



UNIVERSITY OF  
EASTERN FINLAND

**Aki Kotiranta**

**“BECAUSE NORMAL SCHOOLS ARE KIND OF BORING”**

**A Quasi-Experimental Mixed Methods Study on How Educational LARP  
Affects English Language Learning**

**UNIVERSITY OF EASTERN FINLAND**

**October 2019**

**School of Applied Educational Science and Teacher Education**

<b>Tiedekunta – Faculty</b>		<b>Osasto – School</b>	
Filosofinen tiedekunta		Soveltavan kasvatustieteen ja opettajankoulutuksen osasto	
<b>Tekijät – Author</b>			
Aki Kotiranta			
<b>Työn nimi – Title</b>			
”Because Normal Schools are Kind of Boring” A Quasi-Experimental Mixed Methods Study on How Educational LARP Affects English Language Learning			
<b>Pääaine – Main subject</b>	<b>Työn laji – Level</b>	<b>Päivämäärä – Date</b>	<b>Sivumäärä – Number of pages</b>
Kasvatustiede		22.10.2019	68
	Pro gradu -tutkielma	<input checked="" type="checkbox"/>	
	Sivuainetutkielma	<input type="checkbox"/>	
	Kandidaatin tutkielma	<input type="checkbox"/>	
	Aineopintojen tutkielma	<input type="checkbox"/>	
<b>Tiivistelmä – Abstract</b>			
<p>Tämä tutkimus käsittelee eduLARP opetusmetodia ja kuinka kyseinen opetusmetodi vaikuttaa oppilaitten englannin kielen oppimiseen. Tutkimus pohjautuu eksperimentaalisen oppimisen ja tilannesidonnaisen oppimisen teorioille jotka ovat eduLARP opetuksessa keskiössä. Lisäksi tärkeänä osana ovat roolipeliteoriat kuten recentering sekä roolipelien fyysisyys sekä avatar-oppiminen. Toinen tärkeä teoriapohja koostuu vieraan kielen oppimisen teorioista. Tutkimus on toteutettu monimenetelmällisenä kvasikokeena. Tutkimuksen kvantitatiivinen osuus toteutettiin tanskalaisessa Østerskov Efterskolen koulussa. Kvantitatiivinen osuus koostui kahdesta englannin kielitaitotestistä, jotka pidettiin kaikille koulun oppilaille. Testien välillä oli noin seitsemän kuukautta, ja jälkimmäisten testien tuloksia verrattiin aikaisempiin, jolloin muutos voitiin havaita. Tutkimuksen kvalitatiivinen osuus toteutettiin viimeisen kielitaitotestin jälkeen ryhmähaastatteluna, johon osallistui viisi oppilasta. Haastattelulla haluttiin selvittää, miten oppilaat itse kokevat oman osaamisensa englannin kielessä ja kuinka he kokevat oppineensa englantia. Tutkimuksen monimenetelmällinen osuus tapahtui molempien osuuksien analysoinnin jälkeen, jolloin kvalitatiiviset ja kvantitatiiviset tulokset yhdistettiin uudeksi dataksi. Tästä datasta selvitettiin miten oppilaitten oma kokemus selittää kielitaitotestien tuloksia ja päinvastoin. Tutkimuksen tulokset olivat asetetun hypoteesin mukaisia, kvantitatiivisesti voitiin todeta oppilaitten kehittyneen seitsemän kuukauden aikana englannin kielessä. Kvalitatiivisesti saatiin tietää, että oppilaat oppivat englantia eri tavoilla, riippuen opittavasta kielenosasta (kuunneltu, kirjoitettu, luettu kieli). Kun nämä tiedot yhdistettiin, voitiin luoda taulukko josta käy selville, että Østerskov Efterskolen opetusmetodi auttaa kielenoppimisessa ja lisää oppilaitten pystyvyyden tunnetta sekä kielenkäytön rohkeutta.</p>			
<b>Avainsanat – Keywords</b>			
Monimenetelmällinen tutkimus, pelillistäminen, roolipeli, roolipelimetodi, eduLARP, englannin kielen opetus, englannin kielen oppiminen, englannin kieli, vieraan kielen oppiminen, LARP, opetusLARP			

<b>Faculty</b> Philosophical Faculty		<b>School</b> School of Applied Educational Science and Teacher Education	
<b>Author</b> Aki Kotiranta			
<b>Title</b> ” Because Normal Schools are Kind of Boring” A Quasi-Experimental Mixed Methods Study on how Educational LARP Affects English Language Learning			
<b>Main subject</b> Educational Science and Pedagogy	<b>Level</b> Masters’ thesis	<b>Date</b> 22.10.2019	<b>Number of pages</b> 68
<b>Abstract</b> <p>This study focuses on education LARP and how it affects the English language competence of the students in Østerskov Efterskole. The theoretical background consists mainly of different theories that are the base of educational LARP research, such as experimental learning and situated learning. Furthermore, an important part has been received to gamification, recentering, and avatar learning, as well as foreign language learning (L2). This study is a quasi-experimental mixed-methods study. The quantitative portion of the study uses a pre-test post-test design with nonparametric version of the T-tests, and aims to find out if there is a statistically significant change in the results. The qualitative portion is done by a semi-constructed interview with and analysed with content analysis. The qualitative portion is done to gain a better understanding of how the pupils feel about their English language competence and how they have learned English. The mixed portion of the study merges these two data sets into one and aims to explain how the personal experiences of the students explain the results of language competence tests and vice versa. The results of the study are in line with the set hypothesis as the students showed a statistically significant increase in their English language competence. The result of qualitative analysis showed that the different aspects of language competence (listening, reading and writing) were learned from different sources (formal and informal learning sources). The merging of these two data sets revealed that the teaching method in Østerskov Efterskole helps the students in learning English and enhances their language usage confidence.</p>			
<b>Avainsanat – Keywords</b> Mixed methods research, gamification, roleplaying games, roleplaying teaching method, eduLARP, educational LARP, English language learning, L2 learning, English as Second Language, ESL teaching, ESL learning			

## Contents

1. List of Tables .....	II
2. List of Figures .....	II
3. List of Concepts and Abbreviations.....	III
1. Introduction .....	1
2. Research Questions and General Design .....	8
2.1 Research Questions .....	8
2.2 General Design and Paradigm.....	9
2.2.1 Longitudinal Design.....	9
2.2.2 Experimental Design .....	11
2.2.3 Mixed Methods Design .....	12
2.2.4 General Paradigm of Research.....	14
3. Theoretical Background .....	16
3.1 The Game, Roleplaying, and EduLARP .....	16
3.2 Second Language Acquisition .....	25
3.3 What eduLARP Can Offer to L2 Acquisition and Teaching.....	28
4. Research and Analysis Methods .....	30
4.1 Experimental Mixed Methods / Connection Between the Teaching Method and Language Competence.....	30
4.2 Quantitative / English Language Competence.....	33
4.3 Qualitative / How and Where Is the English Language Learned.....	35
5. Results and Discussion .....	41
5.1 English Language Competence .....	41
5.2 How and Where the English Language is learned.....	54
5.3 Connection of the Learning Source and Language Competence .....	59
5.4 Answers to Research questions .....	62
5.5 Discussion.....	63
6. Conclusion and Ethics .....	64
6.1 Conclusion.....	64
6.2 Ethics.....	65
7. References .....	68

## 1. List of Tables

Table 1. Interaction Between the Curricula and the B1 Level .....	6
Table 2. Cognitive, Affective, and Behavioral Dimensions of Student Learning Through EduLARP	23
Table 3. Types of Group Interviews and Dimensions .....	36
Table 4. Descriptive Statistics of Pre-Tests .....	42
Table 5. Tests of Normality for Reading, Listening, Writing and Total Pre-Test Variables .....	44
Table 6. Descriptive Statistics for Post-Test Data .....	45
Table 7. Test of Normality for Post-Tests .....	48
Table 8. Descriptive Statistics of Reading_Pre and Reading_Post .....	49
Table 9. Ranks for Reading_Post – Reading_Pre .....	49
Table 10. Test Statistics Reading_Post -Reading_Pre .....	49
Table 11. Descriptive Statistics for Listening_Pre - Listening_Post.....	50
Table 12. Ranks for Listening_Pre - Listening_Post .....	50
Table 13. Test Statistics for Listening_Pre – Listening_Post .....	50
Table 14. Descriptive Statistics for Writing_Pre – Writing_Post .....	51
Table 15. Ranks for Writing_Pre - Writing_Post .....	51
Table 16. Test Statistics for Writing_Pre - Writing_Post.. .....	51
Table 17. Descriptive Statistics for Total_Pre – Total_Post .....	52
Table 18. Ranks Total_Pre – Total_Post .....	52
Table 19. Test Statistics for Total_Pre – Total_Post .....	52
Table 20. Spearman’s Rho Results .....	53
Table 21. The Joint Display of Mixed Methods Findings .....	61

## 2. List of Figures

Figure 1. Roadmap of the Research .....	4
Figure 2. The Recentering in Larping (Henriksen 2004, 127).....	21
Figure 3. The Normal Q-Q Plot of Reading_Pre .....	43
Figure 4. Normal Q-Q Plot of Listening_Pre.....	43
Figure 5. Normal Q-Q Plot of Written_Pre.....	43
Figure 6. Normal Q-Q Plot of Total Marks .....	44
Figure 7. Normal Q-Q Plot of Reading_Post .....	46
Figure 8. Normal Q-Q Plot of Listening_Post .....	46
Figure 9. Normal Q-Q Plot of Written_Post.....	47
Figure 10. Normal Q-Q Plot of Total_Post .....	47
Figure 11. The Connections Between Aspects and Learning Styles.....	55

### 3. List of Concepts and Abbreviations

<b>Actant</b>	A person, creature, or object playing an active role in a narrative (Oxford Dictionaries 2018).
<b>EduLARP</b>	<b>Educational Live-Action RolePlaying</b> Live-action roleplaying game with educational purpose and goal (Harviainen & Savonsaari 2013; S. L. Bowman & A. Standiford 2015).
<b>EMMD</b>	<b>Experimental Mixed Methods Design</b> An advanced mixed-methods research design that merges a collection of different qualitative methods with a quantitative experimental method.
<b>Experimental learning</b>	“the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience.” (Kolb 1984, 41.)
<b>Formal learning</b>	Learning that happens in an environment that is specifically constructed and organized this goal in mind. Also, learning is usually goal-oriented, validated and measured. (Cedefop 2008.)
<b>Game</b>	A usually competitive imaginary activity that people take part in their freewill (Calliois 2001).
<b>Gamification</b>	The implementation of game-like design and aspects to a service or teaching (Deterding, Dixon, Khaled & Nacke 2011a, 5; 2011b, 3; Hamari 2015).
<b>Informal learning</b>	Learning that occurs outside Formal learning, such as daily activities and life in general. The learning is not organized or structured. Usually unintentional. (Cedefop 2008.)
<b>LARP</b>	<b>Live-Action RolePlaying</b> A game where the player assumes the role of a character in a fictional world. However, when compared to TRPG, the player fully assimilates the character by, for example: dressing up, moving as the character would, and so on. (Pettersson 2005, 22; Leppälähti 2009; Daniau 2016.)

<b>Longitudinal study</b>	The basic underlying idea of a longitudinal study is to find research data over a period of time. (Cohen, Manion & Morrison 2007, 212.)
<b>MMR</b>	<b>Mixed Methods Research</b> A research method that uses both qualitative and quantitative research methods and then mixes or combines them both together. (Denscombe 2010.)
<b>One Group Pretest-posttest design or Time-series design</b>	<p>“A time-series study involves making multiple observations of one or more subjects or cohorts before and after the introduction of an independent variable. The independent variable may or may not be controlled by the researcher.” (Kirk 2013, 12.)</p> <p>Pretest-posttest design is constructed. Cohen et al (2018) mark the pre-test as <math>O^1</math> where the dependent variable has been measured the intervention or influence as X and the pre-test as <math>O^2</math> where the dependent variable is measured again. Then the researcher must compare the differences between the <math>O1</math> and <math>O2</math> to see if the intervention on X had had any effect. The design can be represented as:  <math>O^1 \Rightarrow X \Rightarrow O^2</math>. (Cohen, Manion &amp; Morrison 2018.)</p>
<b>Paradigm</b>	“a philosophical and theoretical framework of a scientific school or discipline within which theories, laws, and generalizations and the experiments performed in support of them are formulated.” (Merriam-Webster, 2019a.)
<b>Physicality</b>	An idea that players in LARP act as themselves, while still playing their character. (Harviainen & Savonsaari 2013.)
<b>PPD</b>	<b>Pre-test-Post-test Design</b>
<b>Pragmatism</b>	“An American movement in philosophy founded by C. S. Peirce and William James and marked by the doctrines that the meaning of conceptions is to be sought in their practical bearings, that the function of thought is to guide action, and that truth is preeminent to be tested by the practical consequences of belief.” (Merriam-Webster, 2019b.)

<b>Quasi-Experiment or Pre-experimental</b>	The quasi-experimental design is used when some or all details of a true experimental design are not present or could not be achieved, for example when dealing with ethical or practical challenges concerning the research subjects or theme. (Kirk 2013, 6-7.)
<b>Recentering</b>	A persons' ability to imagine other possible worlds that are different from our own and the ability to relate to the fictional characters. (Ryan 1991.)
<b>RPG</b>	<b>Roleplaying Game</b> A game where the player assumes a role of a character in a fictional world. The player acts as the conduit between the gaming world and the character's actions. (Pettersson 2005, 22; Leppälähti 2009; Daniau 2016.)
<b>Situated learning - theory</b>	A theory about learning that states: learning does not happen in a vacuum, but everything around the learner affects the possible outcome of the learning. (Lave & Wenger 1991.)  The theory also states that "learning is an integral and inseparable aspect of social practice." (Lave & Wenger 1991, 30.)
<b>TRPG</b>	<b>Tabletop RolePlaying Game</b> A game where the player assumes a role of a character in a fictional world. The game is usually played as a sort of radio theatre, where one of the players is a storyteller and the rest are the protagonists of the story. (Daniau 2016.)
<b>True-experiment Design</b>	The true-experimental design aims to limit the differences of exposure between the control group and the treatment group to only the specified variable, that only the treatment group receives. Both groups are made to be as homogeneous as possible, and the research situation and all external stimuli are chosen to be as identical as possible between the treatment group and the control group. (Abbot, Lee & McKinney 2013, 253.; Cohen, Manion & Morrison 2018.; Creswell & Plano Clark 2018, 108; Kirk 2013, 6.)



