scores (post-test score minus pre-test score). Covariates are added one at a time (for ordinal or nominal variables these are the dummies made from the variable) to verify whether these covariates are significant predictors in the regression model with a pre-test as a covariate.

The significance of the variable 'condition' is checked both after correcting for each covariate that appears to be significant in the regression equation and after correcting for all significant covariates simultaneously. Covariates showing significant correlations with post-test or change scores for Cognition were: *country of birth of the mother* and *country of birth of the father*. Model 4 adds the dummies for country of birth of the mother (reference group is 'country of birth of the mother = The Netherlands'), model 5 adds the 'condition' variable to model 4, model 6 and 7 do the same for country of birth of the father (reference group is 'The Netherlands'), model 8 includes the dummies both for country of birth mother and father and in the final model 9 the condition variable is added to model 8 to estimate the effect of the treatment after controlling for the significant covariates.

## Effect of the treatment on changes in scores for Affect

In table 6.9.1 and 6.9.2 the results of the regression analyses for Affect are presented in a comparable fashion. Again, adding a class level does not significantly improve model fit (model 2), so no class level is needed. The significant covariates are Eibo<sup>44</sup>, time spent watching television, playing games and reading English books and the students' listening/spelling proficiency. Time spent watching television, playing games and reading English books are each measured on five point Likert scales (1 = never; 5= several hours per day).

Again the treatment (condition) has a significant effect ( $p < .05$ ) on the change in Affect both when not correcting for significant covariates (model 3) and after correcting for all significant covariates (model 15 and 17). When not correcting for covariates, the variable condition explains 3.58% of the variance in change in Affect and the difference between the means for control and experimental group amounts to .347 ( $p < .05$ ). After controlling for all significant covariates (model 15) the difference is still .323 ( $p < .05$ ) and condition still explains 3.81% of the differences in change of Affect from pre- to post-test. Significance of covariates is first always determined with one covariate (or all its dummies) in the model. Covariates that were significant when added separately to the model, sometimes become non-significant when added together with all covariates that were significant. In model 16 and 17 (see table 6.9.2) the two covariates that became non-significant when all covariates were added (Time Playing games and Reading English), are removed. The results remain more or less identical. There is a positive effect of the treatment on the development of Affect between pre- and post-test whether controlled for the significant covariates or not.

---

<sup>44</sup> 'Eibo' is a Dutch acronym for: 'Engels in het basis onderwijs', literally 'English in primary education'; in this questionnaire used as variable indicating years of English language teaching received (choices were: no English; 1 year; 2 years; 4 years or 8 years).

**Table 6.8: Results multi-level analyses attitude scores Cognition, effect of the experiment without and with controlling for background variables (student N=123; class N=6) (pre-test scores grand mean centred) (s.e. between brackets)**

model	1	2	3	4	5	6	7	8	9
<b>Fixed part</b>									
Intercept	4.092 (.077)	4.096 (.099)	3.869 (.109)	4.125 (.102)	3.955 (.128)	4.209 (.109)	4.014 (.139)	4.171 (.107)	4.004 (.137)
Pre-test	.073 (.104)	.069 (.105)	.050 (.101)	.065 (.102)	.052 (.100)	.083 (.102)	.061 (.100)	.071 (.099)	.054 (.098)
Condition (exp=1; contr=0)			.414* (.149)		.323* (.151)		.325* (.149)		.287# (.149)
<b>Country of birth mother</b>									
Turkey				.179 (.306)	.069 (.305)			.086 (.495)	-.027 (.492)
Morocco				-.436* (.206)	-.386# (.204)			-.694 (.490)	-.723 (.483)
Surinam				-.222 (.277)	-.181 (.272)			.809 (.485)	.790 (.477)
Aruba				.719 (.488)	.575 (.484)			.440 (.653)	.220 (.654)
other country				.229 (.228)	.197 (.224)			.605 (.349)	.492 (.349)
<b>Country of birth father</b>									
Turkey						.008 (.258)	-.040 (.254)	.047 (.424)	.077 (.418)
Morocco						-.451* (.206)	-.370 (.206)	.219 (.490)	.312 (.485)
Surinam						-.572* (.267)	-.495 (.265)	-1.196 (.467)	-1.118* (.462)
Aruba						.507 (.352)	.494 (.346)	.229 (.477)	.340 (.473)
other country.						-.084 (.230)	-.052 (.226)	-.537 (.351)	-.422 (.351)
<b>Random part</b>									
Students variance	.720 (.092)	.696 (.091)	.678 (.086)	.664 (.085)	.640 (.082)	.655 (.084)	.631 (.080)	.614 (.078)	.596 (.076)
Class variance		.024 (.034)							
Deviance	308.714	307.675	301.270	298.627	294.134	297.025	292.331	289.130	285.496
Ref. model		1	1	1	4	1	6	1	8
% expl. var. students level			5.833%	7.778%	3.614%	9.027%	3.664%	14.722%	2.932%
Fit improvement		$\chi^2 = 1.039$ df=1 p=n.s.	$\chi^2 = 7.444$ df=1 p<.01	$\chi^2 = 10.087$ df=5 p<.10	$\chi^2 = 4.493$ df=1 p<.05	$\chi^2 = 11.689$ df=5 p<.01	$\chi^2 = 4.694$ df=1 p<.05	$\chi^2 = 19.584$ df=10 p<.05	$\chi^2 = 3.634$ df=1 p<.10

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

results of model 1 and 2 show that adding a class level to a model with only a student level does not improve model fit. This means that model 1 is the model of reference for model 3, 4, 6 and 8. Model 3 shows that the treatment group has a significantly larger score on the post-test after controlling for the pre-test than the control group. This difference amounts to .414 and the condition variable explains 5.83% of variance in change in Cognition from pre- to post-test. In model 4 the 'country of birth of the mother' has been added as covariate. Since this variable is measured on a nominal scale with six possible answers, five dummies indicating five different options are added to the regression model. The reference category is 'students whose mother is born in The Netherlands'. The analysis shows that the five dummies together do

not significantly improve model fit ( $p < .10$ ). However, the dummy indicating that the student's mother is born in Morocco has a significant negative effect ( $-.436$ ;  $p < .05$ ). Students with a mother who is born in Morocco compared to students with a mother born in The Netherlands show .436 less development in cognition.

Model 5 shows that adding the variable 'condition' to a model with the dummies for 'country of birth mother' as covariates significantly improves model fit and explains 3.61% of variance in change in Cognition scores. The difference in change between experimental and control group is reduced by these covariates to .323 ( $p < .05$ ). In model 6 and 7 the effect of 'country of birth father' is checked as well as the effect of condition after correcting for country of birth father. Now the five dummies do significantly improve model fit ( $p < .01$ ) and explain 9.03% of total variance in change in Cognition scores. Two out of five regression coefficients are significant. Students with a father born in Morocco show a more negative trend in Cognition scores from pre- to post-test than students with a father born in the Netherlands ( $-.451$ ;  $p < .05$ ) as do students with a father born in Surinam ( $-.572$ ;  $p < .05$ ). After correcting for the country of birth of the father, adding the condition variable to the regression model significantly improves model fit ( $p < .05$ ) and explains 3.66% of variance in change in Cognition scores.

In model 8 all dummies for country of birth of both mother and father are added. These dummies explain 14.72% of change in Cognition scores, which is significant ( $p < .05$ ). None of the dummies shows a significant regression coefficient in this model, which is probably due to collinearity. The final model 9 shows that after correcting for all these dummies, the variable 'condition' is still 1-sided significant at 5% and explains 2.93% of the differences between pre- and post-test for Cognition. The difference between the experimental and the control group in this model amounts to .287. To conclude, this means the treatment shows a positive effect on the change in Cognition and this positive effect remains after controlling for the significant covariates 'country of birth mother' and 'country of birth father'. A noticeable detail of the analyses is that in none of the models fitted the pre-test shows a significant relation to the post-test scores. This means that the rank order of students on Cognition scores at pre-test is not the same rank order at post-test. This absence of a relation between pre- and post-tests does not seem to be caused by the effect

of the treatment, since even in model 9 the pre-test is non-significant. Possibly Cognition scores change relatively much within individuals over time.

### **Effect of the treatment on changes in scores for Affect**

In table 6.9.1 and 6.9.2 the results of the regression analyses for Affect are presented in a comparable fashion. Again, adding a class level does not significantly improve model fit (model 2), so no class level is needed. The significant covariates are Eibo<sup>45</sup>, time spent watching television, playing games and reading English books and the students' listening/spelling proficiency. Time spent watching television, playing games and reading English books are each measured on five point Likert scales (1 = never; 5= several hours per day).

Again the treatment (condition) has a significant effect ( $p < .05$ ) on the change in Affect both when not correcting for significant covariates (model 3) and after correcting for all significant covariates (model 15 and 17). When not correcting for covariates, the variable condition explains 3.58% of the variance in change in Affect and the difference between the means for control and experimental group amounts to .347 ( $p < .05$ ). After controlling for all significant covariates (model 15) the difference is still .323 ( $p < .05$ ) and condition still explains 3.81% of the differences in change of Affect from pre- to post-test. Significance of covariates is first always determined with one covariate (or all its dummies) in the model. Covariates that were significant when added separately to the model, sometimes become non-significant when added together with all covariates that were significant. In model 16 and 17 (see table 6.9.2) the two covariates that became non-significant when all covariates were added (Time Playing games and Reading English), are removed. The results remain more or less identical. There is a positive effect of the treatment on the development of Affect between pre- and post-test whether controlled for the significant covariates or not.

---

<sup>45</sup> 'Eibo' is a Dutch acronym for: 'Engels in het basis onderwijs', literally 'English in primary education'; in this questionnaire used as variable indicating years of English language teaching received (choices were: no English; 1 year; 2 years; 4 years or 8 years).

**Table 6.9.1: Results of the multi-level analyses attitude scores: Affect, effect of the experiment without and with controlling for background variables (student N=123; class N=6) (pre-test and time watching English, playing games and reading English grand mean centred) (s.e. between brackets)**

model	1	2	3	4	5	6	7	8	9	10	11
<b>Fixed part</b>											
Intercept	3.380 (.083)	3.379 (.087)	3.194 (.119)	3.516 (.206)	3.352 (.213)	3.377 (.081)	3.205 (.117)	3.370 (.081)	3.200 (.117)	3.379 (.081)	3.204 (.118)
Pre-test	.177# (.092)	.175# (.092)	.170# (.090)	.198 (.089)	.193 (.087)	.135 (.092)	.131 (.090)	.191* (.090)	.184* (.089)	.147 (.092)	.143 (.091)
Condition (exp=1; contr=0)			.347* (.163)		.370* (.156)		.321# (.160)		.319# (.160)		.328* (.161)
<b>Eibo</b>											
Eibo in Group 7 & 8 (10-12 year olds)				-.017 (.232)	-.070 (.228)						
Eibo in Group 5 to 8 (8-12 year olds)				-.305 (.260)	-.309 (.254)						
Eibo in Group 1 to 8 (4-12 year olds)				-1.444** (.486)	-	1.555*** (.477)					
None				-.215 (.652)	-.233 (.638)						
Time watching TV						.167** (.071)	.157* (.071)				
Time playing games								.146* (.065)	.135* (.064)		
Time reading English										.117# (.065)	.107 (.064)
<b>Random part</b>											
Student variance	.837 (.107)	.832 (.109)	.807 (.103)	.763 (.097)	.730 (.093)	.801 (.102)	.776 (.099)	.803 (.102)	.778 (.099)	.815 (.104)	.789 (.101)
Class variance		.005 (.026)									
Deviance	327.122	327.092	322.650	315.864	310.411	321.772	317.797	322.122	318.217	323.960	319.896
Ref. model		1	1	1	4	1	6	1	8	1	10
% expl. var. student level			3.584%	8.841%	4.325%	4.301%	3.121%	4.062%	3.113%	2.628%	3.190%
Fit improve- ment		$\chi^2 =$ .030 df=1 p=n.s.	$\chi^2 =$ 4.472 df=1 p<.05	$\chi^2 =$ 11.258 df=4 p<.05	$\chi^2 =$ 5.453 df=1; p<.05	$\chi^2 =$ 5.350 df=1 p<.025	$\chi^2 =$ 3.975; df=1 p<.05	$\chi^2 =$ 5.000; df=1 p<.05	$\chi^2 =$ 3.905; df=1 p<.05	$\chi^2 =$ 3.162; df=1 p<.10	$\chi^2 =$ 4.064; df=1 p<.05

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

The variable 'Eibo' has five possible answers (no Eibo, Eibo only in group 8, Eibo only in group 7 and 8, Eibo in groups 5 to 8 and Eibo in groups 1 to 8). For each of these possible answers, separate dummies are entered in the regression equation except for the answer 'Eibo only in group 8'. This means that students that received Eibo only in group 8 are the reference group. Results show that when these dummies are entered in the regression equation (model 4) all regression coefficients of the Eibo-dummies are negative, implying that the students in the reference group had the largest positive change in Affect from pre- to post-test. It may also be concluded that the differences in the change in Affect between pre- and post-test of students who received Eibo from group 1 to 8 is significantly smaller than that of the reference group ( $p<.01$ ). The other groups show less development than the 'English in group 8 only', even those who

received no English in primary school, although these differences are not significant. Interesting is the time spent watching TV (in the English language) and time playing games both show significant positive correlations with more growth of Affect for studying English in school. These variables explain respectively 4.30% and 4.06% of the changes in Affect scores, which is more than is explained by reading English for leisure (2.63%:  $p < .05$  one sided). This effect watching TV and playing games has on the enjoyment (Affect) of learning English is quite interesting. Perhaps the freely chosen out-of-school exposure to English through the TV programmes and games makes the language offered in school more enjoyable. More research is however needed to really clarify the relationship between Affect, watching TV and playing games.

**Table 6.9.2: Continued results of the multi-level analyses attitude scores: Affect, effect of the experiment without and with controlling for background variables (student N=123; class N=6) (pre-test, sum scores for: dictation, time watching English, playing games and reading English grand mean centred) (s.e. between brackets) (continued)**

model	1	12	13	14	15	16	17
<b>Fixed part</b>							
Intercept	3.380 (.083)	3.375 (.080)	3.198 (.115)	3.765 (.204)	3.619 (.211)	3.800 (.204)	3.643 (.211)
Pre-test	.177# (.092)	.122 (.091)	.117 (.089)	.112 (.088)	.110 (.086)	.098 (.086)	.097 (.085)
Condition (exp.=1; contr.=0)			.331* (.157)		.323* (.146)		.339* (.147)
<b>Eibo (ref=</b>							
<b>In group 8 only)</b>							
Eibo in Group 7 & 8 (10-12 year olds)				-.293 (.229)	-.337 (.226)	-.332 (.229)	-.374 (.225)
Eibo in Group 5 to 8 (8-12 year olds)				-.691** (.259)	-.687** (.254)	-.712** (.261)	-.705** (.256)
Eibo in Group 1 to 8 (4-12 year olds)				-1.711*** (.467)	-1.812*** (.460)	-1.793*** (.463)	-1.882*** (.455)
No Eibo				-.368 (.628)	-.386 (.616)	-.425 (.621)	-.444 (.608)
Time watching TV				.123 (.080)	.121 (.078)	.160* (.074)	.151* (.073)
Time Playing games				.084 (.064)	.074 (.063)		
Time reading English				.016 (.065)	.009 (.064)		
Sum dictation		.018** (.006)	.018** (.006)	.017* (.007)	.017** (.006)	.017* (.007)	.017** (.006)
<b>Random part</b>							
Student variance	.837 (.107)	.782 (.100)	.755 (.096)	.656 (.084)	.631 (.080)	.666 (.085)	.638 (.081)
Deviance	327.122	318.763	314.421	297.148	292.360	299.098	293.854
Ref. model	1	12	12	1	14	1	16
% expl. var. student level		6.571%	3.453%	21.625%	3.811%	20.430%	4.204%
Fit improvement		$\chi^2 = 8.359$ df=1 p<.05	$\chi^2 = 4.342$ df=1 p<.05	$\chi^2 = 29.974$ df=8 p<.01	$\chi^2 = 4.788$ df=1 p<.05	$\chi^2 = 28.024$ df=6 p<.01	$\chi^2 = 5.244$ df=1 p<.05

## Effect of the treatment on changes in scores for Subjective Norm

In table 6.10 results of the regression analyses are presented concerning the change in Subjective Norm.