## 1. Introduction

This research studies the possible relationship between the learning results of the students in Østerskov Efterskole in the English language and the eduLARP teaching style. The study is conducted using mixed methods design, and it is a combination of convergent and experimental designs which are explained in more detail in chapter 4.1.3. This research is a spiritual successor of “Roolipelimetodi matematiikan opetuksessa” by Kotiranta (2017), a bachelor’s thesis studying a teaching experiment that used eduLARP method to teach mathematics in an elementary school.

### **Why eduLARP should be researched**

In the Finnish National Core Curriculum for Basic Education 2014, the students are placed in an active role and are involved in the planning and implementing of the teaching, as well as assessment and grading (Opetushallitus 2014, 17). Furthermore, the Finnish National Core Curriculum encourages the teachers and the staff of the school to collaborate and to take part in co-teaching (Opetushallitus 2014, 36). The curriculum states that:

“Oppilaiden osallistuminen oman koulutyönsä ja ryhmänsä toiminnan suunnitteluun on luonteva tapa vahvistaa osallisuutta” (Opetushallitus 2014, 35).

This translates to “The participation of pupils in the planning of school work and group activities is a natural way of strengthening inclusion” (Opetushallitus 2014, 35).

The national curriculum also advises the schools to utilize the subject-wide integration that combines different subjects together. These combinations help the students to understand how different subjects are related to one another, to direct and apply their learning to other subjects, and to provide a feeling of inclusion in a community that constructs knowledge. (Opetushallitus 2014, 31.)

Østerskov Efterskole in Denmark has been using these, or similar, methods and approaches in their teaching for years. The following quotes are taken from an article collection named "Læringsrollespil Undervisningsmetoden på Østerskov Efterskole I Hobro" (Lunau 2014) published by Østerskov Efterskole, where different authors discuss the method used in the school. The following extracts are translated into English for ease of reading by the researcher.

"The teachers work in two teams which take turns teaching one week and preparing and doing evening watches the other week. Every week is set up as a LARP experience by the teaching team of teachers." (Hyltoft 2014, 52.)

"Østerskov Efterskole teaches the pupil to master the context of the surrounding world: processes, systems, and organizations through knowing and mastering his action options and his role options." (Lunau 2014, 3-4.)

Østerskov Efterskole is a boarding school that teaches pupils from age 14 to 17. The school uses a minimum amount of external teaching material, but there are two teams of teachers who plan one week in advance the next week's game and the setting. The school also has a multidisciplinary idea of learning that helps the teachers to cooperate and cover areas of curriculum together. Østerskov Efterskole uses an eduLARP method in their teaching, there are approximately 30+ lessons per week where all lessons take about 45 minutes each, half of the lessons are part of a game that is specific to that week. The game also acts as the overarching theme of the week, which ties all the lessons together. (Hyltoft 2008.) A theme could be anything that the teachers want, for example, Ancient Rome (Østerskov Efterskole 2019).

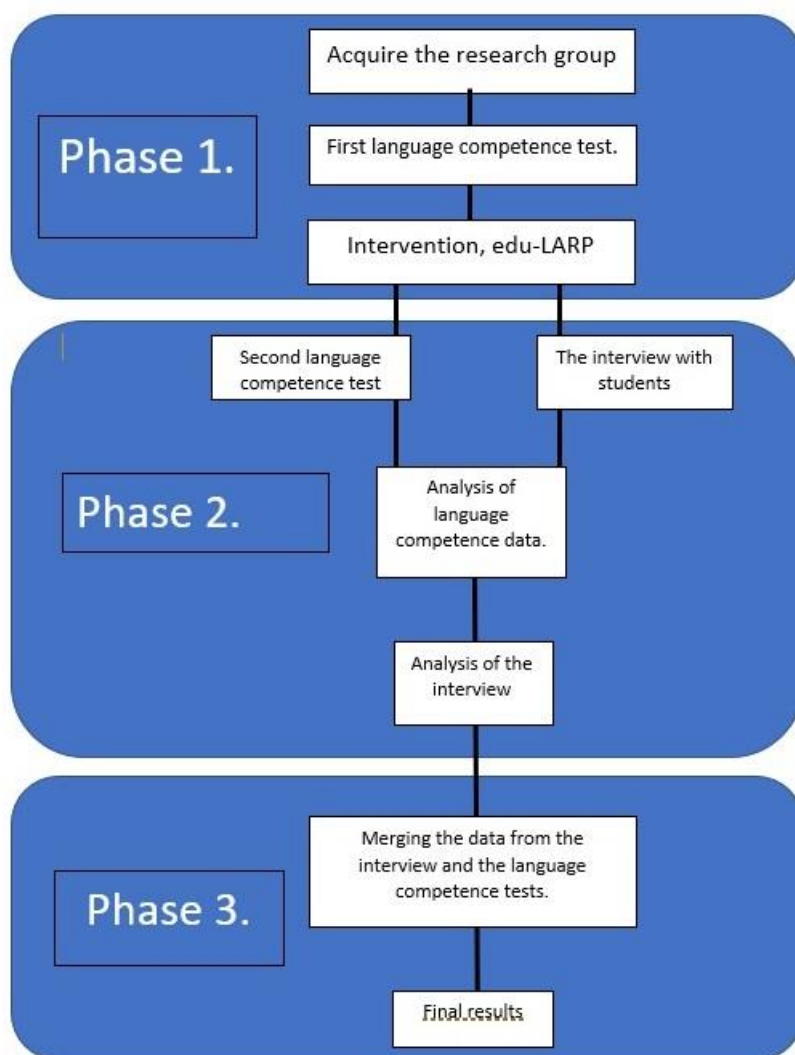
The Østerskov Efterskole and the eduLARP method used there have been studied several times before, for example, Gedde (2013; 2014), Vanek and Peterson (2016), and Hjalmarsson (2011). Furthermore, eduLARP research has been increasing all over the world

(see, Bowman 2014), however, it seems that there has not been a study that aims to find if there is quantitative proof on the actual results of the eduLARP teaching.

The teaching method used in Østerskov Efterskole should be studied in greater detail, to gain a better knowledge of how the eduLARP teaching method works, what are the results of said teaching method, and is there a relation between the teaching method and learning results.

### **The Roadmap**

The research is constructed following the example presented by Creswell and Plano Clark (2017, 266 – 269). This research aspires to follow these instructions and examples. Figure 1 illustrates the roadmap of the research process, that started properly in the spring of 2018 and lasted until the autumn of 2019. The roadmap shows how the research process progressed from acquiring the research group to the final report. The process is divided into three phases, where the first one consists of finding a suitable research group and doing the first English language competency test. The second phase consists of the interview and the second language competency tests, and analyzing the data gathered from the interview and language competency tests. The third phase is reserved for the mixed methods analysis, which means integrating the two data sets that were gathered by the English language competency tests and by the interview. The third phase also includes reporting the results and editing the report.



**Figure 1. Roadmap of the Research**

### **Similarities Between Curricula and B1 Level**

One of the starting points of the study was to compare the Danish national curriculum of English language and culture on the end of seventh grade (see Ministriet for Børn, Undervisning og Ligestilling 2016) to the Finnish national curriculum of English language and culture at the end of seventh grade (see Opetushallitus 2016, 15.4.3) to see if there are any similarities between them. Several similarities were found and those were compared to the standards set by the CEFR and Council of Europe (2018), again, significant similarities were found between all three.

These similarities can be clearly seen in Table 1 below, which is divided into three columns and each column represents a different aspect of English language. These aspects are:

**Writing, Reading, and Listening.** However, there is a cell that covers aspects of reading and listening. This is due to the lack of a specific listening indicators in the Finnish curriculum. Table 1 is also divided into three rows, the first row shows how Finnish curriculum defines a good competence after seventh grade, the second shows how Danish curriculum defines a good competence after seventh grade, and the third row shows the B1 level of competence according to CEFR. The similarities are striking between all three. On writing aspect, the pupil's ability to write a text that conveys a message. On reading and listening aspects the pupil should be able to understand and pinpoint the important part of the text and understand the text satisfactorily.

Table 1. Interaction Between the Curricula and the B1 Level

		Area of competence		
		Writing	Reading	Listening
<b>Finnish curriculum, Pupil competence in the English language after the seventh grade</b>	<p>“ The pupil is able to explain the key points and also some detail of different real-life or fictional topics connected to daily life that interest him or her using a fairly extensive vocabulary and resources of structures as well as some common phrases and idioms. The pupil is able to apply a number of basic rules of punctuation also in expressions that have not been practiced.” (Opetushallitus 2016, 15.4.3.)</p>	<p>“ The pupil understands the main ideas and some details of a clear, nearly regular tempo, standard language speech and popularized written text. The pupil understands speech or written text based on a shared experience or general knowledge. The pupil is able to find the main ideas, keywords, and important detail without preparation.” (Opetushallitus 2016, 15.4.3.)</p>	<p>“ The pupil can understand the main content of simple subjects-texts.</p> <p>The pupil can find specific details in different types of text.</p> <p>The pupil can understand the main content and context of different types of text.” (Ministriet for Børn, Undervisning og Ligestilling 2016.)</p>	<p>“ The pupil can understand the main content of light texts.</p> <p>The pupil can understand details from texts within known topics.</p> <p>The pupil can understand simple, easy-to-understand texts in a natural speaking tempo.” (Ministriet for Børn, Undervisning og Ligestilling 2016.)</p>
<b>Danish curriculum, Pupil competence in the English language after the seventh grade</b>	<p>“ The pupil can write simple, informing texts.</p> <p>The pupil can write small stories.</p> <p>The pupil can write understandable and coherent in English.” (Ministriet for Børn, Undervisning og Ligestilling 2016.)</p>	<p>“ The pupil can understand the main content of simple subjects-texts.</p> <p>The pupil can find specific details in different types of text.</p> <p>The pupil can understand the main content and context of different types of text.” (Ministriet for Børn, Undervisning og Ligestilling 2016.)</p>	<p>“ The pupil can understand the main content of simple subjects-texts.</p> <p>The pupil can find specific details in different types of text.</p> <p>The pupil can understand the main content and context of different types of text.” (Ministriet for Børn, Undervisning og Ligestilling 2016.)</p>	<p>“ The pupil can understand the main content of light texts.</p> <p>The pupil can understand details from texts within known topics.</p> <p>The pupil can understand simple, easy-to-understand texts in a natural speaking tempo.” (Ministriet for Børn, Undervisning og Ligestilling 2016.)</p>
<b>B1 level as set by Council of Europe</b>	<p>“ Can write straightforward connected texts on a range of familiar subjects within his/her field of interest, by linking a series of shorter discrete elements into a linear sequence.” (Council of Europe 2018, 75.)</p>	<p>“ Can read straightforward factual texts on subjects related to his/her field and interests with a satisfactory level of comprehension.” (Council of Europe 2018, 60.)</p>	<p>“ The pupil can understand the main content of simple subjects-texts.</p> <p>The pupil can find specific details in different types of text.</p> <p>The pupil can understand the main content and context of different types of text.” (Ministriet for Børn, Undervisning og Ligestilling 2016.)</p>	<p>“ Pupil can understand straightforward factual information about common every day or job-related topics, identifying both general messages and specific details, provided speech is clearly articulated in a generally familiar accent.” (Council of Europe 2018, 55.)</p> <p>“ Pupil Can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc., including short narratives.” (Council of Europe 2018, 55.)</p>

Consulting Table 1 it can be said that in both curricula the pupils should reach the B1 level in the English language when they graduate from the seventh grade. This comparison is done to show that the findings of this study are comparable to other schools in Finland and Europe, and that the test level of B1 matches the pupils expected competence in English language.

Even though this research is conducted in Finland and the researcher's native language is Finnish, this thesis is written in English. This is because the researcher has more experience in writing academic texts in English than in his native language, and in addition, this way the research report is accessible to a wider audience than just Finnish researchers. Furthermore, since this thesis is studying a Danish school, the report must be understandable to the faculty members and students of Østerskov Efterskole.

All extracts from the Danish curriculum (Ministriet for Børn, Undervisning og Ligestilling 2016) have been translated to English for the ease of reading and understanding. The translation was deemed necessary because the researcher is not fluent in Danish, and because an English version was not available at the time of the study. The translation was made by utilizing two translation programs to minimize the possibility of an error.

## 2. Research Questions and General Design

This section of the report illuminates the research questions and aims set for the study and then goes on to explain the general research design of the study. The research questions are answered in chapter 5.4 and the research design is explored in greater detail in chapter 3.

### 2.1 Research Questions

This study focuses on the teaching of the English language in Østerskov Efterskole in Hobro, Denmark. This focus includes all the students in the school, including special needs students and students with learning difficulties. This focus is broader than first planned, but it offers more diverse data and a greater possibility for analysis.

The aim of this research is to observe the possible connection between the eduLARP the teaching method and the pupils' competence in the English language. This study uses the mixed methods experimental design to integrate the personal experience of the pupils to the quantitative findings. This integration tries to enlighten the reasons behind the results of English language competency tests.

The research questions are as follows:

1. How does the students' competence in English language change during a seven-month period of eduLARP teaching?
2. How and where the pupils have learned the English language?
3. How do the pupils' personal experience of learning English explain results received in English language competency tests?



## 2.2 General Design and Paradigm

The research is conducted as a Longitudinal Mixed Methods Quasi-Experiment. This study aims to firstly, measure the learning results, secondly, study how and where the students have learned English, and finally, attempt to clarify what kind of connection the teaching method has with the results if any. The time frame for the longitudinal study starts during the autumn semester of 2018, the beginning of the school year in Østerskov Efterskole, and ends in the spring semester of 2019. The length is approximately seven months in total.

### 2.2.1 Longitudinal Design

A longitudinal design has been chosen for this research to measure the development of English language competence. The basic underlying idea of a longitudinal study is to find data over a period; however, while the length of the period may vary between studies, the principle remains the same. This study utilizes a cohort study method, where the aim is to follow a specific population over an extended period, on each research point every individual of the research population is measured. (Cohen, Manion & Morrison 2007, 212.)

This research uses both prospective and retrospective methods of a longitudinal study. A prospective method is used when the research population answers a survey, and a retrospective method, when interviewees are looking back on the teaching and what was learned. (Cohen, Manion & Morrison 2007, 212.) This study aims to combine these two by using the experimental mixed methods design.

More specifically the study uses a design of “One Group Pretest-posttest design” (see, Abbott, Lee and McKinney 2013, 254; Cohen, Manion & Morrison 2018, 406) or “Time-series design” (see, Kirk 2013, 12). According to Kirk, a time series design observes the subjects several times prior and after the intervention that acts as the independent variable that can be controlled by the researcher (2013).

This study follows a cohort of approximately 70 pupils over seven months and measures their competence in the English language two times. The independent variable in this study is the education that the pupils receive, and thus it cannot be controlled by the researcher.

Cohen, Manion, and Morrison (2018) provide an example diagram of how the Pretest-posttest design is constructed. They mark the post-test as  $O_1$  where the dependent variable has been measured, the intervention or influence as  $X$ , and the post-test as  $O_2$  where the dependent variable is measured again. Then the researcher must compare the differences between the  $O_1$  and  $O_2$  to see if the  $X$  had had any effect. The design can be represented as:

$$O_1 \Rightarrow X \Rightarrow O_2 \text{ (Cohen, Manion \& Morrison 2018).}$$

Abbot, Lee, and McKinney (2013) introduce five possible complications that the Pretest-posttest design or Time-series design research may have. These complications could affect the research and are necessary to examine for the reliability of the research, these are **History, Maturation, Testing, Instrument, and Attrition**. (Abbott, Lee & McKinney 2013, 255-256)

Firstly, the **History** complication states that the events that happen at the same time with the research might affect the outcomes. Secondly, the **Maturation** complication is where the research subjects, especially when studying children, are grown out or into a phase of development. This might shift their beliefs and responses to variables and affect the research that way. Thirdly, by **testing** the subjects the research might 'tip-off' the underlying goal of the research. Fourth, is the **instrument** complication and it states that the instruments used to measure certain variables are incorrect. Lastly, there is the complication of **attrition**. This happens when test subjects decide to quit the test before the post-test is conducted. (Abbott, Lee & McKinney 2013, 255-256.)

In this study, most of these complications are countered by the short time gap between pre-test and post-test. The gap is approximately seven months, and in this time frame, it is difficult to believe that the study might experience any **maturation** or **attrition** complications. The study, however, is vulnerable to the other three of the complications listed above. **History** complications being the greatest threat to the study, since the students are in constant connection to English language, by internet and other media. To

counter this complication this study uses mixed methods design to see what the impact of the **history** complication on the experiment is. Lastly, two minor threats are **Testing** and **Instrument** complications, however, there is no reason to assume that these might affect the results in any way since the meters used in the tests are certificated language competency tests that were made for this type of use.

### 2.2.2 Experimental Design

The experimental design is in its core a simple design where the treatment group is exposed to the stimulus, that the control group is not. The experimental design aims to limit the differences of exposure between the control group and the treatment group to only the specified variable, that the treatment group receives. This is called a true experimental design. Furthermore, both groups are made to be as homogeneous as possible, the research situation and all external stimuli are shaped to be as identical as possible between the treatment group and the control group. (Abbot, Lee & McKinney 2013, 253.; Cohen, Manion & Morrison 2018.; Creswell & Plano Clark 2018, 108; Kirk 2013, 6.)

However, this study follows a “quasi-experimental” (Cohen, Manion & Morrison 2018; Kirk 2013, 6) or “pre-experimental” design as Abbott and McKinney (2013, 254) name it. The quasi-experimental design is used when some or all the details of a true experimental design are not present or could not be achieved, for example when dealing with ethical or practical challenges concerning the research subjects or theme (Kirk 2013, 6-7).

The quasi-experimental design has been chosen since the topic of the research prevents the usage of true experimental design. For example, the researcher is not in control of how the students learn the English language and are influenced by other sources of English language. Thus, the researcher is not in control of the amount and effect of the exposure. (Cohen, Manion & Morrison 2007, 282).

A true experimental design has been abandoned because the research subjects cannot be divided into a research group and a control group. This is mainly because of the ethical problems caused by the research subject. The students cannot be denied their right to get the teaching promised by the Østerskov Efterskole. To do this would initiate an ethical

conundrum that could possibly undermine the whole research. In addition, the research group is not selected by a truly random selection as it would be selected in a true experimental design, but is selected more by what is available at the time of the research, since all the pupils are part of the study and all the pupils have enrolled for the school in purpose. Possible use of a control group from another school has been rejected because the quantity of the data would become nearly impossible for one researcher to analyze within a reasonable time frame, which could harm the results and the research.

A quasi-experimental design can be greatly affected by nuisance variables, i.e. variables that the researcher cannot control but which still affect the results of the experiment. The nuisance variables can originate from almost anywhere, for example, in this study the greatest impact is produced by the non-informal learning (out of school environment learning) that the students are subjected to. Nuisance variables can distort the research findings so that the seen impact is achieved because of a nuisance variable instead of the desired variable. (Kirk 2013, 5.)

### 2.2.3 Mixed Methods Design

This study uses the mixed methods research design, in which the study uses both qualitative and quantitative methods and then combines the results to form a new data set that is then analyzed (Mark 2015, 4; Mertens & Hesse-Bieber 2013).

Mixed-method research has had multiple names throughout its development. It has been called “Multi-strategy research”, “integrated methods” and “multi-method research”. However, this only shows how many ways the research can be mixed and that mixed methods are only a general term to be used when talking about mixed methods. (Denscombe 2010.) Cohen, Manion, and Morrison explain that the MMR is very elusive for a simple definition. (2018) Therefore, Creswell and Plano Clark (2018) illuminate the multitude of the definitions of the MMR. They state that the definition has shifted focus from the methods used to the most modern focus of methods and contested terrain. (Creswell & Plano Clark, 2018.)

Mixed methods design as a term can be most loosely defined to be a research method that uses both qualitative and quantitative research methods and then mixes or combines them both together (Denscombe 2010). However, this definition is somewhat limited, and it does not give a complete picture on the definition and complexity that is Mixed Methods Research (Or MMR) as Cohen, Manion, and Morrison abbreviate (2018). For example, Denscombe provides more detailed definition where they state that:

“It (MMR) refers to a research strategy that crosses the boundaries of conventional paradigms of research by deliberately combining methods drawn from different traditions with different underlying assumptions.”  
(Denscombe 2010, 137.)

Mixed methods is a research method that aims to combine and mix the data gathered and analyzed by qualitative and quantitative research methods. This data is then organized into specific research models that offer logic and procedures for the study which it aims to frame within theory and philosophy. (Creswell & Plano Clark 2018, 5.) The MMR enables the researcher to investigate the research questions and problems from a multitude of angles, this can be compared to the mono-methodology studies where the study implements only one methodology. Cohen, Manion, and Morrison (2018) state that the MMR provides a “more comprehensible and complete understanding of phenomena” (Cohen, Manion, & Morrison 2018, 33).

The MMR is most commonly used when the researcher needs to improve accuracy, to have a more complete picture on the research target, to compensate the strengths and the weakness of different research methods or, and to develop the analysis or to provide help with the sampling. (Denscombe 2010, 140-143.) Furthermore, according to Mertens and Hesse-Biber (2013), The MMR combines two methods that strengthen each other and therefore provide a better and fuller result.

Cohen et al. (2018) offer the researcher questions to ponder when deciding whether to use the MMR or not. They advise to question what the profit of using the MMR is, and what the downsides of not using it are, and if there is a possibility that by not using MMR the researcher might diminish the quality of the research. (Cohen, Manion, & Morrison 2018.)

MMR has a multitude of different applications that have developed over the years that it has been used by researchers. These applications are called core designs by Creswell and Plano Clark. (2018) All of these designs can be either fixed or emergent, the fixed designs have predetermined the qualitative and quantitative methods, prior to the begin of the study, and emergent design lets the other research design to emerge from the research questions and research problems. These categories are not fixed, but flowing, and some parts of the study might be predetermined and some of it might emerge while doing the research. (Creswell & Plano Clark 2018, 52.)

Creswell and Plano Clark introduce the typologies of MMR, which are numerous and span from 1989 to this day (see Creswell & Plano Clark 2018, 54 - 57), but they have managed to limit the typologies into three core designs which are: Explanatory sequential design, Exploratory sequential design, and Convergent design. However, this study follows a more advanced design of Experimental MMR design to accomplish the research task which is explained in greater detail in chapter 4.3.

#### 2.2.4 General Paradigm of Research

According to Creswell and Plano Clark (2018, 41-42), there are several stances where the researcher may choose when deciding on the paradigm of the research. They introduce four different stances which are: **The 'Best' perspective, the Dialectic perspective, Context perspective, and Scholar community perspective.** This study will follow the stance of the **Context perspective** and the **'Best' perspective.**

This leans heavily on the pragmatism side of the spectrum of paradigms, Creswell and Plano Clark advice Experimental Mixed Methods design to follow Pragmatism (Creswell & Plano Clark 2018, 39). Furthermore, according to Feilzer (2010) pragmatism offers an option to the paradigm war. Feilzer states that when using a pragmatic paradigm, the researcher chooses to be unconcerned of what method they use if the methods chosen will provide the best results (Feilzer 2010).

The issue of pragmatism is also discussed in an article by Burke and Onwuegbuzie (2004) where they state that: "the bottom line is that research approaches should be mixed in

ways that offer the best opportunities for answering important research questions.” (Burke & Onwuegbuzie 2004, 16.) Furthermore, they say that with pragmatism a researcher can cross the gap left by paradigm war, they conclude that pragmatism is not the resolution to all the philosophical problems, and it should not be (Burke & Onwuegbuzie 2004, 17).

### 3. Theoretical Background

The theoretical background chapter of this research is divided into three sections, where the first discusses the games, roleplaying, and eduLARP teaching method, while the second discusses second language acquisition of English. The final section offers a possible connection between eduLARP and second language acquisition.

#### 3.1 The Game, Roleplaying, and EduLARP

In this section, roleplaying is explored from a scientific viewpoint and the theories behind eduLARP are explained. The chapter begins by illuminating the theory of games. It continues with looking at what RPG and LARP are as concepts and what they involve. Thirdly, the study explores the idea of recentering and learning by avatars. Fourthly, the study answers the question of what are the pedagogical theories that act as a base for eduLARP. The section concludes with exploring gamification in the classrooms.

#### **What is a game**

One of the earliest definitions of a game can be found from the book *Man, Play and Games* by Calliois in 1961 (reprinted and edited in 2001). The definition consists of six assumptions that must be true for a game to be a game.

According to Calliois (2001, 9-10), a game must be:

1. **Free:** The participants in the game must take part in their free will. The participants cannot be forced to partake in the game if they do not want to.



2. **Separate:** The game is tied to a certain place and time usually decided prior to the game.
3. **Uncertain:** The game must be somewhat uncertain. The course and outcome of the game cannot be determined beforehand. A game must also allow some degree of innovation for the players.
4. **Unprofitable:** The game cannot generate any wealth or product that is not in the game itself. The situation after the game must be identical to what it was before the game started if the wealth of the players is concerned.
5. **Governed by rules:** In a game, there is a new set of rules that can contradict the rules set by society. However, these rules must be abandoned after the game is over.
6. **Make-believe:** In a game, there is always a hint of make-belief involved. The make-belief works in unison with the rules set by the game. Usually, players are aware of this and accept it freely.

Using these assumptions, a game is easy to identify. For example, if all the players have arrived at the previously agreed location of their free will, the first two assumptions are true. During the game, if the make-belief and uncertainty of the game can be suspended, the next two assumptions of Uncertainty and Unprofitability are true. Then, if the players are committed to the rules, the fifth assumption is also true. Probably the trickiest of these assumptions to fill is the assumption of unprofitability. It becomes important only when the game is ending, and the wealth gathered in the game is re-distributed to the players or forgotten. This becomes an issue when dealing with gambling and other games of chance, where a simple card game transforms into something else, and at the end of the game the players do not return to the identical pre-game status.

Roleplaying is easily fitted into this definition that Calliois presents, however, Syväluoma and Turpeinen (2003, 4), and Leppälahti (2009, 7) state that a roleplaying game does not offer a clear winner or loser and therefore does not fill the assumption 4 completely. In an RPG, all participants (players) come together of their free will, and uncertainty and make-belief are easily suspended over the players. The game follows very detailed rules, which

are usually forced very strictly over the players. Also, the game will not profit anyone and the situation prior to the game is almost identical to the situation after the game.

### **Roleplaying Game**

Roleplaying is usually divided into two main categories Live-Action Roleplaying (LARP for short) and Roleplaying game (RPG for short) (Pettersson 2005, 22; Leppälähti 2009). However, Daniau (2016) includes “Play-by-post Role-Playing Games (PRPG)” and “Video Role-Playing Games (VRPG)” under the umbrella term of roleplaying (Daniau 2016, 426). However, this study focuses mainly on RPG and LARPing, and therefore will leave the PRPG and VRPG out of the theoretical background.

The term roleplaying game is moderately evasive since different rulesets and traditions offer a different definition to RPG and what it is trying to accomplish. Pettersson (2005) defines RPG as an improvisational radio theater, and Daniau (2016, 424) defines RPG as “the progressive creation in a small group of players of a type of collaborative narrative animated by a game master, in which each player takes on the main role.”

Furthermore, this study borrows its definition from several roleplaying books, such as Stars without number (Crawford 2011), Dungeons & Dragons fifth edition: Dungeon masters guide (Gray, Carter, Sims, Wilkes 2014) and Warhammer 40000: Dark Heresy (Barnes, Flack, Mason 2008).

The earliest form of LARP/RPG is the ‘let’s pretend’ game that comes easy to all children. However, the first LARPS that are documented can be traced back to ancient Rome, where the theatre was also a game (Ludi), also, gladiatorial games and naval combat performances involved some roleplaying elements. More recent cousins of LARP are interactive murder mysteries, and the kriegspiel developed by Prussians to train soldiers. In the 20<sup>th</sup> century, the more peaceful LARPs such as mock legislatures and Model United Nations have gained popularity. (Morton 2007.)

There are a few similarities to all RPG’s (see Barnes, Flack, Mason 2008; Crawford 2011; Gray, Carter, Sims, Wilkes 2014). The game is played by several players, where one player takes the role of a gamemaster or the storyteller that controls the game and the story, acts as non-player characters, plans the game session, and tries to control the chaos that is an RPG session. (Gray, Carter, Sims & Wilkes 2014, 4; Pettersson 2005, 27.) All the while, other

players take the role of player characters that act like protagonists of a movie or a novel. Dungeons & Dragons fifth edition defines the player characters role as follows:

“Each player creates an adventurer (also called a character) and teams up with other adventurers (played by friends). Working together, the group might explore a dark dungeon, a ruined city...” (Carter, Sims, Gray & Perkins 2014, p. 5.)

The main difference between a LARP and an RPG is that the player fully immerses into their character, much like in improvisational theater (Harviainen & Savonsaari 2013). The players dress up, speak, move, and act like their character.

It is possible to create a roleplaying game from almost anything in a human spectrum of experience if all the players agree and commit to the imaginary narrative of the game. Heliö (2004) argues that there is a non-spoken agreement between the players of the game. This agreement affects the ordinary things present in the game and could change the meanings of these things to something completely different in the narrative of the game. For example, a plane flying over the player might transform into a dragon in the narrative of the game if the narrative has been accepted by all players. (Heliö 2004, 70.)