**Table 6.10: Results of the multi-level analyses attitude scores Subjective Norm, effect of the experiment without and with controlling for background variables (student N=123; class N=6) (pre-test grand mean centred) (s.e. between brackets)**

model	1	2	3	4	5	6	7	8
<b>Fixed part</b>								
Intercept	3.675 (.081)	3.676 (.087)	3.525 (.119)	3.778 (.108)	3.770 (.117)	3.776 (.096)	4.006 (.198)	2.729 (.434)
Pre-test	.236* (.105)	.260** (.105)	.246** (.104)	.286** (.102)	.276** (.103)	.271** (.103)	.327** (.101)	.309 (.101)
Condition (exp=1; contr=0)			.279# (.162)					.234 (.153)
<b>Country of birth</b>								
<b>mother</b>								
Turkey				-.019 (.326)				
Morocco				-.489* (.220)				
Surinam				-.468 (.296)				
Aruba different				.557 (.515) .057 (.244)				
<b>Country of birth</b>								
<b>father</b>								
Turkey					.027 (.279)			
Morocco					-.445* (.222)			
Surinam					-.463 (.289)			
Aruba different					.076 (.376) .159 (.248)			
<b>Language with</b>								
<b>father</b>								
Usually other language same							-.243 (.263)	
Dutch/different different							-.498* (.224)	
							.039 (.373)	
<b>Eibo</b>								
Eibo in Group 7 & 8 (10-12 year olds)							-.351 (.223)	-.385 (.222)
Eibo in Group 5 to 8 (8-12 year olds)							-.580* (.249)	-.584* (.247)
Eibo in Group 1 to 8 (4-12 year olds)							1.044* (.479)	.951* (.478)
None							-1.410* (.627)	-1.429* (.621)
<b>Random part</b>								
Student variance	.814 (.104)	.809 (.106)	.795 (.101)	.758 (.097)	.765 (.098)	.780 (.099)	.705 (.090)	.692 (.088)
Class variance		.005 (.026)						
Deviance	323.816	323.769	320.886	315.052	316.140	318.497	306.006	303.700
Ref. model		1	1	1	1	1	1	7
% expl. var. student level			2.334%				13.391%	
Fit improve- ment		$\chi^2=.047$ df=1 p=n.s.	$\chi^2=2.930$ df=1 p<.10	$\chi^2=8.764$ df=5 p=n.s.	$\chi^2=7.676$ df=5 p=n.s.	$\chi^2=5.319$ df=3 p=n.s.	$\chi^2=17.810$ df=4 p<.01	$\chi^2=2.307$ pf=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

As can be seen in model 2 adding a class level to the model does not significantly improve model fit. When not correcting for covariates (model 3), the students in the treatment group show a one-sided significant more favourable change in Subjective Norm scores than the control group. Model 4 shows that adding the dummies indicating the country of birth of the mother (reference group is 'The Netherlands') does not significantly improve model fit, although the regression coefficient for the dummy indicating that the mother is born in Morocco is significant. Students with Moroccan mothers show a less favourable change in Subjective Norm scores than students with a mother born in the Netherlands. Model 5 shows that the same is true for the country of birth of the father. From model 6 can be inferred that the dummies indicating the language the student speaks with his or her father (reference group is 'Dutch') does not significantly predict change in Subjective Norm although students who report speaking both Dutch and another language with their father, do show a significantly less positive change in Subjective Norm than the reference group. In model 7 can be seen that the effect of Eibo, the years of English language teaching received, is significant and explains 13.39% of the changes in scores for Subjective Norm from pre- to post-test. The group that only received English language teaching in the final two years does not significantly differ in its change in Subjective Norm from the reference group (English in the final year only). The students who received English in the final four years and those who did not receive English show a significantly more negative change than the reference group, while those who received English the full eight years of primary education show a significantly more positive change than the reference group. In model 8 can be seen that after controlling for Eibo, the variable 'condition' is no longer significant. It must be concluded that the positive effect found from the treatment when not correcting for covariates can be caused by the treatment, but can also be an effect of the amount of Eibo received in primary education.

### **Effect of the treatment on changes in scores for Perceived Behavioural Control**

In table 6.11 the results for Perceived Behavioural Control (PBC) are presented. The data show that adding a class level (see model 2) significantly improves model fit and that 7.9% (.062/.640) of the variance in change in PBC-scores between pre- and post-test is class level variance. Furthermore in model 3 it can be seen that when not correcting for covariates, the treatment group shows a significantly more positive

development in PBC than the control group. The variable 'condition' explains 79.03% of class level variance in the development of PBC-scores from pre- to post-test, which is 7.66% of total variance.

**Table 6.11: Results of the multi-level analyses attitude scores PBC, effect of the experiment without and with controlling for background variables (student N=123; class N=6) (pre-test, sumscore dictation and age are grand mean centred) (s.e. between brackets)**

model	1	2	3	4	5	6	7	8
<b>Fixed part</b>	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)	Coeff. (s.e.)
Intercept	3.821 (.072)	3.820 (.123)	3.582 (.122)	3.719 (.127)	3.675 (.143)	3.288 (.389)	3.823 (.132)	3.584 (.137)
Pre-test	.221* (.096)	.149 (.095)	.160# (.093)	.096 (.096)	.092 (.094)	.138 (.094)	.187* (.094)	.184* (.092)
Condition (exp=1; contr=0)			.452* (.168)					.450 (.190)
<b>Country of birth mother</b>								
Turkey				.527 (.290)				
Morocco				.044 (.209)				
Surinam				-.168 (.261)				
Aruban other country.				.618 (.458) .482* (.223)				
<b>Country of birth father</b>								
Turkey					.280 (.242)			
Morocco					.196 (.219)			
Surinam					-.113 (.256)			
Aruban other country.					.846* (.337) .459* (.228)			
Sumscore dictation						.012* (.006)		
Age							-.355** (.134)	-.331* (.132)
<b>Random part</b>								
Student variance	.644 (.082)	.578 (.076)	.587 (.076)	.543 (.071)	.533 (.070)	.560 (.073)	.538 (.071)	.539 (.071)
Class variance		.062 (.053)	.013 (.024)	.038 (.038)	.054 (.047)	.068 (.056)	.077 (.060)	.026 (.031)
Total var.		.640	.600	.581	.587	.628	.615	.565
Deviance	294.842	288.485	283.884	279.265	278.319	285.168	278.687	274.957
Ref. model		1	2	2	2	2	2	7
% expl. var. student level			-	6.055%	7.785%	3.114%	6.920%	-
% expl var. Class level			79.032%	38.710%	12.903%	-	-	66.234%
% expl var total			7.656%	9.219%	8.281%	1.875%	3.906%	8.130%
Fit improvement		$\chi^2=6.357$ df=1 p<.005	$\chi^2=4.601$ df=1 p<.05	$\chi^2=9.220$ df=5 p=n.s.	$\chi^2=10.166$ df=5 p<.10	$\chi^2=3.317$ df=1 p<.10	$\chi^2=9.798$ df=1 p<.01	$\chi^2=3.730$ df=1 p<.10

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Adding the covariates one at a time shows that the dummies for 'country of birth mother', for 'country of birth father' and the dictation-test do not significantly improve model fit. The variable age however is significant, explaining 6.92% of student level variance and 3.91% of total variance. Older students show a less positive development in PBC-scores between pre- and post-test than younger students. After controlling for this one significant covariate, the effect of condition is not significant ( $t=2.36$ ;  $df=2$ ;  $p<.10$  (1-sided)) and also the model fit improvement is non-significant ( $\chi^2=3.730$ ;  $df=1$ ;  $p<.10$ ). This means a significant positive effect of the treatment is found on the development in PBC-scores when not correcting

for covariates. When correcting for the significant covariate 'age', the variable condition is no longer significant when the strict criterion of Hox (2010) in interpreting the ratio of regression coefficient and standard error as a t-value ( $t=450/190=2.37$ ;  $df=2$ ;  $p>.10$ , 2-sided) is applied, although model fit improvement after adding the condition to the regression model does approach significance ( $p<.10$ ) and also as a z-score the ratio would be significant.

### **Effect of the treatment on change in scores for Intention**

In table 6.12 the results for the 'intention' scores are presented. Adding a class level (model 2) does not significantly improve model fit, so analyses will be conducted with only a student level. In model three the variable 'condition' does not show a significant Wald test ( $t=2.09$ ,  $df=3$ ,  $p<.10$  one-sided). However, the model fit does significantly improve after adding 'condition' to the regression equation ( $\chi^2=4.321$ ;  $df=1$ ;  $p<.05$ ). Possibly this difference is caused by the somewhat strict criterion for the Wald test when used for variables that only show variance at class level. In these instances, following Hox (2010) the number of classes are taken as the sample size and not the number of students. Would the Wald test be interpreted as a z-score, the variable condition is significant. The only significant covariate is 'time watching television'. After correcting for this covariate, the variable 'condition' is no longer significant when considered as a t-score with 3 degrees of freedom (as a z-score again it would be significant). Also model 5 does not show a significant model fit improvement compared to model 4, although chi-square is almost significant.

**Table 6.12: Results of the multi-level analyses attitude scores Intention, effect of the experiment without and with controlling for background variables (student N=123; class N=6) (pre-test and time watching TV, grand mean centred) (s.e. between brackets)**

model	1	2	3	4	5
<b>Fixed part</b>					
Intercept	3.570 (.075)	3.566 (.089)	3.404 (.108)	3.566 (.073)	3.415 (.106)
Pre-test	.118 (.088)	.103 (.088)	.115 (.086)	.112 (.086)	.110 (.084)
Condition (exp=1; contr=0)			.309* (.148)		.283# (.145)
Time watching TV				.156* (.064)	.146 (.063)
<b>Random part</b>					
Student variance	.689 (.088)	.675 (.088)	.666 (.085)	.657 (.084)	.637 (.081)
Class variance		.014 (.027)			
Deviance	303.293	302.872	298.972	297.438	293.675
Ref. model		1	1	1	4
% expl. var. student level			3.338%	7.547%	3.044%
Fit improvement		$\chi^2=.421$ df=1 p=n.s.	$\chi^2=4.321$ df=1 p<.05	$\chi^2=5.855$ df=1 p<.05	$\chi^2=3.763$ df=1 p<.10

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Effects of the language teaching programme based on differentiated instruction on Vocabulary**

Table 6.14 shows that in this regression model, with the post-test score for vocabulary as dependent variable and the pre-test as predictor, adding a class level does not improve model fit. Adding a school level, however, does significantly improve model fit in this regression model. For testing fit improvement after adding an additional level of variance, test chi-square needs to be tested 1-sided (Hox, 2010), so the critical chi-square here is 2.706 (df=1). It is clear that adding a class level (see model 1) to the model with only a student level (model 0) does not significantly improve model fit. Adding a school level (see model 2) does significantly improve model fit. As the fit of model 3 does not differ from the fit of model 2, model 3 has been used as most parsimonious model. When model 3 is compared with model 0 the difference in deviance is 3.452 (660.172-656.720) which is significant (df=1, p(one sided)<.05). Multi-level analysis is needed, although the intra-school correlation is not enormous, with  $(.713/10.692)=.06668$ , that is 7% school-related variance in the learning gains for vocabulary.

**Table 6.14. Multi-level analyses for ‘vocabulary’ (student N=127; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets).**

	model 0	model 1	model 2	model 3
<b>Fixed part</b>				
Intercept	24.530*** (.289)	24.491*** (.375)	24.442*** (.564)	24.442*** (.564)
Pre-test vocabulary	.637*** (.064)	.605*** (.066)	.575*** (.067)	.575*** (.067)
<b>Random part</b>				
Student variance	10.594 (1.330)	10.270 (1.320)	9.979 (1.267)	9.979 (1.267)
Class variance		.349 (.486)	.000 (.000)	
School variance			.713 (.776)	.713 (.776)
Total variance	10.594	10.619	10.692	10.692
Deviance	660.172	659.449	656.720	656.720
Ref. model		0	1	2
Fit improvement		$\chi^2=.723$ df=1 p=n.s.	$\chi^2=2.729$ df=1 p<.10	$\chi^2=.000$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

The data in table 6.15a allow us to test whether the treatment has a significant effect on changes in vocabulary (model 2). As can be seen, this is not the case, model fit improvement and the Wald test for the regression coefficient are both non-significant. Models 3 to 7 show that gender significantly predicts changes in vocabulary between pre- and post-test ( $p<.05$  2-sided). Girls' vocabulary scores progress slower than that of boys. In tables 6.15b to 6.20 the effects of the other covariates are tested. As can be seen, next to gender, the Cito-test score is significant at  $p<.05$ . Time watching TV and time playing games are only one-sided significant at  $p<.05$ .

**Table 6.15a: Results of multi-level analyses for vocabulary, effects of treatment with and without control for covariates (student N=127; class N=6) (pre-test and time watching TV, playing games and reading English are grand mean centred) (s.e. between brackets).**

model	0	1	2	3	4	5	6	7
<b>Fixed part</b>								
intercept	24.442*** (.564)	24.444*** (.642)	25.021*** (.910)	24.428*** (.558)	24.403*** (.587)	24.449*** (.556)	25.733*** (1.303)	24.987*** (.665)
Pre-test Vocabulary	.575*** (.067)	.575*** (.067)	.580*** (.068)	.553*** (.067)	.549*** (.067)	.579*** (.066)	.562*** (.067)	.537*** (.068)
Condition (exp=1; contr=0)		-.004 (.566)						
<b>Eibo</b>								
Eibo in Group 7 & 8 (10-12 year olds)			-.417 (.837)					
Eibo in Group 5 to 8 (8-12 year olds)			-1.009 (.935)					
Eibo in Group 1 to 8 (4-12 year olds)			-1.853 (1.729)					
None			-2.343 (2.379)					
Time watching TV				.483# (.255)				
Time playing games					.409# (.226)			
Time reading English						.278 (.226)		
Country of birth student (1=Dutch; 0=not Dutch)							-1.374 (1.235)	
Gender (0=boy; 1=girl)								-1.186* (.571)
<b>Random part</b>								
Student variance	9.979 (1.267)	9.979 (1.267)	9.743 (1.237)	9.702 (1.232)	9.706 (1.233)	9.866 (1.253)	9.863 (1.253)	9.611 (1.221)
School variance	.713 (.776)	.713 (.779)	.829 (.869)	.699 (.761)	.797 (.840)	.689 (.755)	.790 (.836)	.880 (.903)
Total variance	10.692	10.692	10.572	10.401	10.503	10.555	10.653	10.491
Deviance	656.720	656.720	654.087	653.168	653.515	655.218	655.499	652.524
Ref. model		0	0	0	0	0	0	0
% of expl. Var. school level								
% of expl. Var. student level								3.688%
% of expl. Var. total								1.880%
Fit improvement		$\chi^2 = .000$ df=1 p=n.s.	$\chi^2 = 2.633$ df=4 p=n.s.	$\chi^2 = 3.552$ df=1 p<.10	$\chi^2 = 3.205$ df=1 p<.10	$\chi^2 = 1.502$ df=1 p=n.s.	$\chi^2 = 1.221$ df=1 p=n.s.	$\chi^2 = 4.196$ df=1 p<.05