### **Educational LARP**

There is a clear division between educational roleplaying and entertainment roleplaying. This is mainly due to the goals of the game. While entertainment roleplaying aims only to entertain, an educational roleplaying game aims towards learning. To achieve the best learning results educational roleplay usually mixes the two roleplay genres, LARPing and RPG. (Henriksen 2004, 108.)

Bowman and Standiford (2015, 1) define eduLARP as follows:

“As a pedagogical outgrowth, edu-LARP refers to an educational roleplaying exercise in which participants adopt a new role for a long period of time in a bounded, fictional scenario that may or may not resemble mundane reality. Some edu-LARP scenarios contain rules or win conditions, but not all.”

Educational roleplaying has a connection towards educational drama designs, but it differs in how the role that is played is seen. According to Harviainen and Savonsaari (2013), the

character roles are individuals and persons, not just a dramaturgical role for the actor to interpret and act out according to a script. In an RPG the player has a possibility to create their character and choose the qualities that they want to have. Harviainen and Savonsaari state that the players have the freedom to ignore some of these qualities that make the character. This leaves more room for interpretation that can lead to learning if guided correctly. (Harviainen & Savonsaari 2013, 136.)

Harder (2007) states that educational LARP (from now on referred to as eduLARP) teaching is not limited to any one subject. They continue, that eduLARP helps especially when trying to teach the students what kind of citizen and person they want to become in life. Roleplaying can help the students to see the world from a different perspective than their own. When a student plays a character, they must practice how to entertain a thought without accepting it since the opinions of the character and the player might differ. This also aids the teaching by letting the students experience and explore a variety of different ideas. (Harder 2007.)

According to Harder, eduLARP offers a new context to the teaching and a new environment for the students to practice their abilities. They elaborate on the eduRPG's ability to connect practical knowledge to the theoretical and make the subject at hand feel more concrete to the students. (Harder 2007.)

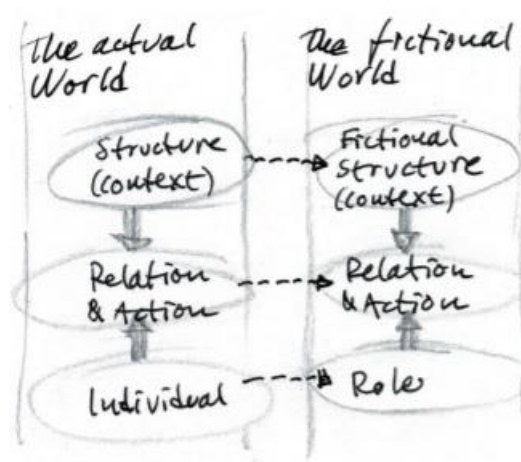
The aspect of combat and roleplaying might be off-putting for some students, however, Thestrup (2007) says that in eduLARP there is a space for everyone and every type of child in the game. According to Thestrup:

“Larp, by the way, includes many other facets that both genders can play with. It is necessary to sneak, run, negotiate, play roles, improvise, innovate and be attentive. You need to practice immersion, participate in a story or play a game with rules. -- There's sufficient space for girls and boys together to find a culture where gender is not viewed as a constraint.” (Thestrup 2007, 225.)

Harviainen and Savonsaari (2013) state that the most important aspect of an eduLARP is the physicality, which means that even when the players are immersed deep into their characters, their thoughts and actions are still their own (personal). For example, if the character learns a way to solve a puzzle the player will also learn the same thing.

The same thing happens when playing a virtual reality game with an avatar. According to Fox, Bailenson, and Binney (2009) when a person sees an avatar of themselves doing something, they are more likely to reproduce the same action again later, for example exercising. They also found that using a model of oneself is more beneficial to the result of learning (Fox & Bailenson 2009; Fox, Bailenson, & Binney 2009).

Physicality leads the player into a dilemma, where as an action is done to the character (e.g. another character hits them with a sword), the same action is also done to the player. This amplifies the actions done and experiences felt to the more embodied level of play. This opens the issue that the player might feel inclined to look for differences between the game world and the actual reality that they are in. (Harvianen 2016.)



**Figure 2. The Recentering in Larping (Henriksen 2004, 127)**

Recentering means the ability to imagine other possible worlds that are different from our own and the ability to relate to the fictional characters. According to Ryan (1991), when a person is exposed to a fictional world or character, that person is recentered into the reality of that world and accepts that reality to become as 'true' as the actual reality that the person resides in (Ryan 1991). Thus, recentering requires that the person involved with the fiction allows it to be an alternative reality. When recentering is successful the reality of the fiction becomes the actual reality of the fictional world (Ryan 1991, 22).

The same kind of thing happens when LARPing or playing an RPG. According to Henriksen (2003), roleplaying offers this recentering in three layers: structure & context, relation & action, and individual layers. In a roleplaying game the player functions in three different levels of reality (see Figure 2). According to Henriksen (2004), the first level focuses on the

recentering of reality to fit the desired fictional world. This includes modifying the history, social structure or even laws of nature itself. Henriksen states that other participants of the game are also important for this since they are part of the world as well and allow interaction between players.

The second level focuses on the recentering of the reality of the player. The player stops being their own person and becomes the character they are playing. For example, in a classroom situation, the students will stop being fourth-grade students and they become pirates in command of a vessel. This recentering includes the different “influential personal actants” (Henriksen 2004, 127) that change the players' actants to the characters' actants. The third level attempts to portray the actions and the relationships in the roleplay itself. The relation recentering facilitates conflict and action between the players and the fictional world they are now present in. This recentering enables the alteration in local actants. (Henriksen 2004, 127-128.)

When combining the ideas of Fox et al. (2009), Harvianen and Savonsaari (2013), Henriksen (2004), and Ryan (1991), the concept of recentering becomes easily tied to physicality and avatar learning. Since test subjects felt a connection to their model (Fox et al. 2009a+b), they were recentered towards the new fictional world of the digital simulation (Ryan 1991). The same thing happens when a person is playing an RPG (Henriksen 2004). Furthermore, because the avatar that the player recenters around is a physical in LARP (Harvianen & Savonsaari 2013), the player might feel a greater connection to it compared to a computer simulation. Therefore, it would seem that if the recentering of reality is successful, learning through an avatar or character could equal or exceed the learning results gained by learning as oneself.

### **Pedagogical Theories Behind EduLARP**

In an article by S. L. Bowman and A. Standiford (2015), it is stated that the eduLARP teaching method follows the theoretical principles of experimental learning by Kolb (1984) and situated learning by Lave and Wenger (1991). The theoretical principles of experimental learning are presented by Kolb in *Experimental Learning* from 1984. Kolb defines learning as “the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience.” (Kolb

1984, p. 41.) An article by Kolb and Kolb (2005) pictures this creation as a circle that includes Concrete Experience, Abstract Conceptualization, Reflective Observation, and Active Experimentation. According to them, in an idealized model of learning, the learner would go through the four stages of experiencing, reflecting, observing and acting. (Kolb & Kolb 2005, p. 194.)

The situated learning theory from 1991 states that learning does not happen in a vacuum, but everything around the learner affects the possible outcome of the learning (Lave & Wenger 1991). The theory states that “learning is an integral and inseparable aspect of social practice” (Lave & Wenger 1991, p. 30). Situated learning also introduces the concept of Legate Peripheral Participation, which says that the learner should become more and more involved with the task at hand as his or her expertise and abilities grow (Lave & Wenger 1991, 34- 37).

The eduLARP method teaches the student with a multitude of learning dimensions. Bowman (2014) has gathered the learning dimensions and the development of the student in a helpful table shown below (Table 2). The table 2 shows the learning dimensions on the left column and how the students benefit from eduLARP on the right.

**Table 2. Cognitive, Affective, and Behavioral Dimensions of Student Learning Through Edu-LARP (Bowman 2014, 115)**

<b>Learning Dimension</b>	<b>Student Development</b>
<b><i>Cognitive</i></b>	<ul style="list-style-type: none"> <li>• Critical ethical reasoning</li> <li>• Exercising creativity, spontaneity, and imagination</li> <li>• Intrinsic motivation*</li> <li>• Improved problem-solving skills</li> <li>• Learning multiple skills and knowledges simultaneously</li> <li>• Self-efficacy, perceived competence</li> </ul>
<b><i>Affective</i></b>	<ul style="list-style-type: none"> <li>• Active engagement*</li> <li>• Enhanced awareness of other perspectives</li> <li>• First-person identification improving emotional investment</li> <li>• Increased empathy</li> <li>• Increased self-awareness</li> <li>• Intrinsic motivation*</li> <li>• Raising social consciousness</li> <li>• Social skills development, e.g. cooperation, debate, negotiation</li> </ul>
<b><i>Behavioral</i></b>	<ul style="list-style-type: none"> <li>• Active engagement*</li> <li>• Exercising leadership skills</li> <li>• Intrinsic motivation*</li> <li>• Improving team work</li> </ul>

\* Literature suggests that intrinsic motivation and active engagement have cognitive, affective, and behavioral dimensions, hence organizing them in multiple categories.<sup>30</sup>

Other researchers have investigated the different aspects of educational roleplaying and have found that there is a multitude of possibilities for implementation. For example, Syväluoma and Turpeinen (2003) in their thesis state that educational roleplaying could bring more excitement and adventure into the classroom situation. They implemented a roleplaying method into a student counseling class, where the students created a character based on their future selves. According to Syväluoma and Turpeinen, the teenaged students can learn about their future professions as well as the mentality that is associated with the said profession. (Syväluoma & Turpeinen 2003.)

### **Gamification**

Gamification is a term used when talking about teaching methods or styles that implement game-like elements and designs. This study uses the definition created by Deterding, Dixon, Khaled and Nacke (2011a, 5; 2011b, 3) and Hamari (2015).

According to Deterding et al. (2011a, 5) and Hamari (2015), when utilizing gamification in the classroom the teacher or educator implements game-like designs rather than game-based technology into teaching. These game designs should provide the player with a rewarding activity that aims to be beneficial. This means that some elements of game design are brought into a non-game situation. The intention behind gamification is not to create a completely new game from scratch but to utilize advantageous game design choices to create a new and motivating game-like environment where the activities are rewarding and target a clear beneficial goal, such as learning. (Deterding et.al. 2011a, 5; Hamari 2015.)

To clarify the definition further, a game is a self-sustained system that feeds on the motivation and commitment of the player. The more you play the more committed you are towards the game. Learning is achieved by implementing a learning goal that the player strives towards. Furthermore, it seems that game-based learning provides optimal results when the players view the game as a learning method, and when the game has been designed with a pedagogical standing point (Whitton 2009). However, producing a feeling of a game is not as simple as it first seems. Huotari and Hamari (2012) say that gamification is not always successful, and it seems to rest solely on the shoulders of the player.



According to them, the gamification of an action or service is an individual experience and unique to every game. (Huotari & Hamari 2012.)

An example of gamification can be found from an article by Pihkala-Posti (2015), where they introduce a Berlin Kompass application that helps the students to practice giving directions in German. They say that the application prevents the students from correcting each other and when given the wrong answer the application gives a special task to the players to complete before allowing them to continue the game. (Pihkala-Posti 2015.)

However, gamification is not without issues, as Whitton (2009) stated that gamification works only if the game is accepted as a source of learning and not a competition. Harviainen, Lainema, and Saarinen (2014) found out that players who participate in an educational game will forfeit the learning task in order to win the game. This is due to the fact that educational games tend to forgo the chance element that is present in many non-educational games. (Harviainen, Lainema & Saarinen 2014.)

### 3.2 Second Language Acquisition

Second language (L2 for short) refers to any language a learner is taking up after their native language. Second Language Acquisition (SLA for short) refers to the act of learning a second language after the first one, e.g. learning English after the native language. In this study concepts of SLA and L2 learning are used almost interchangeably, since both refer to learning another language after the native one.

According to Ellis (2003), there is a structure on how the learners learn L2. The earlier in life a structure of L2 has been learned, the more likely it is that the learner can use it. Also, Ellis's (1997) past research shows that the errors the L2 learners are making state the level of competence the learner is at in SLA at, and that the seriousness of the errors is measured by how much it affects the understandability of the sentence.

The efficiency of SLA depends on several factors that Ellis (1997) divides into external internal factors. External factors refer to the situational conditions the learner is subjected to, for example the teacher's attitude towards the learner, the learner's attitude towards the target language, and the input that the learner is subjected to. Internal factors, on the

other hand, refer to general information that learner has considering languages such as the command of their mother tongue and how they use it to help the learning, general knowledge of the world that helps understanding the input, communication strategies, and the learner's inherent talent with language learning. (Ellis 1997.)

The CEFR (Common European Framework of Reference for Languages: learning, teaching, assessment) direct the language teaching towards the learner's ability to participate in real-life situations and to communicate effectively with people to achieve a goal. CEFR suggested that the assessment of learner's language competence is measured with two parallel categories: communicative ability and language ability. This measurement is also known as an action-oriented approach. (CEFR 2018).

The impact that CEFR has had on the European and Global language pedagogy cannot be denied. After the CEFR was first published in 2001 there has been an excess of discussion, and the governments of Europe were quick to apply the new framework (Figueras 2012). Nonetheless, CEFR has gained some critique, especially for not measuring the L2 students fairly. According to Hulstijn (2007), the CEFR (2001 version) sums up the quality and quantity of language use. This means that the CEFR (2001 version) does not differentiate between how wide the learner's language use is (learner can use a multitude of speech tasks) and how precise the language use is (learner can use very sophisticated language). Hulstijn (2007) says that these two should be separated and the learners should be assessed in both.

CEFR advises educators to view the learners as social agents and language users rather than just language learners. CEFR encourages educators to involve the learner in the learning process using communication. This also implies that the educators must recognize the language learning and teaching to be a social endeavor that involves both participants of the process, learner and educator. CEFR states that the focus of language learning should be on **how to use the language** in contrast to **learning about the language**. (CEFR 2018.) This is very close to a Communicative Language Teaching (CLT for short)

CLT is a language teaching method that highly emphasizes the communication between individuals in the target language. It involves and combines task-based learning with some ideas borrowed from sociolinguistics. CLT has gained popularity in second language teaching in recent years. This increase is due to the world language status of English which

has brought forward the task-based learning and Communicative Language Teaching (Kramsch 2017). According to Kramsch (2017), CLT is based on social scientifically applied linguistic research, meaning applied linguistic theory. From the applied linguistics theory Kramsch points out six assumptions that the CLT is based on:

1. Language is primarily for communication between persons and for information exchange. People learn languages to be able to communicate coherently with each other.
2. CLT emphasizes spoken communication and language by focusing on lexical knowledge, lexicalized grammar and communication skills with the target language.
3. Learning is a part of communication. To facilitate learning the learner should be able to partake in an authentic context.
4. Task-based pedagogy can be used to construct and reconstruct the information in a cognitive process.
5. CLT should only teach good learning and communication techniques.
6. CLT is focused on collaborating teaching styles and tasks where students should work on groups to solve real work tasks to facilitate language learning. (Kramsch 2017, 7-8.)

CLT can also be divided into two versions, strong CLT and weak CLT. In the strong version of CLT, the language users must use language, any form of it, to be able to learn the target language. Howatt (1985) says that in strong CLT a language is acquired by communication between persons and no other means are necessary. The weak version of CLT, on the other hand, gives a more open idea of CLT. Communication is used as a tool to achieve a product, in this case language competence. Weak CLT is used to tie up the communicative aspect of language learning into a wider curriculum. However, with this approach there is a possibility that the learning activities that are communicative might become distant and feel like a separate part of teaching for the pupil. (Howatt 1985.)

Kramsch (2017) also mentions **task-based pedagogy** in the article and it is closely tied to CLT. Ellis (2003) offers six criteria that define what a task, in task-based pedagogy, is:

1. a task must have a plan to follow. The plan directs the activity of the learner towards something. However, the plan itself might not follow this plan as intended and may not result in communicative behavior.
2. a task has to have a gap that can be crossed only by using language. The first criteria do not specify what language or communication tool the participants should use.
3. a task aims at performing a real-world situation where the language is used.
4. a task may involve a mixture of language skills such as listening and producing orally, communicating with someone, or writing. All these are set before the task during the creation of the plan.
5. a task must engage the learner's cognitive processes.
6. the task must have a clear goal that the learner strides towards. Such as conveying a message to another. (Ellis 2003.)

CLT and task-based pedagogy have a clear link to sociolinguistic theories of language acquisition. Verhoeven (1997) presents three aspects that language education has borrowed from sociolinguistics. The first aspect states that language learning is not only a development of the grammatical system but also the increased ability to use language as a tool in language learning and appropriately in the context of surroundings. The second aspect is the vitality of social communication in language learning, and the environment. The third aspect concerns the learner's ability to interact with a wide diversity of people from different backgrounds. (Verhoeven 1997, 403.)

It seems that CEFR follows a sociolinguistic approach towards language acquisition and suggests a task-based communicative pedagogy for all languages as a framework. This is most likely due to an increasing need for communication in our society and in the world in general.

### 3.3 What EduLARP Can Offer to L2 Acquisition and Teaching

From a completely theoretical standing point, it seems that eduLARP has a lot in common with modern theories of L2 acquisition and language teaching in general. When compared to the framework that CEFR (2018) offers, it seems that eduLARP shares multiple theories

that the CEFR framework and the theories it is based on emphasize. This is due to the communicative nature of LARPing and roleplaying in general. An RPG is not possible if the players and the gamemaster do not work together to create a game, and this is not possible without communication between players, gamemaster and non-player characters. (Pettersson 2005; Daniau 2016, 424). Furthermore, according to Thestrup (2007), a vital part of LARP is negotiation and groupwork.

For example, the characters and the game itself can be anything a player can imagine (Heliö 2004). This offers excessive possibilities for communication situations between the characters and the game world. When this notion of communication possibilities is compared to the principles of CLT (see, Krash 1982, 7-8) a clear connection can be seen. Therefore, CLT and task-based pedagogy both closely relate to eduLARP, since it (EduLARP) can offer a safe environment to practice real-life communication and language skills through an immersive simulation where the participants are engaged in the task. For example, Ivanove and Troflimove (2016) tell that eduLARP is used in language teaching in higher education in Russia. For example, in business schools the students practice going through customs and holding business meetings.

Additionally, in eduLARP teaching the student is centered on a situation where they can freely experiment with different communication styles and methods. For example, the character could have a speech impediment that the player must work around. In addition, a LARP setting might have an extremely diverse cast of characters that all must find a way to communicate, and as stated CLT and task-based pedagogy task or teaching situation should accommodate diversity, and the learner should know how to communicate with people of different background.

The task-based pedagogy that Ellis (2001) introduces focuses on meaningful tasks that the learners must complete. These tasks are made important by giving the player characters a drive or a reason to conclude them. The tasks of an RPG or LARP are also usually tied to a story that acts as the plan of the task.

In conclusion, it seems that eduLARP would be an excellent teaching method to answer the goals that CEFR has planned for language learning. This is due to the communicative and task-based nature that RPGs and LARPing have.

## 4. Research and Analysis Methods

This chapter of the thesis discusses the research- and analysis design choices made for this study and explains how the different data is collected and analyzed. The chapter is constructed to focus on one aspect of the mixed methods research (mixed methods, quantitative, qualitative, and) at a time. The first section (4.1) explains how the mixed methods data was collected and analyzed. The second section (4.2) focuses on quantitative data collection and analyzing. Lastly, the third section (4.3) aims to explain how qualitative data was collected and analyzed. The sections are divided into more compact portions, where the main topics are discussed in more detail.

This section of the study is constructed as presented above, for the ease of reading and to highlight **how and why** certain research design choices were made. Even though the study is ultimately concluded by using mixed methods, the presentation of qualitative and quantitative methods and the reasoning behind those methods are of vital importance.

### 4.1 Experimental Mixed Methods / Connection Between the Teaching Method and Language Competence

The mixed methods research is used to see how the personal experiences of the pupils explain the results of English language competency tests, and whether there is a connection between the two. This research method is chosen because of the ambitious goal of this study, and this is deemed to be the best way of establishing a base for future research.

Experimental mixed methods design (or EMMD for short) is an advanced MMR design that merges a collection of different qualitative methods with a quantitative experimental

method. This study follows the mixed methods experimental design, where a qualitative method is added to the basic experimental design. According to Creswell and Plano Clark (2018) the objective of this combination is to enhance the quantitative data with a more personal and deeper touch; to “provide personal, contextual, qualitative experiences drawn from the setting or culture of the participants along with the quantitative outcome measures.” (Creswell & Plano Clark 2018, 108). This means that the data gathered from an interview, for example, is embedded in the test results to gain information on how the subjects have learned.

In EMMD, the data collection is dominated by the decision of when the qualitative data is gathered to maximize the advantage that it provides to the research. Creswell and Plano Clark introduce three possible positions where the qualitative data can be gathered; before, during, or after the quantitative part is collected. Also, the amount of collected data plays an important role. The data can be collected in all the points of the study (as stated above) or just in a singular point. (Creswell & Plano Clark 2018, 197.) They give a variety of reasons to attach the qualitative data to the experimental study. Reasons vary according to the position when the qualitative data was gathered, and the researcher must decide what is a suitable position to gather the data according to their research goals.

The before-position offers the researcher an opening to modify the intervention so that it will fit the subjects or maybe modify the meter on which the subjects are measured. The second position, during, presents the researcher with a possibility to ask questions on how the subjects experience the intervention. The final position takes place after the intervention where the researcher may enquire about the intervention and explore in more detail how the results from the quantitative part came to be. (Creswell & Plano Clark 2018, 108, 198). In this study the gathering of qualitative data happens after the intervention, since the research aims to understand the development of English language competence the after-position is perfect.

Creswell and Plano Clark (2018) state that EMMD follows the post-positivist orientation where the qualitative part is in an insubordinate position when compared to quantitative, especially when the qualitative test is held in during the intervention. However, they also state that there is a clear difference in the philosophy when the qualitative part is prior to or after the intervention.

The implementation of Experimental Mixed Methods Design also presents a challenge of how to diminish the issues caused when the qualitative data is implemented into the experiment. For example, the collected qualitative data might affect the answers in the quantitative part of the test so that the credibility of the research might be in question. To counter this bias in the research, Creswell and Plano Clark suggest that the qualitative data is gathered after the experiment or distribute the qualitative data collection evenly over the control group and the research group. (Creswell & Plano Clark 2018, 108; 199).

Since this study does not contain a control group, the latter possibility of controlling the bias presented is removed and therefore, the first control method is used in this study. Thus, the qualitative data is gathered through an interview after the quantitative data is collected.

### **Experimental Mixed Methods Design**

The EMMD is divided into three clear steps by Creswell and Plano Clark. The researcher must:

1. determine how the qualitative data will be used in the experiment,
2. conduct the quantitative experiment,
3. determine how the qualitative findings enhance the experiment.

(Creswell & Plano Clark 2018, 111.)

For the mixed methods to work properly the two different data sets must be merged into one. Creswell and Plano Clark advice to search for the intersection of the two data sets where they enhance the experiment. This is found when the data sets merge or phase into each other.

### **Analysis**

The Mixed methods data is analyzed by using a joint display model, as stated in Creswell and Plano Clark (2018) The joint display provides the researcher with a possibility to compare the two different data sets parallel to each other and then interpret them together to create a new data set than can be analyzed again to draw conclusions. In this study, the analysis is done by adding the personal experiences of the pupils to the quantitative data received from English language competency tests.



## 4.2 Quantitative / English Language Competence

This study utilizes quantitative methods to measure the pupils' competence and development of the English language using standardized tests provided by the European Consortium for the Certificate of Attainment in Modern Languages (ECL for short). These tests were chosen to ensure that the tests are eligible for the principles of language testing. Tests act as the meter to determine how well the pupils can write English and understand written and spoken language.

The tests will measure the students' competence in three aspects of the English language on B1 level:

- (i) written communication
- (ii) reading comprehension
- (iii) listening comprehension

(ECL 2018b.)

B1 level of CEFR's system was chosen since it is clearly the closest one to the national, both Danish and Finnish, curriculum level where the students should be after the seventh grade.

A pre-test will determine the starting competence of the students at the beginning of the school year before the students have been subjected to the eduLARP method teaching. The tests are coded with the student's name so that the development between the pre-test and the post-test can be assessed for individual students as well as the whole group. A post-test is conducted approximately seven months from the pre-test, and it uses a similar test where the only difference is the questions.

The ECL is an association of different institutions that represent different European languages. It aims to promote the teaching of languages in Europe, provide valid standards for language testing and provide certificates and exams to help individuals in the EU. (ECL 2018a.) The tests are also checked using ECL standards and guidelines. (ECL 2018c.)

The tests are graded by the researcher as follows: The full marks combined from the three tests are 75 and the maximum marks from an individual test are 25. The tests are divided into two parts with ten questions each, therefore, maximum marks from each part of the

test were 12.5 and that divided by ten equals 1.25 marks. In the listening and reading comprehension tests, a wrong answer costs 1.25 marks for the student. In the essay test, a minor spelling error or minor grammatical error costs 0.625 marks, a major grammatical or spelling error takes 1.25 marks, crossing or not filling the word count of the essays (125 words) costs 1.25 points, and not answering the given questions (the test paper provides four questions that the students need to answer per essay) results in losing 1.25 points per question. If the text is severely short (under 50 words) the researcher would fail the task completely.

Having evaluated and calculated the marks, the researcher inserts the results into the SPSS program for analyzing. The coding of the test results is as follows:

- Overall grade/competence in the test: This data shows how the student did in the tests in total.
- Aspect grade/competence: This data shows how the student did in a single aspect.

### **Analysis**

The results of the tests are typed into the SPSS program for analysis. The results are coded by the name of the student so that individual development can be measured and compared. However, before the analysis all participants are assigned a random number by a random number generator in Windows Excel to protect their anonymity. These numbers are then assigned to students in a random order. The participants' scores are recorded into different columns that are tagged by the name of the test and when the test is taken. The total scores of individual participants are calculated by adding together their scores from all three tests and are given their own column.

In an experiment, the stimulus, or variable, is controlled by the researcher who then attempts to find a connection to another variable. Variables are divided into independent variables and dependent variables. The main distinction between these is that the independent variable is the event or stimulus under inquiry. (Kirk 2013, 3.) The independent variable in this study is the change in the students' competence of English language, and the dependent variables are the results they get from the tests.

The variables are named as below:

Reading_Pre,	Reading_Post,
Listening_Pre,	Listening_Post,
Writing_Pre,	Writing_Post,
Total_Pre.	Total_Post.

The ending ‘\_pre’ or ‘\_post’ is added to the variables to define when the variable was measured. Ending ‘\_pre’ means that the variable is measured before the intervention and ending ‘\_post’ means that the variable is measured after the intervention. The data points are added in by hand and double-checked by the researcher after the grading process as well as at the beginning of the analysis process.

The analysis process begins by acquiring the descriptive statistics for all variables after this the variables are tested for normality of distribution with Kolmogorov-Smirnov and Shapiro-Wilk tests. After this, the data is tested for difference between pre-tests and post-tests with paired sample T-test or non-parametric variants for paired samples T-tests.

#### 4.3 Qualitative / How and Where Is the English Language Learned

The qualitative data is collected via a group interview. The aim of the interview is to provide the research with a better understanding of how and where the pupils feel that they learn the English language. The interview tries to identify how different sources of learning help the students learn English, which sources of learning help with which aspect of the English language, how the students feel that the teaching method helps them learn English, and where the students learn English. This interview should give a clear indication of how and which aspects of the English language the students learn. Interview questions can be found in Appendix 1.

Five to eight students are selected from the group that shows the most progress between Test 1 and Test 2. The English teachers are asked which students would be best suited to the group interview. A student group of five to eight is deemed enough to give a proper resemblance to the research group of the study (N=63).

**Table 3. Types of Group Interviews and Dimensions (Fontana 2007, 33)**

Type	Setting	Role of interviewer	Question format	Purpose
Focus group	Formal, preset	Directive	Structured	Exploratory, pretest
Brainstorming	Formal or informal	Nondirective	Unstructured	Exploratory
Nominal / Delphi	Formal	Directive	Structured	Exploratory, pretest
Field, natural	Informal, spontaneous	Moderately nondirective	Very unstructured	Exploratory Phenomenological
Field, formal	Preset In field	Somewhat directive	Semistructured	Phenomenological

SOURCE: Frey and Fontana (1991: 184).

When consulting the Table 3 above this study follows the group interview type that Frey and Fontana (2007) list as “Field, formal” with the role of the interviewer as “somewhat directive”, question format as “Semistructured”, and the purpose of the interview to be “Phenomenological”. These directives are in unison with the methods of responsive interview and group interview. (Fontana 2007, 33.)

A less formal interview provides the researcher with the ability to follow possible threads of information that would otherwise not be found. Furthermore, the atmosphere in the less formal interview offers a more personal connection between the interviewer and the interviewees. (Fontana 2007, 36-37; Curtis 2013, 115.) According to Rubin and Rubin (2005), a qualitative interview is useful when the aim is to reconstruct or understand events and their outcomes (Rubin & Rubin 2005, 3). Sarajärvi and Tuomi (2004) state that open-ended questions and the researcher’s ability to ask some follow-up questions and to direct the interview towards more interesting topics that are more important to the research, will offer more to the research (Sarajärvi & Tuomi 2004, 78). In this study, the aim is to understand the results of the quantitative tests and provide the understanding required to state if there is a connection between the teaching method and the results.