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.15b (continuation): Results of multi-level analyses for vocabulary, effects of covariates (student N=127; class N=6; school N=3) (pre-test is grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4
<b>Fixed part</b>					
Intercept	24.442*** (.564)	24.477*** (.674)	24.171*** (.692)	24.504*** (.591)	24.637*** (.580)
Pre-test Vocabulary	.575*** (.067)	.538*** (.070)	.511*** (.070)	.574*** (.067)	.580*** (.067)
<b>Country of birth mother (ref=The Netherlands)</b>					
Turkey		-1.281 (1.178)			
Morocco		-.730 (.887)			
Surinam		.215 (.982)			
Aruban		3.157 (1.902)			
other country.		.443 (.950)			
<b>Country of birth father (ref=The Netherlands)</b>					
Turkey			-1.037 (1.011)		
Morocco			-.008 (.930)		
Surinam			1.738 (1.033)		
Aruban			-.019 (1.384)		
other country.			1.479 (.956)		
Native tongue mother (0=Dutch; 1=not Dutch)				-.218 (.649)	
Native tongue father (0=Dutch; 1=not Dutch)					-.659 (.663)
<b>Random part</b>					
Student variance	9.979 (1.267)	9.504 (1.207)	9.416 (1.196)	9.973 (1.267)	9.915 (1.259)
School variance	.713 (.776)	.860 (.885)	.843 (.874)	.701 (.765)	.656 (.730)
Total variance	10.692	10.364	10.259	10.674	10.571
Deviance	656.720	651.078	649.860	656.607	655.739
Ref. model		0	0	0	0
Fit improvement		$\chi^2 = 5.642$ df=5 p=n.s.	$\chi^2 = 6.860$ df=5 p=n.s.	$\chi^2 = .113$ df=1 p=n.s.	$\chi^2 = .981$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.16: Results of multi-level analyses for vocabulary, effects of covariate 'age' (student N=127; class N=6; school N=3) (pre-test and age grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	24.424*** (.561)	24.429*** (.586)
Pre-test Vocabulary	.575*** (.067)	.574*** (.067)
Age in days		.251 (.564)
<b>Random part</b>		
Student variance	10.059 (1.283)	10.024 (1.278)
School variance	.700 (.768)	.782 (.837)
Total variance	10.759	10.806
Deviance	652.516	652.329
Ref. model		0
Fit improvement		$\chi^2 = .187$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.17: Results of multi-level analyses for vocabulary, effects of covariate ‘Cito-test’ (student N=96; class N=6; school N=3) (pre-test and Cito-test are grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	25.235*** (.471)	24.796*** (.296)
Pre-test Vocabulary	.623*** (.071)	.612*** (.071)
Cito-test		.081* (.038)
<b>Random part</b>		
Student variance	8.232 (1.204)	8.168 (1.179)
School variance	.318 (.493)	.000 (.000)
Total variance	8.550	8.168
Deviance	477.011	474.055
Ref. model		0
% expl. student var.		.777
% expl. school var.		100.00
% expl. total var.		4.468
Fit improvement		$\chi^2 = 2.956$ df=1 p<.10

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.18: Results of multi-level analyses for vocabulary, effects of covariate ‘Language spoken with mother’(student N=126; class N=6; school N=3) (pre-test is grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	24.463*** (.585)	24.492*** (.614)
Pre-test Vocabulary	.570*** (.067)	.571*** (.068)
<b>Language spoken with mother</b> (Ref. cat. = Dutch)		
Half Dutch, half other language		-.015 (.790)
Other language		-.323 (.947)
<b>Random part</b>		
Student variance	9.992 (1.274)	9.982 (1.273)
School variance	.783 (.835)	.785 (.839)
Total variance	10.775	10.767
Deviance	651.937	651.819
Ref. model		0
Fit improvement		$\chi^2 = .118$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.19: Results of multi-level analyses for vocabulary, effects of covariate ‘Language spoken with father’ (student N=126; class N=6; school N=3) (pre-test is grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	24.441*** (.545)	24.602*** (.561)
Pre-test Vocabulary	.586*** (.068)	.586*** (.068)
<b>Language spoken with father (Ref. cat. = Dutch)</b>		
Half Dutch, half other language		-.237 (.810)
Other language		-.162 (.941)
<b>Random part</b>		
Student variance	8.962 (1.172)	8.965 (1.172)
School variance	.661 (.726)	.623 (.692)
Total variance	9.623	9.588
Deviance	607.807	607.713
Ref. model		0
Fit improvement		$\chi^2 = .094$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Now that it is clear which covariates had a significant effect on the model fit, the following and final table

(Table 6.20) presents the model fit with all significant covariates..

**Table 6.20 (final table): Results of multi-level analyses for vocabulary, effects of treatment after control for covariates (student N=127/96; class N=6; school N=3) (pre-test, time watching TV, time playing games and Cito-test are grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4	5
<b>Sample size</b>	127	127	127	127	96	96
<b>Fixed part</b>						
intercept	24.858*** (.660)	24.967*** (.724)	24.987** * (.665)	25.025** * (.735)	25.490** * (.412)	25.576** * (.517)
Pre-test Vocabulary	.514*** (.068)	.513*** (.068)	.537*** (.068)	.536*** (.068)	.576*** (.071)	.576*** (.071)
Gender (girl=1; boy=0)	-.970# (.584)	-.978# (.584)	-1.186* (.571)	-1.190* (.572)	-1.368* (.580)	-1.374* (.581)
Watching TV (1 = never; 5= several hours per day)	.378 (.261)	.387 (.262)				
Playing games (1 = never; 5= several hours per day)	.210 (.240)	.213 (.240)				
Cito					.086* (.037)	.086* (.037)
Condition (experimental=1; control=0)		-.193 (.552)		-.067 (.556)		-.156 (.570)
<b>Random part</b>						
Student variance	9.329 (1.185)	9.325 (1.184)	9.611 (1.221)	9.612 (1.221)	7.721 (1.114)	7.715 (1.114)
School variance	.849 (.875)	.827 (.857)	.880 (.903)	.872 (.896)	.000 (.000)	.000 (.000)
Total variance	10.178	10.152	10.491	10.484	7.721	7.715
Deviance	648.725	648.603	652.524	652.510	468.655	468.580
Ref. model		0		2		4
Fit improvement		$\chi^2 = .122$ df=1 p=n.s.		$\chi^2 = .014$ df=1 p=n.s.		$\chi^2 = .075$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

In table 6.20 (final table) in model 0 next to the pre-test score the covariates gender, time watching TV and time playing games are included in the model. The Cito-test is not included because this variable has many missing values. In model 1 the variable 'condition' is added to model 0 and shows to be non-significant. The covariates time watching TV and time playing games are also non-significant. In model 2 these non-significant covariates are removed and model 3 shows that adding the variable condition to model 2 does not significantly improve model fit. In model 4 a new model of reference is created with gender and Cito-test as covariates. Both covariates are significant. In model 5 the variable condition is added to model 4, which does not significantly improves model fit. The conclusion is therefore that the treatment both with and without covariates, does not show a significant effect on changes in vocabulary from pre- to post-test.

#### **Effects of the language teaching programme based on differentiated instruction on Listening/Spelling skills (dictation).**

Below the results relating to the changes in the scores of the listening/spelling (dictation) test are presented. Model 2 shows that adding a class level to the model with only a pre-test as predictor yields a significantly improved model fit ( $\chi^2=5.016$ ,  $df=1$ ,  $p<.05$ ). The subsequent addition of a school level does not give a significant improved fit ( $\chi^2=.000$ ,  $df=1$ , n.s.), so the regression analyses for 'dictation' will be performed with a student and class level.

**Table 6.21. Multi-level analyses for 'dictation' post-test with pre-test (standard errors between brackets. (student N=127; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

	model 0	model 1	model 2
<b>Fixed part</b>			
Intercept	55.622*** (.423)	55.395*** (.857)	55.395*** (.857)
Pre-test dictation	.619*** (.035)	.578*** (.042)	.578*** (.042)
<b>Random part</b>			
Student variance	22.630 (2.840)	20.267 (2.606)	20.267 (2.606)
Class variance		3.399 (2.523)	3.399 (2.523)
School variance			.000 (.000)
Total variance		23.666	23.666
Deviance	756.560	751.544	751.544
Ref. model		0	1
Fit improvement		$\chi^2 = 5.016$ $df=1$ $p<.05$	$\chi^2 = .000$ $df=1$ $p=n.s.$

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

The table 6.22a shows that the variable indicating whether students are part of the experimental group correlates significantly with the development of the scores between the pre- and post-test for the dictation ( $\chi^2=4.583$ ,  $df=1$ ,  $p<.05$ ). After controlling for the pre-test score, the experimental group scores 2.7 points higher in the post-test than the control group. The regression coefficient of the 'condition' variable is also one-sided significant ( $t=2.74$ ,  $df=3$ , one-sided  $p<.01$ ). Based on these results it can be concluded that the treatment shows a positive effect on the listening/spelling skills without correcting for covariates.

When subsequently the background variables are incorporated in the regression comparison one by one, 'country of birth' ( $\chi^2=7.473$ ,  $df=1$ ,  $p<.01$ ), 'native tongue mother' ( $\chi^2=4.275$ ,  $df=1$ ,  $p<.05$ ) and 'language spoken with mother' ( $\chi^2=6.437$ ,  $df=2$ ,  $p<.05$ ) turn out to be significant. The 'country of birth father' variable is significant if the Wald test is taken as criterion ( $z=1.975$ ,  $p<.05$ ), but not when the difference in deviance between both models is used ( $\chi^2=3.831$ ,  $df=1$ ,  $p<.10$ ).

Table 6.22a: Results of multi-level analyses for dictation, effects of covariates (student N=127; class N=6; school N=3) (pre-test, time watching TV, time playing games and time reading English grand mean centred) (s.e. between brackets).

model	0	1	2	3	4	5
<b>Fixed part</b>						
intercept	55.395*** (.857)	54.139*** (.714)	56.333*** (1.348)	55.387*** (.860)	55.377*** (.867)	55.393*** (.85)
Pre-test Dictation condition (experimental=1; control=0)	.578*** (.042)	.602*** (.036) 2.697** (.986)	.591*** (.044)	.568*** (.045)	.572*** (.042)	.580*** (.043)
<b>Eibo (reference group = only group 8)</b>						
No eibo			-.154 (3.424)			
Group 7&8 (10-12 year olds)			-.950 (1.263)			
Group 5 to 8 (8-12 year olds)			-1.410 (1.396)			
Group 1 to 8 (4-12 year olds)			-1.543 (2.528)			
Time watching TV (1 = never; 5= several hours per day)				.210 (.395)		
Time playing games (1 = never; 5= several hours per day)					.209 (.328)	
Time reading English (1 = never; 5= several hours per day)						-.097 (.345)
Country of birth student (Dutch=1; not Dutch=0)						
Gender (girl=1; boy=0)						
<b>Random part</b>						
Student variance	20.267 (2.606)	20.610 (2.648)	20.083 (2.583)	20.213 (2.599)	20.177 (2.595)	20.260 (2.605)
Class variance	3.399 (2.523)	.453 (.829)	3.368 (2.501)	3.431 (2.550)	3.506 (2.580)	3.372 (2.506)
Total variance	23.666	21.063	23.451	23.644	23.683	23.632
% of expl. student var.		-				
% of expl. class var.		86.673%				
% of expl. total var.		10.999%				
deviance	751.544	746.961	750.386	751.262	751.141	751.466
Ref. model		0	0	0	0	0
Fit improvement		$\chi^2 = 4.583$ df=1 p<.05	$\chi^2 = 1.158$ df=4 p=n.s.	$\chi^2 = .282$ df=1 p=n.s.	$\chi^2 = .403$ df=1 p=n.s.	$\chi^2 = .078$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.22b (continuation): Results of multi-level analyses for dictation, effects of covariates (student N=127; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4
<b>Fixed part</b>					
intercept	55.395*** (.857)	54.656*** (.765)	54.517*** (.915)	56.852*** (.977)	55.575*** (1.088)
Pre-test dictation	.578*** (.042)	.601*** (.039)	.580*** (.041)	.598*** (.040)	.579*** (.042)
<b>Country of birth mother (ref=The Netherlands)</b>					
Turkey		3.865* (1.698)			
Morocco		.954 (1.241)			
Surinam		.186 (1.397)			
Aruban		2.736 (2.759)			
other country.		2.698* (1.286)			
<b>Country of birth father (ref=The Netherlands)</b>					
Turkey			3.097* (1.441)		
Morocco			1.400 (1.341)		
Surinam			1.469 (1.526)		
Aruban			2.497 (2.073)		
other country.			1.143 (1.350)		
Native tongue mother (Dutch=1; not Dutch=0)				-1.985* (.928)	
Native tongue father (Dutch=1; not Dutch=0)					-.252 (.967)
<b>Random part</b>					
Student variance	20.267 (2.606)	19.510 (2.508)	19.605 (2.520)	19.904 (2.559)	20.275 (2.607)
Class variance	3.399 (2.523)	1.495 (1.405)	2.507 (1.998)	2.125 (1.783)	3.317 (2.474)
Total variance	23.666	21.005	22.112	22.029	23.592
% of expl. student var.				1.791%	
% of expl. class var.				37.482%	
% of expl. total var.				6.917%	
Deviance	751.544	743.442	746.108	747.269	751.477
Ref. model		0	0	0	0
Fit improvement		$\chi^2 = 8.102$ df=5 p=n.s.	$\chi^2 = 5.436$ df=1 p=n.s.	$\chi^2 = 4.275$ df=1 p<.05	$\chi^2 = .067$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.23: Results of multi-level analyses for dictation, effects of covariate ‘age’ (student N=126; class N=6; school N=3) (pre-test and age grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	55.316*** (.875)	55.376*** (.847)
Pre-test dictation	.575*** (.042)	.572*** (.042)
Age in days		-.530 (.818)
<b>Random part</b>		
Class variance	3.587 (2.637)	3.298 (2.470)
Student variance	20.213 (2.610)	20.209 (2.609)
Total variance	23.800	23.507
Deviance	745.589	745.178
Ref model		0
Fit improvement		$\chi^2 = .411$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.24: Results of multi-level analyses for dictation, effects of covariate ‘Cito-test’ (student N=96; class N=6; school N=3) (pre-test and Cito-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	57.282*** (1.414)	55.418*** (1.297)
Pre-test dictation	.537*** (.051)	.535*** (.052)
Cito-test		.069 (.090)
<b>Random part</b>		
Class variance	10.335 (6.754)	7.843 (5.282)
Student variance	10.844 (1.616)	10.953 (1.632)
Total variance	21.179	18.796
deviance	516.152	515.648
Ref. model		0
Fit improvement		$\chi^2 = .504$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.25: Results of multi-level analyses for dictation, effects of covariate ‘language spoken with mother’ (student N=126; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	55.358*** (.860)	54.880*** (.707)
Pre-test dictation	.578*** (.042)	.608*** (.039)
<b>Language spoken with mother (ref. Dutch)</b>		
Half Dutch, half other language		1.190 (1.097)
Other language		3.595** (1.373)
<b>Random part</b>		
Student variance	20.414 (2.636)	19.879 (2.566)
Class variance	3.424 (2.546)	1.595 (1.480)
Total variance	23.838	21.474
% of expl. student var.		2.621%
% of expl. class var.		53.417%
% of expl. total var.		9.917%
deviance	746.573	740.136
Ref. model		0
Fit improvement		$\chi^2 = 6.437$ df=2 p<.05

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.26: Results of multi-level analyses for dictation, effects of covariate ‘language spoken with father’(student N=126; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	55.363*** (.809)	55.354*** (.833)
Pre-test dictation (gm)	.590*** (.043)	.598*** (.043)
<b>Language spoken with father (ref. Dutch)</b>		
Half Dutch, half other language		1.109 (1.252)
Other language		-.618 (1.447)
<b>Random part</b>		
Student variance	21.118 (2.797)	20.953 (2.776)
Class variance	2.822 (2.253)	2.626 (2.133)
Total variance	23.940	23.579
Deviance	714.278	713.063
Fit improvement compared to model ...		0
Fit improvement		$\chi^2 = 1.215$ df=2 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.27 (final table): Results of multi-level analyses for dictation, effects of treatment after control for covariates (student N=126; class N=6; school N=3) (all continuous predictors grand mean centred) (s.e. between brackets)**

model	model 0	model 1
<b>Fixed part</b>		
intercept	29.584*** (2.837)	27.633*** (2.556)
Pre-test dictation	.597*** (.040)	.608*** (.034)
Land of birth student (Netherlands)	-3.632* (1.801)	-3.186 (1.795)
Land of birth father (Netherlands)	-.714 (1.123)	-1.035 (1.019)
<b>Language spoken with mother (ref. Dutch)</b>		
Half Dutch , half other language	.556 (1.205)	.598 (1.187)
Mostly other language	2.331 (1.510)	2.346 (1.485)
Native tongue mother (Dutch)	.000 (.000)	.000 (.000)
condition (experimental)		2.385** (.819)
<b>Random part</b>		
Student variance	18.939 (2.445)	19.122 (2.409)
Class variance	2.040 (1.708)	.000 (.000)
Total variance	20.979	19.122
% of explained student variance		-
% of explained class variance		1.00
% of explained total variance		.089
Deviance	735.185	729.377
Ref. model		0
Fit improvement		$\chi^2 = 5.808$ df=1 p<.025

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Again a significant effect is found of the 'condition' variable on the trend of dictation scores between pre- and post-test after controlling for significant covariates; it even explains all class related variance. So, the treatment shows a significant effect on the trend in dictation scores both with and without correcting for the significant covariates.