When holding an interview, power dynamics always play a key role. According to Curtis (2013), power dynamics between the interviewer and the interviewees are present in an interview. Curtis states that these power dynamics may affect the outcome of the interview, as well as the questions the interviewer feels comfortable asking. (Curtis 2013, 116.)

Furthermore, Wang and Yan (2012) state that the interviewer has greater power status when holding the interview. They emphasize this standing point in their article as follows:

“An interviewer has the right to ask questions to initiate an interview and has the privilege of terminating it. Through the interviewer’s choice of questions, she or he selects the topic of the interview; the interviewer even has the prerogative to ask questions so deliberately designed that no new information is introduced (as in closed-ended surveys).” (Wang & Yan 2012, 234.)

They also state that the interviewee only acts as a source of information to the research that is used to gain results. According to them, the interviewee is only able to ask questions if the researcher permits it. The interviewee is not permitted to change the topic that is discussed, as the researcher will decide it. (Wang & Yan 2012, 234 - 235)

Taking this into consideration, the power dynamic in this study is clear: the interviewer is older and a semi-professional in his field compared to the students he is interviewing. However, the researcher and the participants share a common interest and the participants are familiar with the researcher beforehand. The familiarity between the researcher and the participants might affect the results of the interview in some manner. According to Curtis et.al. (2013) when planning the interview, the researcher should consider several questions concerning the power dynamic of the research.

These power dynamics are shaped by how the researcher chooses to deal with them, which then shapes the interviewees and how they work. This study follows the idea of “the interviewees as conversational partners” (Rubin & Rubin 2005, 14), where the interviewees, or participants, are an active operator in the interview and they can move the conversation towards a common understanding. According to Curtis et.al. (2013), this power dynamic is considered non-hierarchical, they state that:

“The use of interviews in a non-hierarchical way does enable the development of a more democratized research process and facilitate the formation of more reciprocal relationships between the researcher and the participants.” (Curtis et.al. 2013, 117.)

The interview in this study follows loosely the method of responsive interviewing where the researcher and the participants develop a relationship during the interview, and the goal is to develop a deep understanding of the discussed topic. The interview might take a surprising but relevant turn therefore, the design must be flexible to dive deep enough. (Rubin & Rubin 2005, 30 – 35.)

However, because of the nature of this study, fully following the ideals of a responsive interview is not possible and they need to be modified. Rubin and Rubin (2005) say that the stylistic variation in the interview does not produce an issue towards the end results. They emphasize the interaction and flexibility between the researcher and the participant. Furthermore, they state that the researcher is not only there to objectively ask questions and record answers. Contrastingly, they state that the researcher's own emotions and personality should be shown since empathy is a valuable tool to get people to answer questions. (Rubin & Rubin 2005, 32.)

Frey and Fontana (1991) list the advantages of group interviews. According to them, the group interview can provide the researcher with additional data by analyzing the participants' interactions amongst themselves. This interaction might reveal some additional phenomenological dimensions of the research topic that might have been lost in a personal interview. The group interview has some utilitarian advantages over the personal interview as well. By interviewing multiple participants at the same time, the researcher saves valuable time and money during the research process. (Frey & Fontana 1991.)

Considering the familiarity that the students have towards each other, the group provides the participants with a sense of security from being together rather than being alone with the researcher. The group interview will also provide an advantage in what the students can remember, as they are able to remind and encourage each other, and provide additional information concerning the topic. (Fontana 2007, 34; Frey & Fontana 1991; Dawn & Deirdre 2012, 408.)

The interview questions are designed with the second research question in mind and the questions attempt to illuminate from where the students have learned English as well as the motivational aspects of the teaching method. The interview questions are assigned to

four different sets. The first question set covers the overall attitude towards English language and the general competence in English. The second set assesses the impact of informal learning, while the third aims to identify the impact of formal learning, and the final set is to see the impact of eduLARP learning. The aspects of learning were chosen to represent platforms where the students might learn English.

Organization for Economic Co-operation and Development (OECD for short) defines four aspects of learning as follows. First of these aspects is Formal learning that has an explicit goal and participants partake in the teaching with a clear goal of acquiring skills or competences. The second aspect is non-formal learning, where the learners gain the knowledge or skills by taking part in an activity that has no specific learning object, however, the learners are aware that they are learning. The third aspect is semi-formal learning where the individuals learn about topic A but also simultaneously learn about topic B. The final aspect is informal learning, where the learning is done without any knowledge of learning goals or any knowledge that the individual is learning. (OECD 2007.)

Non-formal learning and informal learning aspects are combined in this study under the term of informal learning, while aspects of formal learning and semi-formal learning are both represented by the term formal learning. This is done in order to keep the number of aspects needed to discuss during the interview within a controllable limit, as well as to limit the length of the interview and thus aid the analysis.

There are two challenges when conducting the qualitative interview in this research, Firstly, the language barrier and secondly, unfamiliarity with the interviewer and the situation. However, there is a simple and effective solution to these challenges. The issue of a language barrier between the Danish students and the researcher is avoided with the usage of translation programs, recorded interview and the time reserved for the interview. During the interview, the students are assured that they can use either English or their native language to answer the questions. It is made clear that they can discuss with each other, and that they may use a translator software provided by the researcher. These precautions are taken to ensure that the students feel comfortable enough to answer the questions and that the interaction between the researcher and the students is as fluent as possible. The unfamiliarity of the researcher and the situation is countered by holding a group interview instead of a personal interview, as has already been discussed.

## **Analysis**

The qualitative data is analyzed using the analysis of content. The analysis is made by thematizing and coding the contents of the interview and trying to find out how the students learn each aspect of the English language.

An analysis of content was chosen, because it will reveal more detail on how the aspects of English are learned in the context of the chosen themes. It will also illuminate the underlying themes of language learning and give some extra information on how the students assess learning certain aspects of English language. In the analysis process, the themes were compared to the three aspects of English that the quantitative part of the research studied. The interview was analyzed with the goal to find out how and where the pupils learn the three aspects of language (reading, listening, and writing) and how they feel about learning English.



## 5. Results and Discussion

This chapter illuminates the results of the study, answers the research questions, and discusses how the said results fit into the provided theory background. Firstly, by analyzing the quantitative data gathered from English language competency tests. The second section of this chapter is dedicated to the analysis of how and where the language is learned, and the third part of the chapter looks at the mixed methods data that merges the two data sets. The final part of this chapter answers to the research questions and discusses how the results fit into the theoretical background.

### 5.1 English Language Competence

This portion of the thesis focuses on explaining the results of the qualitative tests by going through all results with extreme detail. The measured variables are:

Reading_Pre,	Reading_Post,
Listening_Pre,	Listening_Post,
Writing_Pre,	Writing_Post,
Total_Pre	Total_Post

The data points were added in by hand and double-checked by the researcher after the grading process, and at the beginning of the analysis process. For the sake of clarity, the variables are first looked as two groups (Pre-test and Post-test) and after that, all variables

are explained in more detail. After a detailed inquiry, the variables are run through non-parametric variants of T-tests and ANOVA as explained in chapter 4.1.

The analysis process began with acquiring the exploratory statistics for all variables and testing if the variables follow the normal distribution. The variables are divided into two groups for ease of understanding and use. The first group is pre-tests and the total of those tests and the second is the post-test data and the total.

Table 4 shows the listwise valid value of participants as well as the descriptive statistics of all Pre-test variables. The listwise value is 60 (N=60). The means for the variances are; Reading\_Pre mean is 17.433 (6.019), Listening\_Pre mean is 17.594 (6.387), Written\_Pre mean is 10.626 (10.329), and the Total\_Pre mean is 46.385 (16.984) marks. From the data presented in Table 4 can be seen that the SD of all the variables is relatively high.

**Table 4. Descriptive Statistics of Pre-tests**

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Reading_Pre	63	1.00	25.00	17.4325	6.01865	-1.070	.302	.242	.595
Listening_Pre	63	.00	24.37	17.5941	6.38686	-1.751	.302	2.504	.595
Written_Pre	60	.00	25.00	10.6262	10.32874	.190	.309	-1.731	.608
Total_Pre	63	11.25	73.25	46.3852	16.98373	-.378	.302	-.887	.595
Valid (listwise)	N 60								

Table 4 also shows the kurtosis and skewness of the pre-test data. According to the data, the Reading\_Pre skewness is -1.07 (0.302) and kurtosis is 0.242 (0.595). the same data for Listening\_Pre is skewness = -1.751 (0.302) and kurtosis = 2.504 (0.595). For Writing\_Pre The skewness is 0.19 (0.309) and kurtosis is -1.731 (0.608). For the total marks, the skewness is -0.378 (0.302) and kurtosis is -0.887 (0.595). These results (skewness, kurtosis, and SD) and the graphical expression of data by Q-Q plots (as seen in figures 3-6) all point towards the conclusion that the data would not follow the normal distribution.

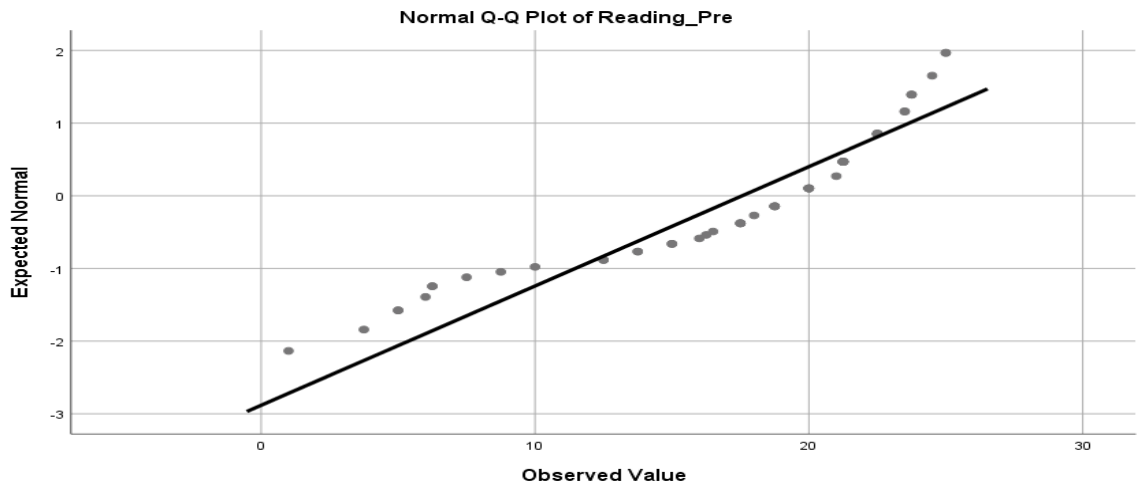


Figure 3. The Normal Q-Q Plot of Reading\_Pre

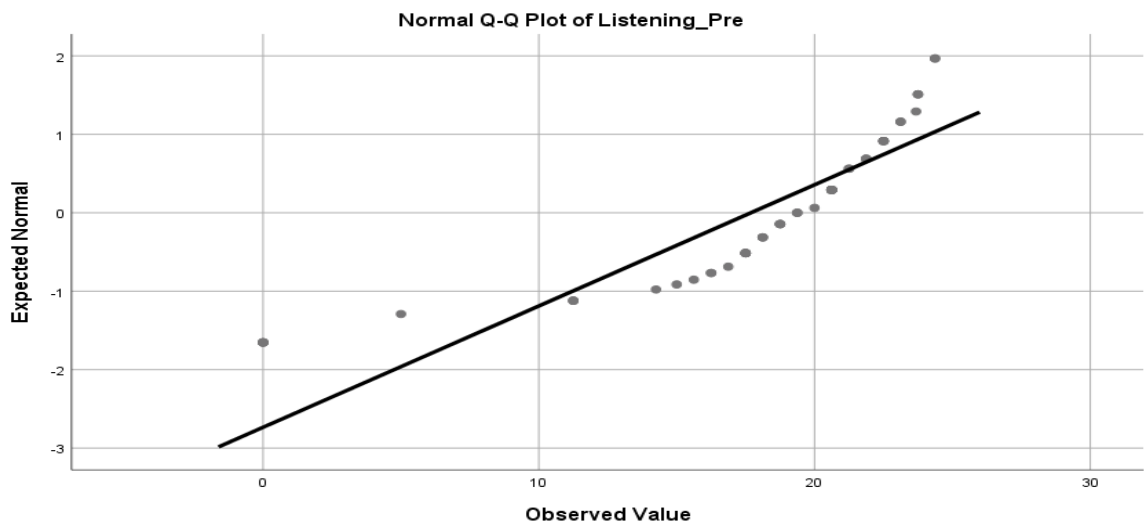
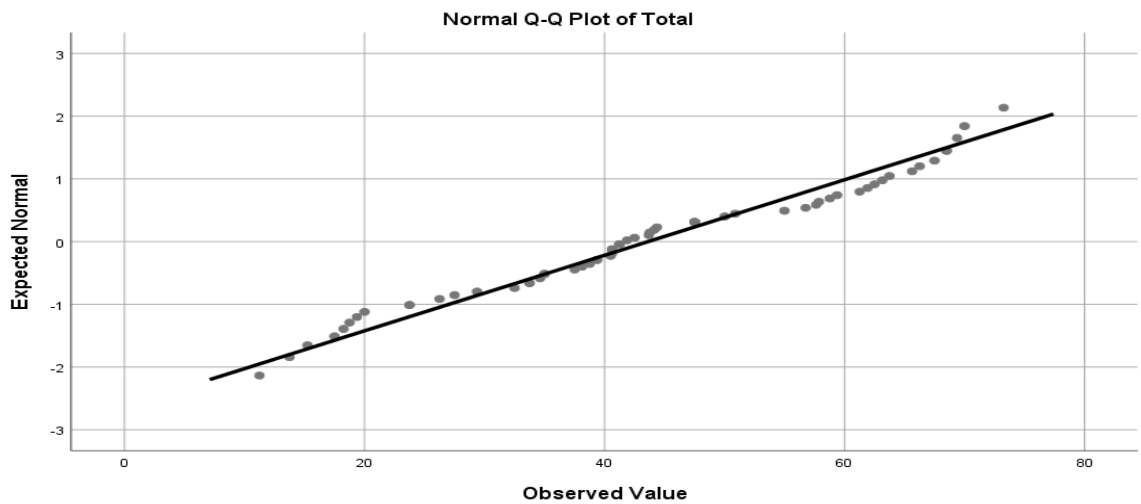


Figure 4. Normal Q-Q Plot of Listening\_Pre



Figure 5. Normal Q-Q Plot of Written\_Pre



**Figure 6. Normal Q-Q Plot of Total Marks**

To test the normality of variables, all variables were run through Kolmogorov-Smirnov and Shapiro-Wilk tests. As seen in Table 5 Kolmogorov-Smirnov test clearly shows that Reading\_Pre (D =0.179, P <0.000), Listening\_Pre (D =0.238, P <0.000), and Written\_Pre (D =0.248, P <0.00) do not follow the normal distribution. However, the variable of total marks does follow the normal distribution, since the Kolmogorov-Smirnov test returns the result of (D =0.085, P >0.05). Since Kolmogorov-Smirnov tends to favour rejecting the null hypothesis Shapiro-Wilk test was done beside it. Shapiro-Wilk returns the same conclusion: Reading\_Pre (D =0.0.869, P <0.000), Listening\_Pre (D =0.762, P <0.000), and Written\_Pre (D =0.798, P <0.00) and the total marks (D =0.964, P>0.05). This confirms the idea that all variables, except total marks, are non-normally distributed and therefore to test them further non-parametric tests should be used.

**Table 5. Tests of Normality for Reading, Listening, Writing and Total Pre-test Variables**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Reading_Pre	.179	60	.000	.869	60	.000
Listening_Pre	.238	60	.000	.762	60	.000
Written_Pre	.248	60	.000	.798	60	.000
Total_Pre	.085	60	.200*	.964	60	.075

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

After this, the same analysis was done to the post-test variables. Table 6 shows that listwise valid value of participants is 47. The means for the variances' are; Reading\_Post mean = 22.086 (4.309), Listening\_Post mean = 20.347 (5.221), Written\_Post mean is 15.343 (8.860), and the total mean = 51.608 (19.680) marks. From Table 6 can be seen that the points have increased in all variables when comparing to pre-tests.

Table 6 It also provides a more detailed look at the post-test data. According to the data, the skewness of Reading\_Post is -2.444 (0.306) and kurtosis is 5.947 (0.604). the same data for Listening\_Post is Skewness = -2.453 (0.302) and kurtosis = 6.723 (0.595). For Written\_Post The Skewness is -0.673 (0.325) and kurtosis in -1.058 (0.639). For the Total\_Post the skewness is -0.999 (0.293) and kurtosis is 0.249 (0.578). These results and the graphical expression of data by Q-Q plots (as seen in graphs 5-8) all point towards that the data would not follow the normal distribution.

**Table 6. Descriptive Statistics for Post-test Variables**

	N	Minimum	Maximum	Mean	Std.	Skewness	Kurtosis		
					Deviation			Std. Error	Statistic
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Reading_Post	61	5.00	25.00	22.0861	4.30888	-2.444	.306	5.947	.604
Listening_Post	63	.00	25.00	20.3472	5.22080	-2.453	.302	6.723	.595
Written_Post	54	.00	25.00	15.3425	8.86006	-.673	.325	-1.058	.639
Total_Post	67	.00	73.75	51.6075	19.67984	-.999	.293	.249	.578
Valid N (listwise)	48								

According to the Kolmogorov-Smirnov and Shapiro Wilk test of the post-tests (Table 7), the post-test data seems not to follow the normal distribution. According to Kolmogorov-Smirnov test: Reading\_Post (D =0.284, P <0.000), Listening\_Post (D =0.156, P <0.000) Writing\_Post (D =0.209, P <0.00) and Total\_Post (D =0.172, P <0.000) variables do not follow the normal distribution. To reassure this result a Shapiro-Wilk test was performed on the same data. Reading\_post (D =0.620, P <0.000), Listening\_Post (D =0.862, P <0.000), Writing\_Post (D =0.845, P <0.000), and the Total\_Post (D =0.868, P <0.000). These findings are in unison to Normal Q-Q plot graphs (figures 7-10) this reinforces the conclusion that the Post-test variables do not follow the normal distribution. Therefore, to calculate

differences between Pre-test and post-test non-parametric versions of T-tests and ANOVA are used.

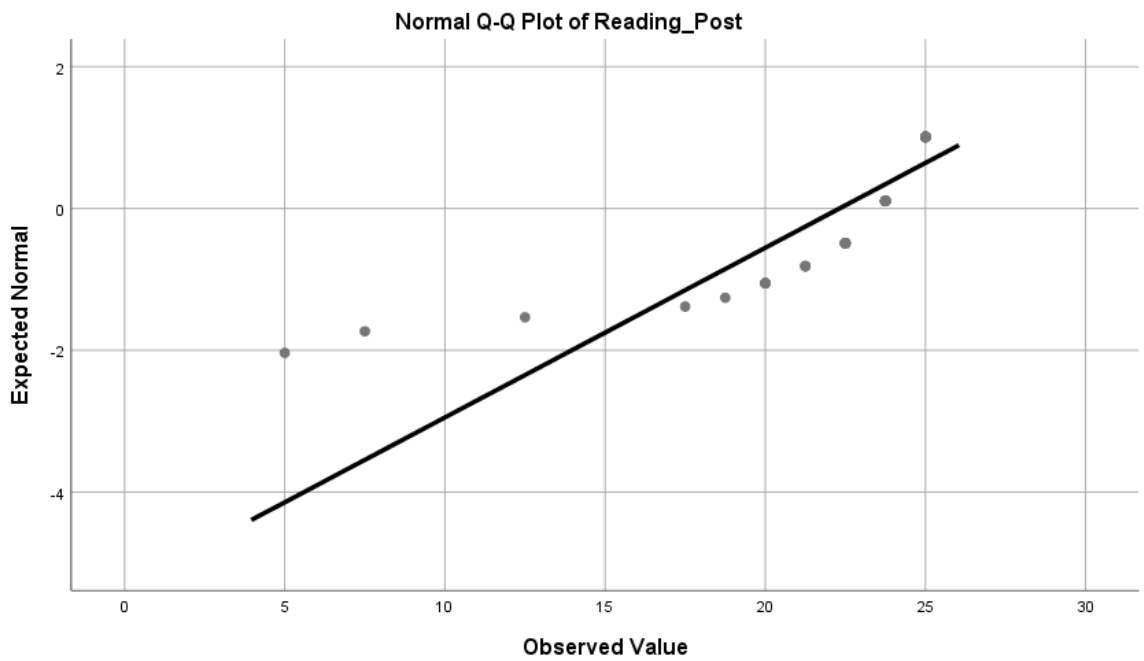


Figure 7. Normal Q-Q Plot of Reading\_Post

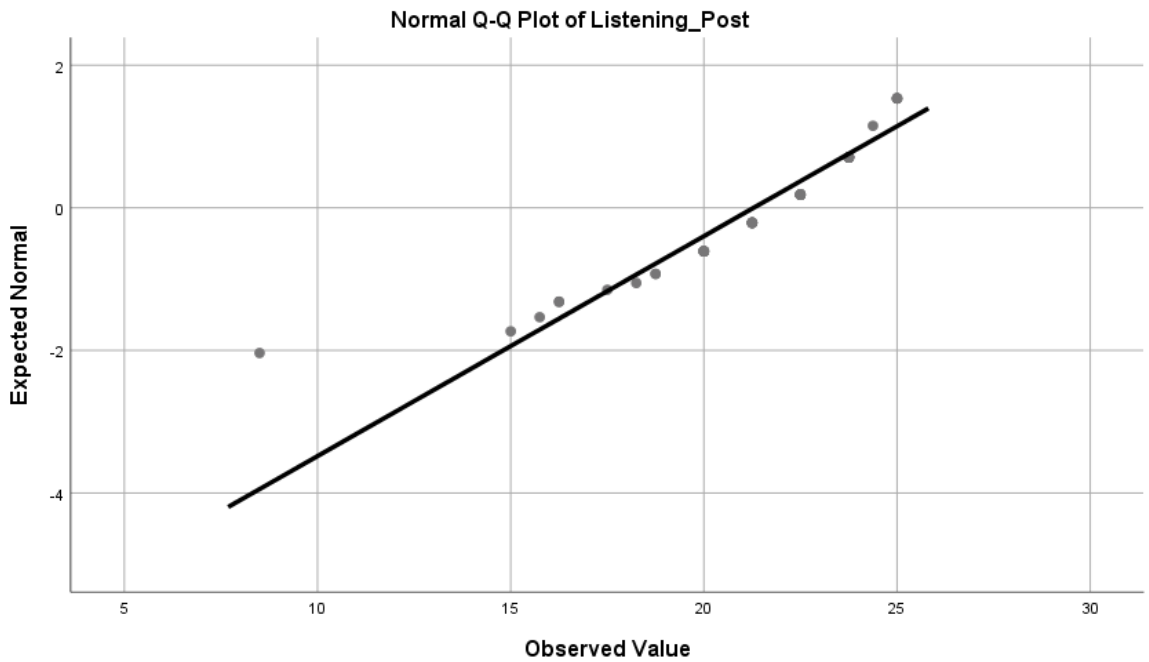
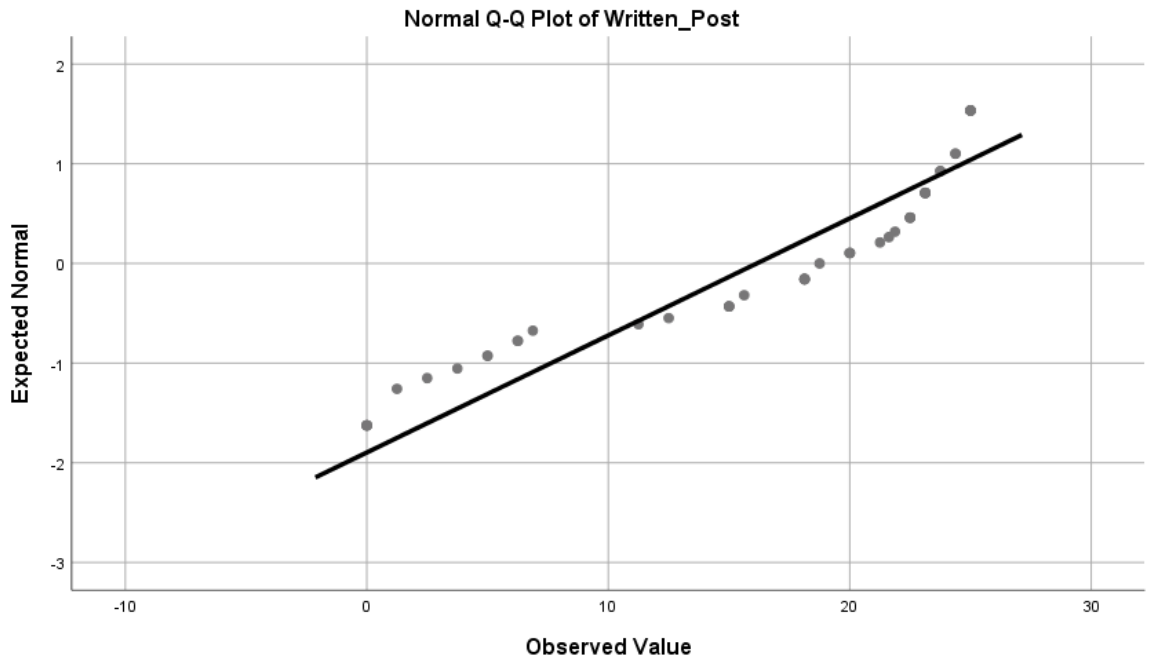
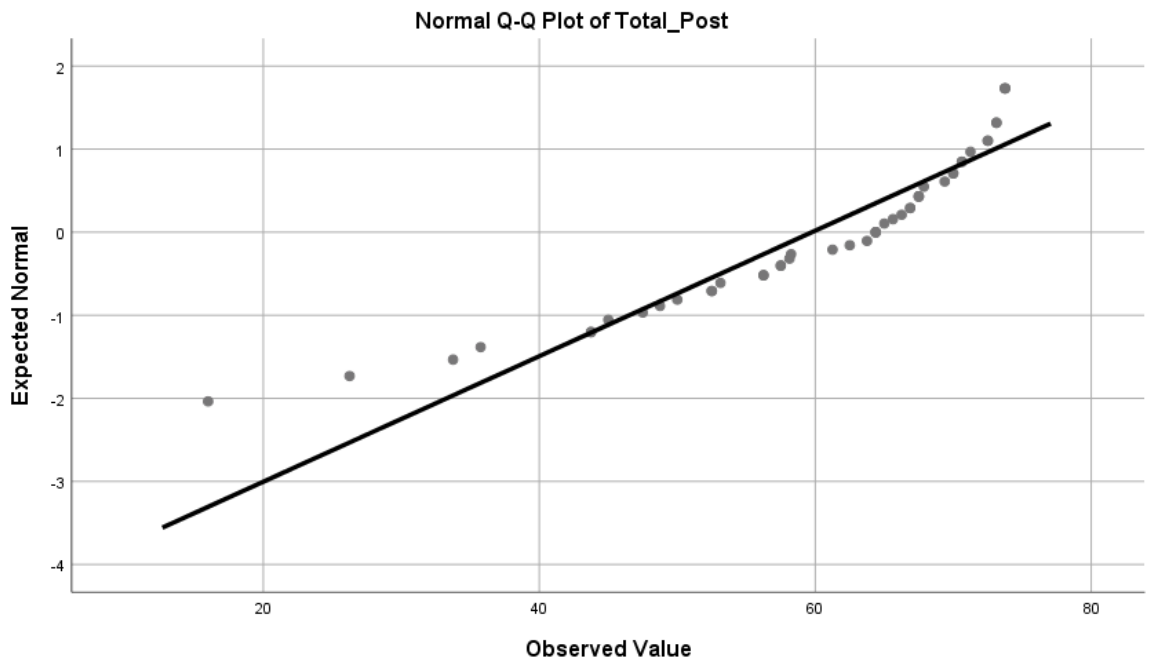


Figure 8. Normal Q-Q Plot of Listening\_Post



**Figure 9. Normal Q-Q Plot of Written\_Post**



**Figure 10. Normal Q-Q Plot of Total\_Post**

**Table 7. Test of Normality for Post-test Variables**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Reading_Post	.286	48	.000	.615	48	.000
Listening_Post	.161	48	.003	.857	48	.000
Written_Post	.210	48	.000	.840	48	.000
Total_Post	.174	48	.001	.863	48	.000

a. Lilliefors Significance Correction

Therefore, non-parametric variants of T-test and ANOVA tests are used to analyse the connection between the pre-tests and post-test and to see if there is any correlation between the two data sets

After these findings, the data was transformed using SPSS transform command. This was used to transform the data to fit a normal distribution. However, the Kolmogorov-Smirnov test and Shapiro-Wilk test still returned results that pointed towards the non-normally distributed data.