Quite noticeable are the covariates referring to having Dutch speaking parents vs. having parents that speak other languages. Students with a Dutch background (country of birth pupil, native tongue mother and language spoken with mother) consistently show a less positive trend on their scores for dictation. Perhaps the children with other language backgrounds than Dutch have developed more metalinguistic awareness, are more eager to add an international language to their language skills repertoire, or the selection based on language background coincides with other traits not researched in this study. An explanation for these differences in trends cannot be based on the data gathered in this study and further research is needed to find out what might cause this effect.

Top

## Effects of the language teaching programme based on differentiated instruction on Reading skills, A

### (Editing test-scores for correctly underlined words 'Edecor')

The effects of the language teaching programme based on differentiated instruction on Reading skills will be described below in two parts. The first part, A, deals with the editing test-scores for correctly underlined words ('Edecor'), the second part, B, deals with the editing test-scores for incorrectly underlined words ('EdIncor'). This sub-division is based on the correlations calculated between the sum scores of the different tests as has been described in the analyses chapter above. The correlations, which can be found in the appendix (Table AP10C1), clearly shows two different things have been measured and both 'Edecor' and 'EdIncor' are scores for different traits of the reading skill.

**Table 6.28. Multi-level analyses for correctly underlined words in the editing test for reading 'EdCor' post-test with pre-test (student N=127; class N=6; school N=3) (all continuous predictors grand mean centred) (s.e. between brackets)**

model	0	1	2	3
<b>Fixed part</b>				
Intercept	-.052 (.056)	-.112 (.178)	-.112 (.226)	-.094 (.201)
Pre-test editing correct	.789*** (.057)	.590*** (.061)	.582*** (.062)	.606*** (.064)
<b>Random part</b>				
Student variance	.394 (.049)	.298 (.038)	.298 (.038)	.329 (.042)
Class variance		.175 (.109)	.064 (.064)	
School variance			.113 (.128)	.113 (.098)
Total variance	.394	.473	.475	.442
Deviance	242.235	222.300	221.028	227.235
Ref. model		0	1	2
Fit improvement		$\chi^2 = 19.935$ df=1 p<.001	$\chi^2 = 1.272$ df=1 p=n.s.	$\chi^2 = 6.207$ df=1 p<.02

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Table 6.28 shows that adding a class level to the model with only the pre-test as predictor yields a significantly improved model fit ( $\chi^2=19.935$ , df=1, p<.001). The subsequent addition of the school level does not give a significant improved fit ( $\chi^2=1.272$ , df=1, n.s.), although school related variance is greater than the class related variance. The deviance of a model with a student- and a school-level is, on the other hand, much larger than a model with student and class-level. For this reason the addition of a class-level and not a school-level to the model has been chosen.

**Table 6.29a: Results of multi-level analyses for correctly underlined words in the editing test for Reading 'Edcor', effects of treatment and of covariates (student N=127; class N=6; school N=3) (pre-test, time watching TV, time playing games and time reading English grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4	5	6	7
<b>Fixed part</b>								
intercept	-.112 (.178)	-.293 (.229)	-.127 (.215)	-.113 (.175)	-.115 (.181)	-.108 (.177)	.180 (.268)	-.098 (.184)
Pre-test editing correct	.590*** (.061)	.595*** (.061)	.590*** (.061)	.568*** (.063)	.575*** (.063)	.580*** (.061)	.598*** (.061)	.590*** (.061)
group (experimental)		.362 (.322)						
<b>Eibo (reference group = only Group 8)</b>								
No Eibo			.293 (.414)					
Eibo in Group 7 & 8 (10-12 year olds)			-.031 (.148)					
Eibo in Group 5 to 8 (8-12 year olds)			.075 (.163)					
Eibo in Group 1 to 8 (4-12 year olds)			.219 (.302)					
Time watching TV (1 = never; 5= several hours per day)				.072 (.046)				
Time playing games (1 = never; 5= several hours per day)					.041 (.040)			
Time reading English (1 = never; 5= several hours per day)						.078# (.041)		
Country of birth (Netherlands)							-.309 (.213)	
gender (girl)								-.030 (.098)
<b>Random part</b>								
Student variance	.298 (.038)	.299 (.038)	.294 (.038)	.293 (.038)	.295 (.038)	.290 (.037)	.294 (.038)	.298 (.038)
class Variance	.175 (.109)	.141 (.090)	.177 (.110)	.169 (.105)	.181 (.112)	.173 (.108)	.172 (.107)	.176 (.110)
Total variance	.473	.440	.471	.462	.476	.463	.466	.474
% of expl. student var.							2.759%	
% of expl. class var.							1.156%	
% of expl. total var.							2.114%	
Deviance	222.300	221.148	220.327	219.859	221.235	218.702	220.216	222.204
Ref. model		0	0	0	0	0	0	0
		$\chi^2 =$ 1.152 df=1 p=n.s.	$\chi^2 =$ 1.973 df=4 p=n.s.	$\chi^2 =$ 2.441 df=1 p=n.s.	$\chi^2 =$ 1.065 df=1 p=n.s.	$\chi^2 =$ 3.598 df=1 p<.10	$\chi^2 =$ 2.084 df=1 p=n.s.	$\chi^2 =$ .096 df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.29b (continuation): Results of multi-level analyses for correctly underlined words in the editing test for Reading ‘Edcor’, effects of covariates (student N=127; class N=6; school N=3) (all continuous predictors grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4
<b>Fixed part</b>					
intercept	-.112 (.178)	-.139 (.182)	-.125 (.185)	-.030 (.192)	-.196 (.200)
Pre-test editing correct	.590*** (.061)	.588*** (.066)	.604*** (.067)	.595*** (.061)	.586*** (.061)
<b>Country of birth mother (ref=The Netherlands)</b>					
Turkey		-.015 (.211)			
Morocco		.068 (.164)			
Surinam		-.070 (.177)			
Aruban		-.066 (.356)			
other country.		.209 (.167)			
<b>Country of birth father (ref=The Netherlands)</b>					
Turkey			.053 (.182)		
Morocco			.044 (.178)		
Surinam			-.034 (.194)		
Aruban			-.073 (.268)		
other country.			.062 (.174)		
Native tongue mother (Dutch)				-.113 (.115)	
Native tongue father (Dutch)					.118 (.120)
<b>Random part</b>					
Student variance	.298 (.038)	.293 (.038)	.298 (.038)	.297 (.038)	.296 (.038)
Class variance	.175 (.109)	.167 (.104)	.164 (.103)	.166 (.104)	.183 (.114)
Total variance	.473	.460	.462	.463	.479
Deviance	222.300	219.859	221.867	221.340	221.340
Ref. model		0	0	0	0
		$\chi^2 = 2.441$ df=5 p=n.s.	$\chi^2 = .433$ df=5 p=n.s.	$\chi^2 = .960$ df=1 p=n.s.	$\chi^2 = .960$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.30: Results of multi-level analyses for correctly underlined words in the editing test for Reading ‘Edcor’, effects of covariate ‘age’ (student N=126; class N=6; school N=3) (pre-test and age grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	-.115 (.178)	-.110 (.176)
Pre-test editing correct	.592*** (.062)	.593*** (.062)
Age in days		-.017 (.100)
<b>Random part</b>		
Student variance	.300 (.039)	.301 (.039)
Class variance	.175 (.109)	.172 (.107)
Total variance	.475	.473
Deviance	221.416	221.389
Ref. model		0
Fit improvement		$\chi^2 = .027$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.31: Results of multi-level analyses for correctly underlined words in the editing test for Reading ‘Edcor’, effects of covariate ‘Cito-test’ (student N=96; class N=6; school N=3) (all continuous predictors grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	.111 (.175)	.126 (.081)
Pre-test editing correct	.546*** (.061)	.521*** (.062)
Cito-test		.030*** (.009)
<b>Random part</b>		
Student variance	.255 (.038)	.256 (.038)
Class variance	.150 (.104)	.017 (.020)
Total variance	.405	.273
% of expl. student var.		-
% of expl. class var.		88.667%
% of expl. total var.		32.593%
Deviance	153.830	145.559
Ref. model		0
Fit improvement		$\chi^2 = 8.271$ df=1 p<.005

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.32: Results of multi-level analyses for correctly underlined words in the editing test for Reading ‘Edcor’, effects of covariate ‘language spoken with mother’ (student N=126; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	-.117 (.178)	-.142 (.175)
Pre-test editing correct	.590*** (.062)	.595*** (.061)
<b>Language spoken with mother (ref. Dutch)</b>		
Half Dutch, half other language		.064 (.139)
Mostly other language		.165 (.169)
<b>Random part</b>		
Student variance	.300 (.039)	.298 (.039)
Class variance	.175 (.109)	.163 (.103)
Total variance	.475	.461
Deviance	221.161	220.154
Ref. model		0
Fit improvement		$\chi^2 = 1.007$ df=2 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.33: Results of multi-level analyses for correctly underlined words in the editing test for Reading ‘Edecor’, effects of covariate ‘language spoken with father’ (student N=120; class N=6; school N=3) (all continuous predictors grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	-.122 (.182)	-.076 (.194)
Pre-test editing correct	.569*** (.063)	.567*** (.062)
<b>Language spoken with father (ref. Dutch)</b>		
Half Dutch, half other language		-.016 (.149)
Mostly other language		-.399* (.172)
<b>Random part</b>		
Student variance	.297 (.039)	.281 (.037)
Class variance	.183 (.114)	.204 (.126)
Total variance	.480	.485
% of expl. student var.		5.387%
% of expl. class var.		-
% of expl. total var.		-
Deviance	210.220	204.666
Fit improvement compared to model ...		0
		$\chi^2 = 5.554$
		df=2
		p<.10

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

The Cito-test scores have a significant effect. Time reading English books for leisure and the language spoken with father are also significant, albeit only at p<.10 (or 5% one-sided).

**Table 6.34 (final table): Results of multi-level analyses for correctly underlined words in the editing test for Reading ‘Edecor’, effects of treatment after control for covariates (student N=90; class N=6; school N=3) (pre-test, Cito-test and time reading English grand mean centred) (s.e. between brackets)**

model	0	1	3	4
<b>Fixed part</b>				
intercept	.248* (.107)	.070 (.084)	.126 (.081)	-.025 (.075)
Pre-test editing correct	.490*** (.060)	.496*** (.060)	.521*** (.062)	.516*** (.061)
Cito-test	.037*** (.010)	.030*** (.007)	.030*** (.009)	.027*** (.007)
Time reading English (1 = never; 5= several hours per day)	.049 (.040)	.030 (.039)		
<b>Language spoken with father (ref. Dutch)</b>				
Half Dutch, half other language	.089 (.139)	.166 (.133)		
Mostly other language	-.375* (.149)	-.296* (.144)		
condition (experimental)		.320** (.100)		.305** (.102)
<b>Random part</b>				
Student variance	.213 (.033)	.215 (.032)	.256 (.038)	.247 (.036)
Class variance	.037 (.032)	.000 (.000)	.017 (.020)	.000 (.000)
Total variance	.250	.215	.273	.247
% of expl. student var.		-		3.5
% of expl. class var.		100%		100%
% of expl. total var.		14.000%		9.524%
Deviance	123.140	117.031	145.559	138.260
Ref. model		0		3
		$\chi^2 = 6.109$		$\chi^2 = 7.299$
		df=1		df=1
		p<.02		p<.01

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Without controlling for covariates no significant difference in growth is found for the scores for correctly underlined words in the reading test between the experimental and control group. After correcting for significant covariates a difference is found for the scores for correctly underlined words in the reading test, when only correcting for Cito-scores as well as when correcting for the Cito-test and the covariates that were significant at 10% (=5% one-sided).

### Effects of the language teaching programme based on differentiated instruction on Reading skills, B

#### (Editing test-scores for incorrectly underlined words, 'EdIncor')

**Table 6.35. Multi-level analyses for incorrectly underlined words in the editing test for Reading 'EdIncor' post-test with pre-test (student N=127; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1	2	3
<b>Fixed part</b>				
Intercept	.027 (.076)	.094 (.210)	.101 (.302)	.101 (.302)
Pre-test editing incorrect	.556*** (.085)	.333*** (.089)	.316*** (.084)	.316*** (.084)
<b>Random part</b>				
Student variance	.737 (.092)	.572 (.074)	.561 (.071)	.561 (.071)
Class variance		.237 (.153)	.000 (.000)	
School variance			.259 (.222)	.259 (.222)
Total variance	.737	.809	.820	.820
Deviance	321.633	302.951	296.092	296.092
Ref. model		0	1	2
		$\chi^2 = 18.682$	$\chi^2 = 6.859$	$\chi^2 = .000$
		df=1	df=1	df=1
		p<.001	p<.01	p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Table 6.35 above shows that adding class-level to the model yields a significant improved model fit, the subsequent addition of school level again significantly improves the fit. The class related variance then becomes nil. For this reason a model with only a student- and a school-level is used. This model does not fit significantly worse than the model with three levels. That is why the most parsimonious model has been chosen with a student- and school-level. In table 6.36a model 1 shows that the variable indicating whether students were placed in the experimental- or control group does not significantly predict growth in scores of incorrect underlined words in the editing test. Furthermore the effects of some covariates are presented in the table, as well as in the following tables. None of the covariates is a significant predictor of the trend in incorrectly underlined words in the editing test for Reading (EdInc) , except the Cito-test. Gender appears

to be significant only at 10% (5% one sided). In table 6.41 (final table) the effect of the condition after correcting for the Cito-test and gender is presented.

**Table 6.36a: Results of multi-level analyses for incorrectly underlined words in the editing test for Reading ‘EdIncor’, effects of treatment with and without control for (student N=127; class N=6; school N=3) (pre-test, time watching TV, time playing games and time reading English grand mean centred) (s.e. between brackets)**

	model 0	model 1	model 2	model 3	model 4	model 5	model 6	model 7
<b>Fixed part</b>								
intercept	.101 (.302)	.172 (.321)	.175 (.346)	.101 (.295)	.106 (.306)	.100 (.300)	-.208 (.412)	-.015 (.318)
Pre-test editing incorrect (gm)	.316*** (.084)	.295*** (.086)	.308*** (.084)	.297*** (.086)	.287*** (.086)	.314*** (.084)	.305*** (.084)	.302*** (.083)
condition (experimental)		-.130 (.138)						
<b>Eibo (reference group = only in Group 8)</b>								
No Eibo			-.037 (.568)					
Eibo in Group 7 & 8 (10-12 year olds)			-.097 (.201)					
Eibo in Group 5 to 8 (8-12 year olds)			-.032 (.224)					
Eibo in Group 1 to 8 (4-12 year olds)			-.355 (.417)					
Time watching TV (1 = never; 5= several hours per day)				-.063 (.063)				
Time playing games (1 = never; 5= several hours per day)					-.069 (.055)			
Time reading English (1 = never; 5= several hours per day)						-.018 (.054)		
Country of birth (Dutch)							.328 (.292)	
Gender (girl)								.247# (.133)
<b>Random part</b>								
Student variance	.561 (.071)	.556 (.071)	.557 (.071)	.557 (.071)	.554 (.070)	.561 (.071)	.555 (.071)	.545 (.069)
School variance	.259 (.222)	.278 (.237)	.265 (.226)	.248 (.213)	.266 (.228)	.256 (.220)	.267 (.228)	.278 (.237)
Total variance	.820	.834	.822	.805	.820	.817	.822	.823
% of expl. student var.								2.852%
% of expl. school var.								-
% of expl. total var.								-
Deviance	296.092	295.220	295.184	295.088	294.538	295.977	294.841	292.683
Ref. model		0	0	0	0	0	0	0
Difference in -2logl.		.872	.908	1.004	1.554	.115	1.251	3.409
Difference df		1	4	1	1	1	1	1
P value		n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	P<.10
		$\chi^2 = .872$ df=	$\chi^2 = .908$ df=4	$\chi^2 = 1.004$ df=1	$\chi^2 = 1.554$ df=1	$\chi^2 = .115$ df=1	$\chi^2 = 1.251$ df=1	$\chi^2 = 3.409$ df=1
		p=n.s.	p=n.s.	p=n.s.	p=n.s.	p=n.s.	p=n.s.	p<.10

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.36b (continuation): Results of multi-level analyses for incorrectly underlined words in the editing test for Reading ‘EdIncor’, effects of treatment with and without control for covariates (student N=127; class N=6; school N=3) (all continuous predictors grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4
<b>Fixed part</b>					
intercept	.101 (.302)	.081 (.332)	.070 (.332)	.148 (.325)	.177 (.332)
Pre-test editing incorrect	.316*** (.084)	.296*** (.083)	.297*** (.084)	.315*** (.083)	.305*** (.085)
<b>Country of birth mother (ref=The Netherlands)</b>					
Turkey		-.268 (.275)			
Morocco		.358 (.219)			
Surinam		-.093 (.235)			
Aruban		-.045 (.460)			
other country.		-.040 (.226)			
<b>Country of birth father (ref=The Netherlands)</b>					
Turkey			-.139 (.241)		
Morocco			.336 (.236)		
Surinam			-.028 (.253)		
Aruban			-.036 (.340)		
other country.			-.029 (.234)		
Native tongue mother (Dutch)				-.065 (.156)	
Native tongue father (Dutch)					-.106 (.164)
<b>Random part</b>					
Student variance	.561 (.071)	.535 (.068)	.542 (.069)	.560 (.071)	.559 (.071)
School variance	.259 (.222)	.301 (.255)	.295 (.251)	.265 (.227)	.276 (.235)
Total variance	.820	.836	.837	.825	.835
Deviance	296.092	290.660	292.163	295.917	295.687
Ref. model		0	0	0	0
Fit improvement		$\chi^2 = 5.432$ df=5 p=n.s.	$\chi^2 = 3.929$ df=5 p=n.s.	$\chi^2 = .175$ df=1 p=n.s.	$\chi^2 = .405$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.37: Results of multi-level analyses for incorrectly underlined words in the editing test for Reading, effects of covariate ‘age’ (student N=126; class N=6; school N=3) (pre-test and age grand mean centred) (s.e. between brackets)**

Model	0	1
<b>Fixed part</b>		
intercept	.102 (.302)	.100 (.300)
Pre-test editing incorrect	.316 (.084)	.315 (.084)
Age in days		.020 (.136)
<b>Random part</b>		
Student variance	.566 (.072)	.566 (.072)
School variance	.259 (.222)	.256 (.219)
Total variance	.825	.822
Fit (-2loglikelihood)	294.807	294.787
Ref. model		0
		$\chi^2 = .020$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.38: Results of multi-level analyses for incorrectly underlined words in the editing test for Reading ‘EdIncor’, effects of covariate ‘Cito-test’ (student N=96; class N=6; school N=3) (pre-test and Cito-test grand mean centred) (s.e. between brackets)**

model	model 0	model 1
<b>Fixed part</b>		
intercept	-.004 (.291)	-.130 (.134)
Pre-test editing incorrect (gm)	.275*** (.080)	.243** (.081)
Cito-test		-.033*** (.010)
<b>Random part</b>		
Student variance	.150 (.022)	.148 (.022)
School variance	.241 (.206)	.041 (.040)
Total variance	.391	.189
% of expl. student var.		1.333%
% of expl. school var.		82.988%
% of expl. total var.		51.662%
Deviance	101.037	94.799
Ref. model		0
		$\chi^2 = 6.238$
		df=1
		p<.025

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.39: Results of multi-level analyses for incorrectly underlined words in the editing test for Reading ‘EdIncor’, effects of covariate ‘Language spoken with mother’(student N=126; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	.101 (.302)	.079 (.309)
Pre-test editing incorrect	.317*** (.084)	.316*** (.083)
<b>Language spoken with mother (ref. Dutch)</b>		
Half Dutch, half other language		.180 (.189)
Mostly other language		-.113 (.224)
<b>Random part</b>		
Student variance	.566 (.072)	.559 (.071)
School variance	.259 (.222)	.267 (.228)
Total variance	.825	.826
Deviance	294.794	293.389
Fit improvement		$\chi^2 = 1.405$
		df=2
		p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.40: Results of multi-level analyses for incorrectly underlined words in the editing test for Reading ‘EdIncor’, effects of covariate ‘Language spoken with father’ (student N=120; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	.113 (.290)	.075 (.303)
Pre-test editing incorrect	.411*** (.100)	.394*** (.100)
<b>Language spoken with father (Ref. = Dutch)</b>		
Half Dutch, half other language		.274 (.209)
Mostly other language		.008 (.241)
<b>Random part</b>		
Student variance	.580 (.076)	.570 (.074)
School variance	.237 (.205)	.255 (.220)
Total variance	.817	.825
deviance	283.620	281.809
Ref. model		0
Fit improvement		$\chi^2 = 1.811$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Only the Cito variable turned out to have a significant effect (gender only one-sided), but as no effect on gender had been anticipated, it has been taken out (in model 2 and 3). Now the treatment shows a positive effect after controlling for Cito (with or without controlling for gender) The reported effect is based on the  $\chi^2$  of the condition, as the t for models 1 and 3 (model 1,  $t = -.194/.078=2.48$ ; model 3,  $t = -.200/.078=2.56$ ), are both non-significant. This means hox is too strict, since the chi-squares are significant.

**Table 6.41 (final table): Results of multi-level analyses for incorrectly underlined words in the editing test for Reading ‘EdIncor’, effects of treatment after control for covariates (student N=96; class N=6; school N=3) (pre-test and Cito-test grand mean centred) (s.e. between brackets)**

model	0	1	2	3
<b>Fixed part</b>				
intercept	-.228 (.146)	-.073 (.184)	-.130 (.134)	.016 (.174)
Pre-test editing incorrect	.251 (.081)	.217 (.079)	.243 (.081)	.208 (.079)
Cito-test	-.033 (.010)	-.027 (.011)	-.033 (.010)	-.027 (.011)
Gender (girl)	.104 (.079)	.092 (.077)		
Group (experimental)		-.194* (.078)		-.200* (.078)
<b>Random part</b>				
Student variance	.144 (.021)	.134 (.020)	.148 (.022)	.136 (.020)
School variance	.047 (.045)	.075 (.045)	.041 (.040)	.069 (.063)
Total variance	.191	.209	.189	.205
% of expl. student var.		6.944%		8.108%
% of expl. school var.		-		-
% of expl. total var.		-		-
deviance	93.104	87.217	94.799	88.661
Ref model		0		2
Fit improvement		$\chi^2 = 5.887$ df=1 p<.025		$\chi^2 = 6.138$ df=2 p<.025

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

## Effects of the based on differentiated instruction on Writing skills

**Table 6.42. Multi-level analyses for ‘writing’ post-test with pre-test (student N=127; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1	2
<b>Fixed part</b>			
Intercept	1.906*** (.108)	1.896*** (.152)	1.896*** (.152)
Pre-test writing	.597*** (.095)	.572*** (.095)	.572*** (.095)
<b>Random part</b>			
Student variance	1.484 (.186)	1.413 (.182)	1.413 (.182)
Class variance		.070 (.080)	.070 (.080)
School variance			.000 (.000)
Total variance	1.484	1.483	1.483
deviance	410.577	408.593	408.593
Ref. model		0	1
Fit improvement		$\chi^2 = 1.984$ df=1 p=n.s.	$\chi^2 = .000$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

Table 6.42 shows that adding class- or school-level does not yield a significantly improved model fit, so the writing scores will be analysed uni-level.

**Table 6.43a: Results of multi-level analyses for writing, effects of treatment and covariates (student N=127; class N=6; school N=3) (pre-test, time watching TV, time playing games and time reading English grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4	5	6	7
<b>Fixed part</b>								
intercept	1.906** * (.108)	1.561** * (.154)	1.866** * (.287)	1.895** * (.107)	1.897** * (.108)	1.905** * (.108)	1.790** * (.461)	1.769** * (.149)
Pre-test writing (gm)	.597*** (.095)	.561*** (.093)	.595*** (.095)	.547*** (.097)	.584*** (.095)	.595*** (.096)	.596*** (.095)	.591*** (.095)
condition (experimental)		.644*** (.211)						
<b>Eibo (reference group = only in Group 8 (11-12 year olds)</b>								
No Eibo			-.376 (.899)					
Group 7&8 (10-12 year olds)			.090 (.322)					
Group 5 to 8 (8-12 year olds)			-.124 (.356)					
Group 1 to 8 (4-12 year olds)			.874 (.667)					
Time watching TV (1 = never; 5= several hours per day)				.180# (.097)				
Time playing games (1 = never; 5= several hours per day)					.106 (.087)			
Time reading English (1 = never; 5= several hours per day)						.012 (.087)		
Country of birth (the Netherlands)							.122 (.474)	
Gender (girl)								.284 (.215)
<b>Random part</b>								
Student / total variance	1.484 (.186)	1.383 (.174)	1.451 (.182)	1.446 (.181)	1.467 (.184)	1.484 (.186)	1.484 (.186)	1.464 (.184)
% of expl. variance		6.806%		2.561%				
deviance	410.577	401.586	407.701	407.215	409.099	410.559	410.511	408.842
Ref. model		0	0	0	0	0	0	0
Fit improvement		$\chi^2= 8.991$ df=1 p<.005	$\chi^2 =$ 2.876 df=4 p=n.s.	$\chi^2 =$ 3.362 df=1 p<.10	$\chi^2= 1.478$ df=1 p=n.s.	$\chi^2= .018$ df=1 p=n.s.	$\chi^2= .066$ df=1 p=n.s.	$\chi^2= 1.735$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

The treatment turns out to be a significant predictor for the development in writing scores between pre- and post-test ( $t=3.05$ ;  $df=3$ ; one-sided  $p<.05$ ) and the model fit also improves significantly after adding the 'condition' variable to the model ( $\chi^2=8.991$ ,  $df=1$ ,  $p<.005$ ).

**Table 6.43b (continuation): Results of multi-level analyses for writing, effects of treatment with and without control for covariates (student N=127; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1	2	3	4
<b>Fixed part</b>					
intercept	1.906*** (.108)	1.942*** (.149)	1.916*** (.153)	2.082*** (.198)	1.923*** (.198)
Pre-test writing	.597*** (.095)	.583*** (.096)	.584*** (.093)	.609*** (.096)	.599*** (.097)
<b>Country of birth mother (ref=The Netherlands)</b>					
Turkey		-.181 (.450)			
Morocco		.013 (.304)			
Surinam		-.535 (.364)			
Aruban		.799 (.714)			
other country.		.063 (.327)			
<b>Country of birth father (ref=The Netherlands)</b>					
Turkey			-.300 (.369)		
Morocco			.097 (.291)		
Surinam			-.669# (.369)		
Aruban			1.449** (.498)		
other country.			-.032 (.314)		
Native tongue mother (ref=Dutch)				-.252 (.237)	
Native tongue father (ref=Dutch)					-.026 (.239)
<b>Random part</b>					
Total/student variance	1.484 (.186)	1.439 (.181)	1.333 (.167)	1.471 (.185)	1.484 (.186)
Deviance	410.577	406.634	396.908	409.451	410.566
Ref. model		0	0	0	0
Fit improvement		$\chi^2 = 3.943$ df=1 p=n.s.	$\chi^2 = 13.669$ df=5 p<.05	$\chi^2 = 1.126$ df=1 p=n.s.	$\chi^2 = .011$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.44: Results of multi-level analyses for writing, effects of covariate 'age' (student N=126; class N=6; school N=3) (pre-test and Cito-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
Intercept	1.305*** (.140)	1.931*** (.105)
Pre-test writing	.636*** (.094)	.634*** (.093)
Age in days		-.303 (.202)
<b>Random part</b>		
Total/student variance	1.305 (.140)	1.392 (.175)
Deviance	401.483	399.255
Ref. model		0
Fit improvement		$\chi^2 = 2.228$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.45: Results of multi-level analyses for writing, effects of covariate ‘Cito-test’ (student N=96; class N=6; school N=3) (pre-test and Cito-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
Intercept	1.990*** (.131)	1.967*** (.131)
Pre-test writing	.488*** (.116)	.490*** (.116)
Cito-test		-.002 (.017)
<b>Random part</b>		
Total/student variance	1.642 (.237)	1.642 (.237)
deviance	320.066	320.047
Ref. model		0
Fit improvement		$\chi^2 = .019$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.46: Results of multi-level analyses for writing, effects of covariate ‘language spoken with mother’ (student N=120; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	1.913*** (.109)	1.831*** (.129)
Pre-test writing	.595*** (.096)	.604*** (.096)
<b>Language spoken with mother (ref=Dutch)</b>		
Half Dutch, half other language		.194 (.282)
Mostly other language		.391 (.361)
<b>Random part</b>		
Total/student variance	1.495 (.188)	1.479 (.186)
deviance	408.272	406.844
Ref. model		0
Fit improvement		$\chi^2 = .019$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.47: Results of multi-level analyses for writing, effects of covariate ‘language spoken with father’ (student N=126; class N=6; school N=3) (pre-test grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
Intercept	1.942*** (.110)	1.900*** (.130)
Pre-test writing (gm)	.576*** (.096)	.581*** (.098)
<b>Language spoken with father (ref.=Dutch)</b>		
Half Dutch, half other language		.081 (.313)
Mostly other language		.091 (.359)
<b>Random part</b>		
Total/student variance	1.463 (.189)	1.462 (.189)
Fit (-2loglikelihood)	386.195	386.082
Ref. model		0
Fit improvement		$\chi^2 = .113$ df=1 p=n.s.

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)

**Table 6.48 (final table): Results of multi-level analyses for writing, effects of treatment after control for covariates (student N=127; class N=6; school N=3) (pre-test and time watching TV grand mean centred) (s.e. between brackets)**

model	0	1
<b>Fixed part</b>		
intercept	1.926*** (.151)	1.561*** (.192)
Pre-test writing	.543*** (.095)	.510*** (.092)
Time watching TV (1 = never; 5= several hours per day)	.165# (.095)	.135 (.093)
<b>Country of birth father (ref=The Netherlands)</b>		
Turkey	-.226 (.367)	-.351 (.358)
Morocco	.034 (.289)	.178 (.284)
Surinam	-.695# (.365)	-.511 (.359)
Aruban	1.399** (.493)	1.388** (.477)
other country.	-.109 (.313)	-.023 (.305)
condition (experimental)		.607** (.207)
<b>Random part</b>		
Total/student variance	1.302 (.163)	1.220 (.153)
% of expl. variance		6.298%
Deviance	393.935	385.613
Ref. model		0
Fit improvement		$\chi^2 = 8.322$ df=1 p<.01

#=sig at 10% (=5% one sided); \*=sig. at 5%; \*\* sig. at 1%; \*\*\*=sig. at 0.1%. (n.s.=non-significant)  
t-group=2.981; df=6-3-1=2; p<.05 (one-sided)

In the tables with the results of the analyses indicating whether covariates are significant, only 'time watching television' approaches significance. A covariate 'approaching significance' is reported this way because it is significant for the one-sided Wald test, but not significant for IGLS-deviance. Even after controlling for this variable the effect of the treatment remains significant and explains 6.6% variance.

**10.d. : Correlations of Reading skills scores (Editing test-scores for correctly and incorrectly underlined words, 'EdCor' and 'EdIncor')**

Correlations

		SumzEdCorJul	SumzrEdIncJul	SumzrEdCorSO	SumzrEdIncSO
		sum z scores	sum z scores	sum z scores	sum z scores
		Editing Correct	Editing Incorrect	Editing Correct	Editing Incorrect
		July	July	Sept-October	Sept-October
				combined z	combined z
				scores	scores
SumzEd Correct July	Pearson Correlation	1	-,661**	,774**	-,568**
	Sig. (2-tailed)		,000	,000	,000
	N	127	127	127	127
SumzrEd Incorrect July	Pearson Correlation	-,661**	1	-,498**	,501**
	Sig. (2-tailed)	,000		,000	,000
	N	127	127	127	127
Sumzr Ed Correct Sept. – Oct.	Pearson Correlation	,774**	-,498**	1	-,429**
	Sig. (2-tailed)	,000	,000		,000
	N	127	127	145	145
Sumzr Ed Incorrect Sept. – Oct.	Pearson Correlation	-,568**	,501**	-,429**	1
	Sig. (2-tailed)	,000	,000	,000	
	N	127	127	145	145

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

## 11. Delphi Round 1 -Rephrased Complete

### 1. Articulation between levels and connection problems

#### In primary education:

- English teaching starts at different ages(4 - 8 -10);
- the amount of English teaching differs enormously;
- the quality of English teaching differs enormously;
- students have not been offered enough **language learning** strategies and skills;
- more able students have not been challenged enough and enter secondary education ill-prepared for the challenges it brings;
- English language skills are not part of the (Cito) test in the final year;
- Cito scores do not correctly reflect student aptitude and abilities.

#### In secondary education:

- English is not tested upon entry, not even formatively;
- language skills levels can differ from unexposed to the language to (near) Native speaker quality;
- it is difficult to challenge 30-ish children with at least 15 different language skills levels;
- students lack reading skills;
- out of school exposure (like gaming and watching TV) influences language skill development. Sometimes students know, for instance, words like *flint* and *armour* because of online role-playing games;
- students are not offered the extended instruction or additional help they have grown accustomed to;
- some children need to work very hard to meet standard requirements, while others need to wait a few years before they are challenged on their own level;
- the additional materials needed to help children deal with their deficiency are not readily available;
- it takes a year to get all students on the approximately needed level to continue their English training in year 2 of secondary school.

### 2. Solutions for the improvement of articulation between levels

#### Primary education:

- should have English language skills levels tested in the final year
- should have English language skills levels standardised (through lists of vocab and skills)
- should have English language skills levels enforced
- should have English language skills levels standardised and enforced
- should have English language skills levels standardized and tested
- should have English language skills levels standardized, enforced and tested
- should expose children more to the English language. For instance through the internet, songs and video;

#### Secondary education:

- should have English language skills levels tested at the start of the year;
- should offer the students training in study skills;
- should offer the students help (materials, planners, meetings and online activities) to help bridge the gap towards their expected language skills level;
- needs to have less well trained students work harder to meet the required language skills level;

- should offer theme based, cross-curricular activities to have students work more creatively with their knowledge and skills (like the Expo lessons at Wolfert).
- should spend more time on productive skills, like they do in primary education
- should have more meetings a week to deal with these connection problems one way or another.

**In general:**

- For a real solution we need to change education in general: teaching grammar is not what students need in real life and later education. Teaching is outdated in the sense that we teach knowledge and skills that do not relate to what is needed in real life in a way that doesn't fit the current learning styles of the children. Learning outside the classroom is extremely different from the teaching we offer in schools.

**3. Opinion of the treatment in general:**

- The idea to differentiate in itself is great
- The treatment is interesting
- The treatment is doable
- The treatment fits the first year's students willingness to learn
- The treatment offers students the skills and attitudes that relate to today's society.
- The treatment is a nice mix of principles and activities.
- The treatment is really productive and should be continued
- The treatment forces students to produce language, which shows them what they actively need to work on progress.
- The treatment lead me to change course book activities according to the needs and abilities of the children.
- The treatment lead me to differentiate on activity level.

**I am able to do the whole treatment as proposed**

**I am able to do most of the treatment as proposed**

**I am able to largely do the treatment as proposed**

**I am able to partly do the treatment as proposed**

**I am able to hardly do the treatment as proposed**

**I am not able to do the treatment as proposed**

**Reasons for not doing (parts of) the treatment as proposed:**

- I need to have complete ownership of the treatment to be able to implement it
- I need to have more materials
- I need to have more support
- I always change the assignments and activities according to my teaching preferences and student requirements
- I only change my teaching in small steps
- The coursebook is more readily available and easier to work with
- The logistical pressure of organising, assessing and appreciating differentiated levels
- The treatment is too far removed from my every day teaching
- The treatment is too far removed from the teaching approaches used in my school

#### 4. Opinion of the treatment in Parts

##### a. Opinion of the treatment; Tests:

The language skills tests used to measure the learning gains of the treatment were not part of the treatment itself. All participants, in the experimental group as well as the control group, however, reacted positively to the administered test.

Reactions like: "Very interesting to be informed of initial language skills levels and end-of-year year abilities." (Delphi participant A – Jan. 2015) indicate the desire for a test to help, formatively as well as summatively, assess language skills levels of students.

##### b. Opinion of the treatment; Rubrics:

- doable, as long as it is clear and concretely applicable for the students.
- doable, when offered digitally
  
- difficult to consistently use throughout the year;
  
- too abstract for lowest level of the vocational stream;
- too general for children to effectively work with;
- too time consuming (a whole lesson in introduction) to explain how it works;
- too much information in one document;
  
- needs to be made more concrete per topic/grammar point/skill/etc. instead of referring to the general CEFR;
- The use of the rubric becomes repetitive and boring because the intervals in the rubric are too great for students to experience progress when they weekly reflect on the (half-baked) products with the rubric in hand.

##### c. Opinion of the treatment; Chapter wide assignments:

- Doable
- Doable Positive, but difficult to consistently use throughout the year.
- Doable, works well
- the cyclical nature of the assignments made revising work effective
  
- Students need to be trained in this approach, they either do something now, or not do it at all...
- We need to train the children in taking time and revising work to produce the best they have to offer.

##### d. Opinion of the treatment; Cooperative learning:

###### Working together:

###### 1. we have positive experiences with cooperative learning

- Cooperative learning is already in use in HAVO 4, so why not start in year 1
- Cooperating in the sense of helping each other does work
- When children experience they need each other's input and work in order to perform as expected (like in a Jigsaw activity) they are more motivated to invest in a cooperative learning activity.

###### 2. we have negative experiences with cooperative learning:

- difficult to consistently use throughout the year.
- Differences in ability vary too much to make CL effective.
- Teacher skills and student cooperation are a prerequisite for the effectiveness.

- Students need a lot of structure and the freedom given within CL activities might be quite a challenge.
- 24+ students in a classroom impedes the possibilities you have
- lower levels are unable to help each other to progress

**Giving feedback:**

- Giving feedback is a skill that needs to be trained or preferably has been introduced to the pupils in primary schools
- Some students had a problem with giving and/or receiving feedback.
- Social structures and group chemistry impedes effectively giving feedback.
- Peer feedback is already in use in HAVO 4, so why not start in year 1
- You might be able to give structure to the requested feedback based on a form where the possible sentences are (partly) given. You then give VMBO possible compliments or points to work on where students only have to choose or circle. HAVO and VWO can be asked to comment on qualities of the products based on introductory sentences.
- Peer assessment and feedback are not given fairly because of favouritism
- Clarity of instructions, including examples and choice, help to bridge the gap with the missing metalinguistic abstract terminology needed to effectively give feedback on products.

**e. Opinion of the treatment; Differentiated product assessment**

**Positive**

- It is very important to challenge them on their own level.
- We differentiate assessments and marking of products based on achievement.
- As long as differentiated assessment is consistently used and clearly explained, students accept the shift from assessing correctness towards assessing achievement relative to commitment.
- This works when consistently used.
- we differentiate by offering differently streamed tests to the student.
- Doable: there is however the danger that other factors cloud, or colour your judgement.
- This works best if continued right down to the final and central exams. Why allow children to perform far beyond attainment targets if you cannot reward them with the appropriate level diploma for the subject.

**Negative**

- too complicated
- unfair

**f. Opinion of the treatment; Choice offered in process/product:**

- It works; it helps to engage students
- We do it and are pleased with the results
- Nice and doable.
- Doable It helps a bit motivationally speaking, but the choices given did not make them owner of the products. They still feel it is something the teacher/school/coursebook requires from them.
- In interaction we individually differentiate between what each student must do and which activities and exercises are optional
- Coaching is very important when you offer choice, because the choices made by students need to relate to the most effective learning process for each individual.

**5. How to implement the offered differentiated approach:**

**The panel says the following things would make it easier to differentiate teaching:**

- Working with tablets/iPads
- Class size of no more than 24 students
- Class formation while taking ability levels and percentage of students with behavioural problems into account
- Classes timetabled together to enable reorganisation in ability groups.

**When you work with rubrics to involve students in the educational process**

- Rubrics need to be phrased on the language level of the students
- Rubrics need to show example language for each level
- Rubrics need videos to help explain the levels and its use

**Teacher involvement:**

- Teachers need intensive training in differentiated teaching before implementation
- Teachers need to have complete ownership of the differentiated teaching approach.
- Teachers need the incentive of mutual dependency and agreement on differentiated teaching
- Teachers need to have assignments and materials offered beforehand
- Teachers need to work together to develop assignments and materials
- Teachers need help, guidance and coaching to develop assignments and materials
- Teachers need help, guidance and coaching in the implementation of differentiated teaching
- Coaching in the implementation needs to be available on demand.
- Teachers need a reader or planner that weekly tells them what to do.

**Practical introduction:**

- Teachers all need to buy into the new approach in order to change teaching in one go with the whole team working together.
- Teachers need to showcase the new approach in one or two groups, before implementing it in the rest of the rest of the school
- Teachers need extra time, or hours outside of the normal meetings, to help introduce the extra materials and procedures like rubrics and assignments.
- The introduction into the treatment needs to be plenary for students.
- Progressively introduce and teach students the expected self-responsibility and study-skills
- Based on experience with ability grouping, the advice is to differentiate with the whole year, stream or even school. A unique position for a subject in an applied approach will always lead to controversy on organisational/didactical/pedagogical levels
- We should have the final exams differentiated according to ability instead of stream. Otherwise in class differentiation is only superficial and does not allow children to show their full potential.

**6. ...or another differentiated approach:**

- There is not one method that will work for all students, so the magical word is tailor-made education
- No, other approaches are either too general and abstract or too all-embracing to enable effective implementation without too much training and development.
- Flip the classroom is very interesting.
- The use of more online materials might help differentiate

**7. Anything else you would like to add:**

- The differences between the secondary school streams should be taken into account: VMBO students need more and clear structure and guidance before they can get to work than HAVO/VWO students.
- The desired end-result of knowledge and skills should be packaged as a product on which the children structurally work throughout the year: this gives their invested time and effort aim as well as something to hold on to.
- I am afraid the children are already too used to the Dutch 'zesjescultuur', or, 'culture of C's'. Ownership of the educational process is alien to them and demands more involvement in schoolwork than what they have grown accustomed to.
- The workload is keeping us from wholeheartedly joining experiments like this; the number of classes, number of children in classes, preparation, correction and other paperwork weighs down heavily on us.

## 12. Delphi Round 1 -Rephrased and summarised for round 2

### Introduction to round 2

Below you will find a summary of all opinions and arguments won in the first interview round.

The complete listing of all that has been offered by all participants can be found in the document: 'Delphi Round 1 -Rephrased Complete'. Both the summary as well as the complete list of all contributions are offered to you as preparation for the second round of interviews.

Although the aim of the interview rounds is to find consensus on this complex problem, it is important for you to gain insight in the reasoning of the rest of the panel. This allows you to reconsider your point of view or further build your argumentation to state your case.

Round 2 of the interviews consists of a recording of your reactions to the propositions found below, instead of open questions as used in round 1. The request is to consider the propositions and arguments on their merits, with the document including the complete list as background information on the argumentation. It is up to you to decide whether (and/or to which extent) you agree or disagree with what you find below. In both cases you are free to restate, change or add arguments to help clarify your position, or challenge the validity of what others have claimed.

### 1. Articulation between levels and connection problems

#### The panel seems to agree on the following propositions concerning primary education:

- a. English language skills levels differ enormously; from 'unexposed' to 'years ahead of their peers'
- b. Difference in English language skills levels originates from:
  - the start of English teaching at different ages(4 - 8 -10);
  - differences in amount of English teaching;
  - differences in quality of English teaching ;
  - differences in approach to English teaching;
- c. Students lack language learning strategies and skills.
- d. More able students have not been challenged enough in primary education and enter secondary education ill-prepared for the challenges it brings.
- e. 'Cito' scores do not correctly reflect student aptitude and abilities, nor do they offer information on student's English language skills levels.

#### The panel seems to agree on the following propositions concerning secondary education:

- a) English language skills levels are not tested upon entry, not even formatively.
- b) It is difficult to challenge 30-ish children with at least 15 different language skills levels.
- c) Out of school exposure (like gaming and watching TV) influences language skill development. Sometimes students know, for instance, words like *flint* and *armour* because of online role-playing games;
- d) Students are not offered the extended instruction or additional help they have grown accustomed to in primary education.
- e) Additional materials needed to help children deal with their deficiency are not readily available

### 2. Solutions for the Connection problems

#### The panel seems to agree on the following propositions:

- a. English language skills levels should be tested
- b. English language skills levels should be standardized for the end of primary education.
- c. We should offer the students help (extra time, study skills, materials, planners, meetings and online activities) to help bridge the gap towards their expected language skills level.

- d. We should offer theme based, cross-curricular activities to have students work more creatively with their knowledge and skills.
- e. We should spend more time on productive skills, like they do in primary education.

**3. Opinion of the treatment in general:**

On the treatment as a whole the panel was rather divided. On one end of the spectrum we find qualifications like:

“...is really productive and should be continued.”

while on the other end:

“...is too far removed from everyday teaching...”.

With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and **restate your point of view** (including the most important argument(s))?

**4. Opinion of the treatment in Parts**

**a. Opinion of the treatment; Tests:**

The language skills tests used to measure the learning gains of the treatment were not part of the treatment itself. All participants, in the experimental group as well as the control group, however, reacted positively to the administered test.

Reactions like: “Very interesting to be informed of initial language skills levels and end-of-year year abilities.”(Delphi participant A – Jan. 2015) **leads to the following proposition:**

**Primary and secondary education need an English language skills test to measure student language skills levels at the end of primary education/start of secondary education.**

**b. Opinion of the treatment; Rubrics:**

In general, the panel was rather critical about the possibilities the rubrics seemed to offer. This leads to the following proposition:

**Rubrics can only be used when they contain small, but clear and concrete steps (e.g. through instructional videos) that the children can relate to.**

**c. Opinion of the treatment; Chapter-wide assignments:**

**Chapter-wide assignments can be introduced when the students are trained to take time and revise in order to produce the best they have to offer.**

**d. Opinion of the treatment; Cooperative learning:**

On the topic of **Cooperative learning** the panel was rather divided. On one end of the spectrum we find qualifications like:

*Students do it in year 4, so starting in 1 should be possible; they are especially motivated when they experience that they need each other’s input to perform(like in jigsaw activities*

while on the other end:

*Student abilities vary too much to make it an effective learning activity.*

One argument, although completely valid, has been left out, because it is posed an organisational prerequisite. The first question concerning this point is whether you agree that **24+ students in a classroom makes it very difficult, if not impossible, to effectively offer cooperative learning activities?**

The second question goes back to cooperative learning in general. With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and **restate your point of view** (including the most important argument(s))?

**On 'giving feedback' the panel seems to agree on the following propositions:**

- i. Giving feedback is a skill that needs to be trained or preferably has been introduced to the pupils in primary schools
  - ii. Giving feedback needs to be offered in a very structured way, with scaffolding to help the students that do not yet have the words to do so effectively.
  - iii. Favouritism needs to be consistently dealt with.
- e. Opinion of the treatment; Differentiated product assessment**

On the topic of **differentiating the assessments of products** the panel was rather divided. On one end of the spectrum we find qualifications like:

*"Works when it is well explained, consistently used and continued including the final exams",*

while on the other end:

*"It is too complicated and unfair to the students that have to work harder although they are ahead of the rest".*

This discussion touches upon the essence of the difference between 'regular' and differentiated education. When we no longer offer all children the same teaching and the same standards

With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and **restate your point of view** (including the most important argument(s))?

**f. Opinion of the treatment; Choice offered in process/product:**

**On 'offering choice in process/product' the panel seems to agree, with the prerequisite that you need to coach the students into their having their individual choices relate to the most effective learning process.**

**5. How to implement the offered differentiated approach:**

**The panel says the following things would make it easier to differentiate teaching:**

- Working with tablets/iPads
- Class size of no more than 24 students
- Class formation while taking ability levels and percentage of students with behavioural problems into account
- Classes timetabled together to enable reorganisation in ability groups.

On the topic of **'teacher involvement'** the panel broadly agreed on **the following propositions:**

- Teachers all need to buy into the new approach in order to change teaching in one go with the whole team working together.
- Teachers need to have complete ownership of the differentiated teaching approach.

- Teachers need intensive training in differentiated teaching before implementation
- Teachers need help, guidance and coaching when collaboratively developing assignments and materials
- Teachers need help, guidance and coaching in the implementation of differentiated teaching
- Coaching in the implementation period needs to be available on demand.
- Teachers need extra time, or teaching-hours outside of the normal meetings, to help introduce the extra materials and procedures like rubrics and assignments.
- Progressively introduce and teach students the expected self-responsibility and study-skills
- Teachers need the incentive of mutual dependency and agreement on differentiated teaching

#### 6. **Teacher knowledge, skills and attitude:**

This Delphi research has been used to answer the research questions of skills (future) teachers would need to have to effectively differentiate their teaching. Please feel free to **add, or counter-argument the skills listed**. From the panel's responses in round 1 the following teacher skills have been compiled:

##### **Knowledge**

###### **Teachers need to have in-depth knowledge of:**

- different possibilities to differentiate (maybe even expertise)
- the planned English teaching curriculum (to allow freedom of adaptation)
- new developments
- possibilities to expand their pedagogical repertoire,

##### **Skills**

###### **Teachers need to be able to:**

- develop, change or find (online) additional assignments, materials, planners and activities
- completely control classrooms (discipline and order)
- organise activities on different levels,
- give instructions on different levels
- teach language learning strategies
- teach language study skills
- teach self-responsibility
- give feedback on different levels
- train the children in taking time and revising work to produce the best they have to offer
- offer theme based activities
- offer cross-curricular activities
- offer speaking and writing activities frequently and consistently
- assess achievements and progress on different levels

##### **Attitude**

###### **Teachers need to be:**

- appreciative of the different language skills levels
- willing to change assignments and activities because of student needs and abilities
- willing to constructively communicate with primary education,
- willing to constantly develop and expand their pedagogical repertoire.

7) **Discussion:**

The panel seems to be divided on topics like 'cooperative learning' (2d), differentiating product assessments (2e) and the treatment in general (3). The divergence in arguments and experience appears to be largely based on a differing wider view on education. With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and **give your point of view on this observation** (including the most important argument(s))?

**13. Delphi Round 2-Rephrased and summarised for round 3 divided in in italics (round 1) and italics as well as bold (round 2).**

**Introduction to round 2**

Below you will find a summary of all opinions and arguments won in the first two interview rounds. The complete listing of all that has been offered by all participants can be found in the document: 'Delphi Round 2 -Rephrased Complete'. Both the summary as well as the complete list of all contributions are offered to you as preparation for the third Delphi round. Although the aim of the interview rounds is to find consensus on this complex problem, it is important for you to gain insight in the reasoning of the rest of the panel. This allows you to reconsider your point of view or further build your argumentation to state your case.

Round 3 of the interviews consists of a recording of your reactions to the propositions found below, instead of open questions as used in round 1. The request is to consider the propositions and arguments on their merits, with the document including the complete list as background information on the argumentation. It is up to you to decide whether (and/or to which extent) you agree or disagree with what you find below. In both cases you are free to restate, change or add arguments to help clarify your position, or challenge the validity of what others have claimed.

**1. Articulation between levels and connection problems**

**The panel seems to agree on the following propositions concerning primary education:**

- a. *English language skills levels differ enormously; from 'unexposed' to 'years ahead of their peers'*
- b. *Difference in English language skills levels originates from:*
  - *the start of English teaching at different ages(4 - 8 -10);*
  - *differences in amount of English teaching;*
  - *differences in quality of English teaching ;*
  - *differences in approach to English teaching;*
- c. *Students lack language learning strategies and skills.*
- d. *More able students have not been challenged enough in primary education and enter secondary education ill-prepared for the challenges it brings.*
- e. *'Cito' scores do not correctly reflect student aptitude and abilities, nor do they offer information on student's English language skills levels.*
- f. ***students aware of their lacking English language education, feel their deficiency as an unbridgeable gap and sometimes develop inhibitions towards foreign language learning.***

**The panel seems to agree on the following propositions concerning secondary education:**

- a) *English language skills levels are not tested upon entry, not even formatively.*
- b) *It is difficult to challenge 30-ish children with at least 15 different language skills levels.*
- c) *Out of school exposure (like gaming and watching TV) influences language skill development. Sometimes students know, for instance, words like flint and armour because of online role-playing games;*
- d) *Students are not offered the extended instruction or additional help they have grown accustomed to in primary education.*
- e) *Additional materials needed to help children deal with their deficiency are not readily available*
- f) ***information acquired upon entry does not automatically lead to a differentiated approach;***
- g) ***in the current approach to English teaching some children need to work very hard to***  
***Solutions for the Connection problems***

## 2. Solutions for the Connection problems

The panel seems to agree on the following propositions:

- a. *English language skills levels should be tested*
- b. *English language skills levels should be standardized for the end of primary education.*
- c. *We should offer the students help (extra time, study skills, materials, planners, meetings and online activities) to help bridge the gap towards their expected language skills level.*
- d. *We should offer theme based, cross-curricular activities to have students work more creatively with their knowledge and skills.*
- e. *We should spend more time on productive skills, like they do in primary education.*

## 3. Opinion of the treatment in general:

On the treatment as a whole the panel was rather divided. On one end of the spectrum we find qualifications like:

*"...is really productive and should be continued."*

while on the other end:

*"...is too far removed from everyday teaching..."*.

With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and restate your point of view (including the most important argument(s))?

- ***The treatment is doable, but hard work***
- ***The treatment is more child centred as opposed to taking the whole group through the book***
- ***The positive attitude towards school and the subject with which students enter the classroom is more important than cognitive achievements and the individual attention given helps to aid that positive attitude.***
- ***I need to have more time***
- ***You need student cooperation***
- ***Students can fail or be disappointed after being challenged on their own level***

## 4. Opinion of the treatment in Parts

### a. Opinion of the treatment; Tests:

The language skills **tests** used to measure the learning gains of the treatment were not part of the treatment itself. All participants, in the experimental group as well as the control group, however, reacted positively to the administered test.

*Reactions like: "Very interesting to be informed of initial language skills levels and end-of-year year abilities."(Delphi participant A – Jan. 2015) leads to the following proposition:*

*Primary and secondary education need an English language skills test to measure student language skills levels at the end of primary education/start of secondary education.*

***Testing should be used formatively. Not just at the threshold of secondary education, but as part of the educational process.***

### b. Opinion of the treatment; Rubrics:

In general, the panel was rather critical about the possibilities the **rubrics** seemed to offer. This leads to the following proposition:

*Rubrics can only be used when they contain small, but clear and concrete steps (e.g. through instructional videos) that the children can relate to.*

**c. Opinion of the treatment; Chapter-wide assignments:**

*Chapter-wide assignments can be introduced when the students are trained to take time and revise in order to produce the best they have to offer.*

**d. Opinion of the treatment; Cooperative learning:**

On the topic of **Cooperative learning** the panel was rather divided. On one end of the spectrum we find qualifications like:

*Students do it in year 4, so starting in 1 should be possible; they are especially motivated when they experience that they need each other's input to perform (like in jigsaw activities)*

*while on the other end:*

*Student abilities vary too much to make it an effective learning activity.*

- 1. we have positive experiences with cooperative learning**
  - **Cooperative learning is more than just having the stronger help the weaker students, which becomes ineffective over time.**
  - **For effective CL activities the teacher needs to be in control but accept the fact that not 100% of the children are constantly 100% focussed during a CL activity. The number of unfocussed children will be less than in lecture type teaching, but as a teacher you won't notice.**
  - **CL will only work when it is a commonly used activity within the school**
  
- 2. we have negative experiences with cooperative learning:**
  - **progress of higher level students is delayed by lesser students**
  - **the age group is more sensitive to the social context, group, and peer, pressure and chemistry**
  - **teachers need to keep track of personal learning gains within the group achievement, especially when the ability levels differ greatly.**

One argument, although completely valid, has been left out, because it is posed as an organisational prerequisite. *The first question concerning this point is whether you agree that 24+ students in a classroom makes it very difficult, if not impossible, to effectively offer cooperative learning activities?*

***Would make organising CL easier, although the physical limitations, like the size of the classrooms, and the skills of the teachers weigh heavily on this prerequisite.***

The second question goes back to cooperative learning in general. With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and restate your point of view (including the most important argument(s))?

*On 'giving feedback' the panel seems to agree on the following propositions:*

- iv. *Giving feedback is a skill that needs to be trained or preferably has been introduced to the pupils in primary schools*
- v. *Giving feedback needs to be offered in a very structured way, with scaffolding to help the students that do not yet have the words to do so effectively.*
- vi. *Favouritism needs to be consistently dealt with.*

- vii. *Giving feedback is a skill we need to offer as it is needed in real life as well*
- viii. *Dealing with favouritism is hard work although it is part of the pedagogical responsibility.*
- ix. *Teachers need to invest enormous amounts of energy in teaching the children to give honest and constructive feedback, to prevent demotivation through publicly and harshly given feedback.*
- x. *You can teach the children the language needed to give feedback by starting off with substitution tables, because asking them to 'give a compliment' is too open.*

**e. Opinion of treatment; Differentiated product assessment**

On the topic of differentiating the assessments of products the panel was rather divided. On one end of the spectrum we find qualifications like:

*"Works when it is well explained, consistently used and continued including the final exams",*

while on the other end:

*"It is too complicated and unfair to the students that have to work harder although they are ahead of the rest".*

This discussion touches upon the essence of the difference between 'regular' and differentiated education. When we no longer offer all children the same teaching and the same standards

With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and restate your point of view (including the most important argument(s))?

- *We should differentiate on the goals as well; a girl training to become a hair-dresser should, for instance, be offered the opportunity to invest in job related English only*
- *This works best if continued right down to the final and central exams. Why allow children to perform far beyond attainment targets if you cannot reward them with the appropriate level diploma for the subject. On of that, the marks for English on the diploma should be linked to the CEFR, for it to have a real international civil effect.*
- *differentiation on subject level in the final exams would move us towards the Anglo-Saxon educational system.*
- *It is a step away from the culture of C's in which a 'C' is good enough (Dutch: 'zesjes-cultuur')*
- *It would be unfair to not challenge students with a head start; high marks for inferior performance is not a success experience*

*In itself a good idea, but when it is an exception to the rest of the system, it is ineffective and costs an enormous amount of time and effort.*

**f. Opinion of the treatment; Choice offered in process/product:**

*On 'offering choice in process/product' the panel seems to agree, with the prerequisite that you need to coach the students into their having their individual choices relate to the most effective learning process.*

- *Offering choice is very important, we can no longer organise education the way we ourselves experienced it.*

- **You can also offer choice in part of the instruction by offering, or referring to, 'flip the classroom' like explanations online(e.g. [www.meestergijs.nl](http://www.meestergijs.nl))**

**5. How to implement the offered differentiated approach:**

The panel says the following things would make it easier to differentiate teaching:

- *Working with tablets/iPads*
- *Class size of no more than 24 students*
- *Class formation while taking ability levels and percentage of students with behavioural problems into account*
- *Classes timetabled together to enable reorganisation in ability groups.*

**Practical introduction:**

- ***Teachers need to understand differentiation is not doing the same things you did before, but now on as many different moments as there are children in the group.***
- ***Teachers need to feel safe in their team to wholeheartedly join in a new development and share positive as well as negative experiences.***
- ***Weekly progress meetings are the least you need as support.***
- ***To really cooperate it would be better to timetable classes together.***
- ***Developing, working, observing and evaluating as a team is an important prerequisite for continuous team development***

On the topic of **'teacher involvement'** the panel broadly agreed on **the following propositions:**

- *Teachers all need to buy into the new approach in order to change teaching in one go with the whole team working together.*
- *Teachers need to have complete ownership of the differentiated teaching approach.*
- *Teachers need intensive training in differentiated teaching before implementation*
- *Teachers need help, guidance and coaching when collaboratively developing assignments and materials*
- *Teachers need help, guidance and coaching in the implementation of differentiated teaching*
- *Coaching in the implementation period needs to be available on demand.*
- *Teachers need extra time, or teaching-hours outside of the normal meetings, to help introduce the extra materials and procedures like rubrics and assignments.*
- *Progressively introduce and teach students the expected self-responsibility and study-skills*
- *Teachers need the incentive of mutual dependency and agreement on differentiated teaching*
- ***teacher involvement leads to quality***
- ***Top down changes will not stick***
- ***Always involve the teacher***
- ***Huge changes like these need to be carried by the whole team; actually the choice would be to participate or find another school.***
- ***Teachers need to work together to develop assignments and materials, as this is highly effective.***
- ***Coaching in the implementation needs to be available on demand, as coaching and guidance are the most important ingredient in this mix.***

## 6. Teacher knowledge, skills and attitude:

This Delphi research has been used to answer the research questions of skills (future) teachers would need to have to effectively differentiate their teaching. Please feel free to **add, or counter-argument the skills listed**. From the panel's responses in round 1 the following teacher skills have been compiled:

### Knowledge

#### Teachers need to have in-depth knowledge of:

- *different possibilities to differentiate (maybe even expertise)*
- *the planned English teaching curriculum (to allow freedom of adaptation)*
- *new developments*
- *possibilities to expand their pedagogical repertoire,*
- ***In depth knowledge of what is expected per level: BB/BK/TG/TL/H/V - A1/A2/B1/B2/C1, because that***
- ***I agree with the list of desired knowledge, but it is all time consuming and expensive.***

### Skills

#### Teachers need to be able to:

- *develop, change or find (online) additional assignments, materials, planners and activities*
- *completely control classrooms (discipline and order)*
- *organise activities on different levels,*
- *give instructions on different levels*
- *teach language learning strategies*
- *teach language study skills*
- *teach self-responsibility*
- *give feedback on different levels*
- *train the children in taking time and revising work to produce the best they have to offer*
- *offer theme based activities*
- *offer cross-curricular activities*
- *offer speaking and writing activities frequently and consistently*
- *assess achievements and progress on different levels*
- ***Use of internet to find language learning activities and tools on sites.***
- ***Being 100% in control does not mean your students are silent and focussed the whole time you teach. It means you offer the best learning experience possible in the current situation and are actually aware of the moments you have to let go of your plan because the situation (children, colleagues, out-of-class incidents, current events, etc..) and take care of the group and its individuals before you can think about your subject again.***

### Attitude

#### Teachers need to be:

- *appreciative of the different language skills levels*
- *willing to change assignments and activities because of student needs and abilities*
- *willing to constructively communicate with primary education,*
- *willing to constantly develop and expand their pedagogical repertoire.*

***The lists have added value and credibility because they have been constructed by colleague English teachers in the field.***

**8) Discussion:**

The panel seems to be divided on topics like 'cooperative learning' (2d), differentiating product assessments (2e) and the treatment in general (3). The divergence in arguments and experience appears to be largely based on a differing wider view on education. With, or without, the help of the complete list of reactions in the other document, would you be so kind as to weigh and **give your point of view on this observation** (including the most important argument(s))?

***We all are stuck in an English teaching approach which is part of the whole educational system in the Netherlands; this system is more important than the individual student's needs.***

***The way we now teach English has students grow accustomed to achieving through short cycle vocab and grammar tests which are not directly related language skills;***

***Most importantly we need to let go of the coursebook and take responsibility for the educative process in our classrooms instead of having our teaching dictated to us by a publisher.***

***I am sometimes cynical about 'new' developments that have been tried and tested before...***

***I am between the depicted juxtaposed teaching with fixed mind-set and growth mind-set. I am all for the individual growth and attention, but what the students could achieve is sometimes not what they want to achieve.***

***Challenging students on their individual level and growth according to ability is an ideal situation according to most, but honestly speaking almost everyone adapts to the current system and most effort goes to keeping everyone inside the expected bandwidth.***

***The system as we use it prevents us from change towards the desired approach. The problem with that position is that it is a vicious circle: society demands something from us, but we are society. Within the current system our students are expected to perform within the bandwidth of the stream they are in. A lot of effort goes to underachievers we help to climb towards the minimally required standard, leaving those that need no extra help to their own devices. This situation might contribute to a 'culture of C's'.***

***The Dutch 'zesjescultuur', or, 'culture of C's' is very present in Dutch education and society. Only a small minority of students is genuinely interested in personal development; the majority endures education and is more interested in the social interaction of school life beside the educational process.***

***The publications, ideas and examples for a differentiated approach were known before our school joined in this experiment. Applying these principles and ideas to teaching activities cost a lot of extra time and effort teachers seem unable to put in.***

***Teaching students to work with their personal aptitudes, abilities, limitations and learning style preferences is more important than cognitive achievements.***

***If our current teaching had been a success experience for all we would not need experiments like this.***

***I really welcome the possibility for students to take final exams on different levels.***

*I really appreciate the effort taken in this research to have teachers collaborate in the development of a new educational development as well as have them contribute to an open discussion on the effectiveness and feasibility of the experiment worked on.*

*Dutch children are the happiest (<http://www.unicef-irc.org/Report-Card-11>)*

## 14. Delphi Round 3 proposed text

### Introduction to round 3

Below you will find a summary of all opinions and arguments won in the first two interview rounds.

The complete listing of all that has been offered by all participants can be found in the document:

'Delphi Round 2-Rephrased and summarised for round 3 divided in italics(round 1) and italics as well as bold (round 2)', in which all arguments have been sorted according to topic and position. The

summary below has been written as a report to allow nuance in opinion without having to list

differently phrased arguments to clarify positions. Both the summary as well as the complete list of all contributions are offered to you as preparation for the third Delphi round.

Round 3 consists of a request for your written reaction to the summary below.

As stated above, the summary is offered as a story. This means that your personally stated contribution might not be immediately detectable, but is still part of the combined contributions of the panel. Possibly you might find the wording more careful, or circumspect, than how you would have stated your opinion due to the careful wording needed to include the offered nuancing of earlier statements and arguments by other panel members.

The aim of this third round is to ratify the story this panel has to tell about the connection problem and differentiation as possible solution. In that light the request is to consider the story, its positions and arguments on their merits, with the document including the complete list as background information on the argumentation. Please respond with an affirmative e-mail if your position and arguments are adequately represented in the story. Should you feel your position has been violated or misrepresented in the story, than please feel free to mail me how your position or arguments might be restored to their intended contribution to the story.

The panel agrees there is a **connection problem** between primary education and secondary education in general and for English teaching in particular. English language skills levels of students entering the first year of secondary education differ enormously. These differences, by one of the panel-members

described as 'unexposed' to 'years ahead of their peers', in English language skills levels have grown in the past decade because of vvto. Different starting moments (at ages 4, 8 or 10) and the related differences in quantity and quality of the received English language teaching in primary school has widened the gap between primary and secondary education. The view on students' English language skills is further obscured by out-of-school exposure (like gaming and watching TV). Students have sometimes, for instance, acquired words like *flint* and *armour* because of their involvement in online role-playing games.

Because of the above we cannot simply discriminate between students having had English, or not. The example has been given of more able students who have not been challenged enough in primary education and enter secondary education ill-prepared for the challenges it brings, because they believed they would be well able to perform. On the other hand there are sometimes students who are painfully aware of the lack of English language education in their primary school career. They feel their deficiency as being an unbridgeable ability gap from their peers and even have, because of that feeling, developed inhibitions towards foreign language learning in general.

On a more general note the panel observed that students coming from primary education lack language learning strategies and skills needed in secondary education.

According to the panel the 'Cito' scores, previously used to indicate the expected success students would have in a certain stream, did not correctly reflect student aptitude and abilities, nor did it offer information on student's English language skills levels. The law has changed in 2014-2015, no longer demanding a test score for an secondary school stream advice, but relying solely on the primary school teachers' indication. This development is viewed with interest by all, because the effects of this change will become apparent in the academic year starting September 2015.

The above does not mean the panel feels the connection problem lies solely on the primary school side of the divide. Secondary education plays a part in this as well.

The teaching approaches used in secondary school do not relate to what the students have experienced in their first eight years of education. Students are not offered the extended instruction or additional help they have grown accustomed to in primary education and additional materials, needed to help children deal with a deficiency, are not readily available. The panel is aware of the fact that students who previously received none, or little, English language teaching, really need to work hard in the first year of secondary school. The same goes for the teachers, who find it difficult to serve and challenge around 30 children with at least 15 different language skills levels. In that sense the panel respects the primary school approach and would like to apply some of its methodology. The general feeling is that more time could be spent on productive skills, like speaking, and students with deficiencies should be helped more structurally.

Ideally the students would be offered more differentiated, theme based, cross-curricular education to have students work more creatively with their knowledge and skills.

A differentiated approach, like the treatment used in the experiment, was considered to possibly be part of a solution for the connection problem. The treatment was generally very well received, with some criticism and notes on the different parts that made up the treatment, which will be dealt with below (highlighted in bold) in separate paragraphs.

One recurrently referred to activity stands out, especially because it was not part of the treatment. All participants, from both the experimental and control group, reacted positively to the pre- and post-**test** used for the experiment. It triggered reactions like: “Very interesting to be informed of initial language skills levels and end-of-year year abilities”(Delphi participant A – Jan. 2015).

This positive attitude towards testing goes beyond the mere measurement of abilities. The idea is to use tests formatively, as part of the educational process, not just as a threshold for secondary education. In that sense the panel would very much welcome and appreciate a clear standard for English language teaching used in primary education. The information of students having had two,

four, or eight years of English language teaching does not give any guarantee of acquired language skills.

**Rubrics** can only be used when the students can relate to them. This means that the language used needs to be more on their level and the steps need to be small, clear and concrete (e.g. through instructional videos).

**Chapter-wide assignments** were well received. The personal involvement of the students in the assignments, process and product helps them to be more engaged. Chapter-wide assignments can be introduced when students are trained to take time and revise their work in order to produce the best they have to offer. Revision work is made effective by the cyclical nature of the assignments.

**Cooperative learning** yielded several different reactions. During the interviews it turned out the panel had different views on what cooperative learning entailed and the experiences with these type of activities differed. All agreed it was hard work, required quality teacher's skills and it would help to limit the number of students in a classroom to approximately 24.

The '**giving feedback**' part of the treatment has been discussed in two ways. The first was giving feedback to peers in the process of learning and cooperation. This was viewed as a real life skill, which needs to be offered and trained in school. Students need to acquire the words with which they can effectively give feedback and have a safe classroom environment to do so.

The second version of giving feedback referred to the feedback given to peers on a product, like a presentation. The inconsistent harshness, fairness, or even favouritism of peers giving feedback means teachers need to invest in the pedagogical climate in the classroom to make it effective.

**Differentiating the assessments of products** was something the panel came to think of as more positive between the second and third round, although some reservations remained. The whole panel agreed this would work best if continued right down to the final and central exams. On top of that, the marks for English on the diploma should be linked to the CEFR, for it to have a real international effect.

Differentiation was also seen as a step away from the *culture of C's* in which the highest possible achievement (an 'A') is seen as not desirable, because a 'C' is good enough (Dutch: 'zesjes-cultuur').

The expressed reservations mainly referred to the effectiveness and investments needed when this level of differentiation is used only in English teaching and an exception to the rest of the subjects.

**Offering choice in process/product** was unanimously agreed upon as being good practice, with the prerequisite that you need to coach the students into having their individual choices relate to the most effective learning process. Choice might also be offered in the instruction by referring to online explanations (flip the classroom)

According to the panel **the implementation of a differentiated approach** relies heavily on the school environment. The panel was clear on the need of the whole team being involved from the start of the implementation process. All teachers need to have complete agreement on, and ownership of, the differentiated teaching approach the school is going to work with and they all need to join in the effort to bring the approach about. Only then will you have the needed safe environment in which teachers are willing to share their successes and failures and join wholeheartedly in work on a new development. To prevent the pitfalls of misconceptions, teachers need intensive training in differentiated teaching before implementation and collaboratively develop the needed materials, assignments and activities, as this is viewed to be highly effective.

On the practical side of the implementation process, the panel indicated the need for frequent and consistent help, guidance and coaching. Structurally planned time for development, weekly progress meetings with the team and coaching on demand were seen as desirable support. Class size, classes timetabled together per subject and class formation, in which ability levels and percentage of students with behavioural problems are being taken into account, would help. Just like digitally available materials, or even in class tablets/iPads.

The panel felt the list of **knowledge, skills and attitude** (future) teachers would need to have to effectively differentiate their teaching was concise and credible because it has been co-constructed

with English teaching colleagues. The lists were seen as mainly the task for teacher training and an arduous part of the life-long learning part of the job. Teachers need to have in-depth **knowledge** of:

- teaching strategies and methodological possibilities, including differentiation,
- the expectations of what the English language teaching curriculum is supposed to offer on the different secondary levels and who this relates to the CEFR per year.

Teachers need to be to have the **skills** that allow them to:

- develop, change or find (online) additional assignments, materials, planners and activities,
- plan, organise, teach, instruct, coach and give feedback on different levels in one classroom,
- teach language learning strategies, study skills and how students can take responsibility for their language learning process,
- control classrooms (discipline and order) to offer the best learning experience possible in the current situation, while being aware of the moments they have to let go of planning because the situation (children, colleagues, out-of-class incidents, current events, etc..) and take care of the group and its individuals before they can think about the subject again,
- train children in taking time to produce work and revise it to produce the best they can achieve,
- offer theme based, cross-curricular activities and speaking and writing activities frequently and consistently,
- relate individual achievements and the quality of products to expected progress and possible achievements.

Teachers need to have a positive **attitude** towards:

- the different language skill levels in a classroom,
- changing assignments and activities to meet student needs and abilities,
- constructive communication with primary education,
- constant development and expansion of their methodological repertoire,
- keeping up with educational developments in general and language teaching in particular.

During the first two rounds a **discussion** arose around the implementation of differentiated language teaching. The divergence in arguments and experience appeared to be based on a general view on education. The 'culture of C's' and is very present in Dutch education and society. Only a small minority of students is genuinely interested in personal development; the majority endures education and is more interested in the social interaction of school life beside the educational process.

The way English is taught at this moment doesn't help to move away from this culture as it has students grow accustomed to achieving through short cycle vocab and grammar tests which are not directly related to language skills. Within the current teaching approach our students are expected to perform within the bandwidth of the stream they are in and a lot of effort goes to underachievers who

are helped to climb towards the minimally required standard, leaving those that need no extra help to their own devices.

Challenging students on their individual level and growth according to ability would be the ideal situation according to most. The majority of secondary school teachers, however, uses the current approach to (language) teaching and offers full frontal, whole class teaching which leads to the same coursebook based tests for everyone at the same time.

Although some panel members are slightly cynical about 'new' developments as offered in the described treatment, the general feeling is that it is important to let go of the coursebook. There is a desire to take responsibility for the educative process in the English language teaching classrooms instead of having teaching prescribed by a publisher. The individual student's needs and abilities are viewed as important enough to move away from current practice. In that sense the panel feels a drive for change. The publications, ideas and examples for a differentiated approach were known before schools joined in this experiment. Applying these principles and ideas to teaching activities, on the other hand, costs a lot of extra time and effort the teachers just seem unable to put in. The current workload keeps them from wholeheartedly starting, joining or continuing experiments like the one at hand; the number of classes, number of children in classes, preparation, correction and other paperwork weigh down heavily on them.

Apart from the pressure from the present workload, there is also a feeling that a sizeable part of the investments needed to have the treatment succeed needs to be shared; at least school wide.

The more pedagogical part of teaching, like giving feedback, cooperative learning, rubrics and differentiation, is something that needs to be common ground for all teachers and students throughout the school. The shared responsibility combined with students who are accustomed to the activities would make a differentiated approach feasible.

Some of the panel members said that maybe the time was right for differentiation. Apart from pressure from the ministry and inspection, the recent pilot to offer students a possibility to take final

exams on different levels per subject was warmly welcomed. Differentiation of teaching in the first year of secondary education might then continue towards another level exam.

To round off, two remarks from the interview beyond the direct scope of the research. The first is concerned with the experience of the experiment and this Delphi research on the implementation of the treatment: *“The collaboration in the treatment and the interviews on its implementation in this*

*research as well as the attempt to find out what the feasibility is of educational reform are really appreciated by the teachers involved.”* And the final point, that did not answer one of the research

questions at hand, but is noticeable in itself, was a point made by one of the panel members: *“Let’s keep in mind the children in the Dutch educational system are the happiest of the developed world.”<sup>46</sup>*

---

<sup>46</sup> <http://www.unicef-irc.org/Report-Card-11>)

## **12. Curriculum Vitae**

Ton de Kraay was born on August 2, 1966 in Rotterdam, The Netherlands. After obtaining his teaching qualification (B.Ed.) for primary education in 1988 he worked in a primary school until 2000. He then started to teach English in secondary education as part of his English language teaching qualification (B.Ed.). After having started working at the Rotterdam University of applied sciences School of Education as English language teaching pedagogy lecturer, he completed a third teaching qualification in 2004 (B.Ed.). Ton is co-author of 'Four-Leaf Language Clover – A tool for teachers of English in Dutch primary schools' (European Language Label winner 2009) and 'And yet all different; handreiking voor een curriculum Engels op de pabo', which have been published during his work for the ICLON foreign language expertise centre at Leiden University.

Since 2005 Ton works closely together with the EarlyBird agency, developing in-service training programmes and an early English teaching minor programme for the primary school department of the school of education of Rotterdam University. He is a member of several national consultative bodies and has contributed to the Pepelino project of the ECML.

He started his PhD research in February 2011 with Prof. Kees de Bot as promotor at the University of Groningen and is based at Rotterdam University where he combines his work as lecturer of language teaching pedagogy for the school of education with his research at the Research centre Urban Talent.