After this, the pre-test data were compared to post-test data to find if there is a statistically significant difference between them. This was achieved by using nonparametric variant for paired samples T-test, more specifically, a Wilcoxon signed-rank test on all variable pairs. (Pairs being Reading\_Pre – Reading\_Post, Listening\_Pre - Listening\_Post, Writing\_Pre - Writing\_Post, Total\_Pre - Total\_Post. ) After the test, the following formula was used to

calculate the effect size of the Wilcoxon signed-rank test:  $r = \frac{Z}{\sqrt{N}}$  The following paragraphs will demonstrate the results of these tests.

The first test was done to the Reading\_Pre (Mean = 17.432, SD = 6.019) – Reading\_Post (Mean = 22.092, SD = 4.274) data pair. The test showed that in seven months there has been a statistically significant increase in the students reading competence from the pre-test to post-test with a significant effect size between the measurements, (Z = -4.512, p < 0.000, r = 0.688) as seen in Table 10. Indeed, the difference between the median marks from the tests was five marks which are a significant effect of group as Table 8 Shows. The Table 9 shows that there are more positive ranks (Written\_Post > Written\_Pre) (N = 35)



than negative ranks (Reading\_Post < Reading\_Pre) (N = 5). This means that there has been a positive change in the results that the students gained from the tests.

**Table 8. Descriptive Statistics of Reading\_Pre and Reading\_Post**

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Reading_Pre	63	17.4325	6.01865	1.00	25.00	15.0000	18.7500	21.2500
Reading_Post	62	22.0927	4.27374	5.00	25.00	22.1875	23.7500	25.0000

**Table 9. Ranks for Reading\_Post – Reading\_Pre**

		N	Mean Rank	Sum of Ranks
Reading_Post - Reading_Pre	Negative Ranks	5 <sup>a</sup>	15.00	75.00
	Positive Ranks	35 <sup>b</sup>	21.29	745.00
	Ties	3 <sup>c</sup>		
	Total	43		

a. Reading\_Post < Reading\_Pre

b. Reading\_Post > Reading\_Pre

c. Reading\_Post = Reading\_Pre

**Table 10. Test Statistics Reading\_Post -Reading\_Pre <sup>a</sup>**

	Reading_Post - Reading_Pre
Z	-4.512 <sup>b</sup>
Asymp. Sig. (2-tailed)	.000

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Secondly, the Listening\_Pre - Listening\_Post pair was compared to each other. The test showed that in seven months there has been a statistically significant increase in the students listening competence from the pre-test to post-test with medium effect size between the measurements ( $Z = -2.778$ ,  $p = 0.005$ ,  $r = 0.414$ ). Indeed, the difference between the median marks from the tests was 1.25, as table 10 shows. The Table 11 shows that there are more positive ranks (Written\_Post > Written\_Pre) (N = 26) than negative ranks (Reading\_Post < Reading\_Pre) (N = 13). This means that there has been a positive change in the results that the students gained from the tests.

**Table 11. Descriptive Statistics for Listening\_Pre - Listening\_Post**

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Listening_Pre	63	17.5941	6.38686	.00	24.37	16.8700	18.7500	21.8700
Listening_Post	64	20.3809	5.18619	.00	25.00	20.0000	21.2500	23.7500

**Table 12. Ranks for Listening\_Pre - Listening\_Post**

		N	Mean Rank	Sum of Ranks
Listening_Post - Listening_Pre	Negative Ranks	13 <sup>a</sup>	14.69	191.00
	Positive Ranks	26 <sup>b</sup>	22.65	589.00
	Ties	6 <sup>c</sup>		
	Total	45		

- a. Listening\_Post < Listening\_Pre
- b. Listening\_Post > Listening\_Pre
- c. Listening\_Post = Listening\_Pre

**Table 13. Test Statistics for Listening\_Pre – Listening\_Post**

	Listening_Post - Listening_Pre
Z	-2.778 <sup>b</sup>
Asymp. Sig. (2-tailed)	.005

- a. Wilcoxon Signed Ranks Test
- b. Based on negative ranks.

Thirdly, the Writing\_Pre (Mean = 11.358, SD = 10.602) - Writing\_Post (Mean = 15.632, SD = 8.683) pair was compared to each other. The test showed that in seven months, there has been a statistically significant increase in the students writing competence from the pre-test to post-test with medium effect size between the measurements ( $Z = -1.992$ ,  $p < 0.05$ ,  $r = 0.314$ ) as seen in Table 13. Indeed, the difference between the median marks from the tests is +6.255, as table 12 shows. In table 11. Can be seen that there are more positive ranks (Written\_Post > Written\_Pre) (N = 21) than negative ranks (Reading\_Post < Reading\_Pre) (N = 11). This means that there has been a positive change in the results that the students gained from the tests.

**Table 14. Descriptive Statistics for Writing\_Pre - Writing\_Post**

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Written_Pre	63	11.3583	10.60249	.00	26.00	.0000	11.8700	23.1200
Written_Post	53	15.6320	8.68321	.00	25.00	6.5625	18.1250	23.1250

**Table 15. Ranks for Writing\_Pre - Writing\_Post**

		N	Mean Rank	Sum of Ranks
Written_Post - Written_Pre	Negative Ranks	11 <sup>a</sup>	14.32	157.50
	Positive Ranks	21 <sup>b</sup>	17.64	370.50
	Ties	8 <sup>c</sup>		
	Total	40		

a. Written\_Post < Written\_Pre

b. Written\_Post > Written\_Pre

c. Written\_Post = Written\_Pre

**Table 16. Test Statistics for Writing\_Pre - Writing\_Post<sup>a</sup>**

	Written_Post - Written_Pre
Z	-1.992 <sup>b</sup>
Asymp. Sig. (2-tailed)	.046

a. Wilcoxon Signed Ranks Test b. Based on negative ranks.

Finally, the Total\_Pre (Mean = 11.358, SD = 10.602) - Total\_Post (Mean = 15.632, SD = 8.683) pair was compared. The test showed that in seven months there has been a statistically significant increase in the students' writing competence from the pre-test to post-test with a medium effect size ( $Z = -2.462$ ,  $p < 0.05$ ,  $r = 0.355$ ) as seen in Table 19. Indeed, the difference between the median marks from the tests is +6.255, as Table 16 shows. Table 18 shows that there are more positive ranks (Written\_Post > Written\_Pre) (N = 34) than negative ranks (Reading\_Post < Reading\_Pre) (N = 14).

**Table 17. Descriptive Statistics for Total\_Pre – Total\_Post**

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Total_Pre	63	46.3852	16.98373	11.25	73.25	34.6200	47.5000	61.8800
Total_Post	67	51.6075	19.67984	.00	73.75	42.5000	56.2500	67.5000

**Table 18. Ranks Total\_Pre – Total\_Post**

		N	Mean Rank	Sum of Ranks
Total_Post - Total_Pre	Negative Ranks	14 <sup>a</sup>	24.86	348.00
	Positive Ranks	34 <sup>b</sup>	24.35	828.00
	Ties	0 <sup>c</sup>		
	Total	48		

a. Total\_Post < Total\_Pre

b. Total\_Post > Total\_Pre

c. Total\_Post = Total\_Pre

**Table 19. Test Statistics for Total\_pre – Total\_Post<sup>a</sup>**

	Total_Post - Total_pre
Z	-2.462 <sup>b</sup>
Asymp. Sig. (2-tailed)	.014

a. Wilcoxon Signed Ranks Test b. Based on negative ranks.

The main conclusion to draw from these calculations made on the variable pairs is that there is a statistically significant difference between the pre-tests and post-tests. Because of the statistical difference between the pre-test – Post-test pairs, a test of correlation is in order to determine the direction and the strength of the development.

The correlation test was chosen to be the Spearman’s rho test. This test was chosen since one of the measured variables follows a normal distribution. The Spearman’s rho test returns the following results (Table 20). According to these results there is statistically significant correlation between all the pre-test post-test pairs. Reading\_Pre – Reading\_Post ( $r = 0.488$   $p = 0.001$   $N = 43$ ), Listening\_Pre – Listening\_Post ( $r = 0.470$   $p = 0.001$   $N = 45$ ), Written\_Pre – Written\_Post ( $r = 0.515$   $p = 0.001$   $N = 38$ ), and Total\_Pre – Total\_Post ( $r = 0.605$   $p = 0.000$   $N = 48$ ).

**Table 20. Spearman's Rho Results of Pre-Tests and Post-Tests**

			Reading_Post	Listening_Post	Written_Post	Total_Post
Spearman's rho	Reading_Pre	Correlation Coefficient	.488**	.468**	.476**	.541**
		Sig. (2-tailed)	.001	.001	.002	.000
		N	43	45	40	48
	Listening_Pre	Correlation Coefficient	.541**	.470**	.405**	.524**
		Sig. (2-tailed)	.000	.001	.010	.000
		N	43	45	40	48
	Written_Pre	Correlation Coefficient	.164	.267	.515**	.410**
		Sig. (2-tailed)	.306	.083	.001	.005
		N	41	43	38	46
	Total_Pre	Correlation Coefficient	.387*	.485**	.640**	.605**
		Sig. (2-tailed)	.010	.001	.000	.000
		N	43	45	40	48

As table 20 shows there are statistically significant positive correlations between all the variable pairs, this means that as time progresses there is an increase in the students' competence in English. Table 20 has been modified to only show the correlation between the pre-test and post-test results.

### Result

There is a statistically significant increase in the students' competence in English for over seven months. This can be seen in table 20 as well as by comparing the mean and median scores between the pre-tests and post-tests.

## 5.2 How and Where the English Language Is Learned

This part of the analysis is structured as follows, firstly a report on how the analysis was done, secondly, a report on the overall findings that were found and the figure which highlights the connections between the aspects of English and the learning style. And thirdly, there are several paragraphs that explain in greater detail the found connection between the aspect and the learning style. Finally, the last part of this chapter will give an account of what all this means. This part aims to answer the second research question.

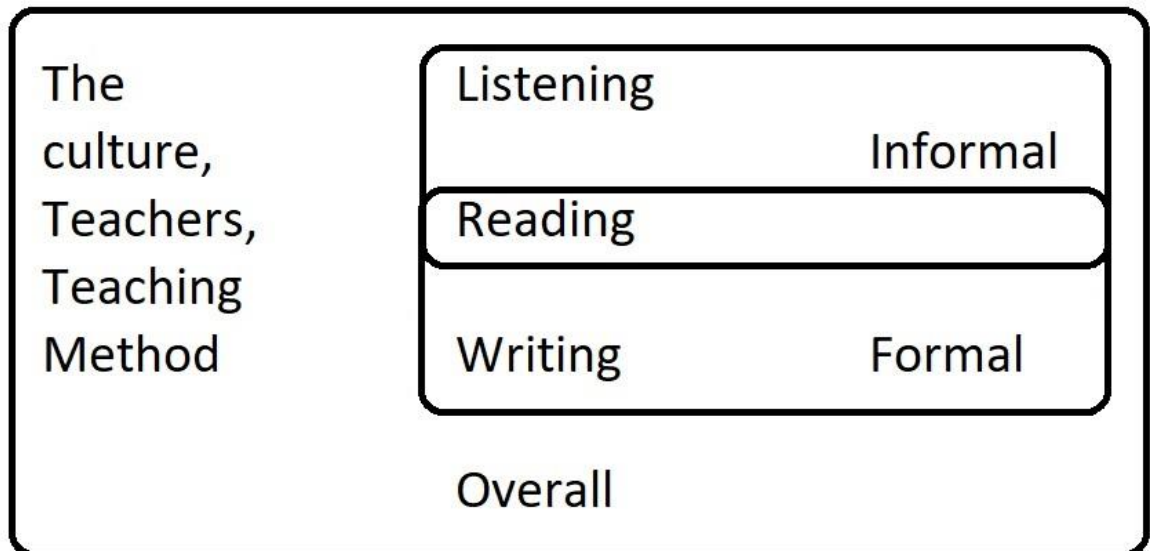
The interview took approximately an hour to complete, and it followed the directions set in the chapter 4.2. The interview had five participants plus the researcher, during the interview the participants were advised to talk with each other and use electronic aids to counter the language barrier between the researcher and the participants.

The interview was constructed around two main themes: **Formal Learning** and **Informal Learning**. Both themes were first approached from a more general aspect and then with a more direct question. This focus was chosen because of the second research question, where the aim was to find how and where the pupils were learning the English language. During the interview, the researcher explained the simplified versions of formal learning and informal learning to the participants in order to complete the direct inquires. A non-formal source of learning was left out since it was tied into formal learning.

To maintain the anonymity of the participants, their names have been changed to random Finnish names that have no connection to the participants. During the interview, a questionnaire framework was used to keep the focus of the interview on the chosen themes. (The framework can be seen in appendix 1.)

The initial analysis of the content revealed several overlapping themes that have affected the learning of the participants, in a deeper analysis these themes were looked through the lenses of the three aspects of English (Reading, writing, listening) and total scores of the tests. It became clear that these lenses combined different themes together. Figure 10 highlights the strongest link between the aspect of the English language and the learning source. The figure shows that Listening has the **strongest** link towards informal learning

and writing is associated strongly with formal teaching. However, reading as an aspect is associated with both sources of learning. The overall learning of the language is related to the culture of the school and the teachers and the methods that are used to teach the students.



**Figure 11. The Connections Between Aspects and Learning Styles**

### **Overall Learning of English Language**

The interview presented a theme that was more prominent than first thought. The participants said that more than just a teaching method, the **school's culture** promotes the learning of the English language. This theme includes both sources of learning (formal and informal) and it engulfs all aspects of English as well. To begin with, all the participants agree that there is a lot of English used; between sentences in speech, as a slang between themselves, and as a part of their daily life.

**Tiina:** "Because sometimes I speak (English) to friends and some people who don't speak"

**Mika:** "Yeah you also use English between the sentences, like it is a kind of slang that..."

(Tiina and Mika discuss the usage of English in their day to day lives.)

**Tiina:** "Well I have been using English a lot more"

**Researcher:** "Why?"

**Tiina:** "Because people speak English here. It's a part of the daily life"

**Liisa:** "It's part of the culture as well"

**Tiina:** "Yeah it's just of part of it in here umm, so a lot of people umm, talk in English, speak English and that has made me speak a lot more English and made me use more English and therefore I have trained English more and... become better at it"

Furthermore, the participants said that the pupils use English daily, in games, on the internet, and with each other. This puts a great emphasis on informal learning of the language, since the students are using it a lot with each other and on the internet, games, and social media.

**Tiina:** "English where I speak it more and listen more to it than reading it you know; I do also read it when I am on my phone sometimes"

**Liisa:** "It (Social media) is usually only English is also on facebook I read posts in English and follow English sites and so on. So, reading and writing there, there is not much speaking in social media"

**Martti:** "I am a gamer. I play (video games) a lot, I can tell you that I got most of the English by gaming and by going to a game and just play with international friends"

**Liisa:** "I communicate in English umm, in social media even with people who are Danish too"

Nevertheless, the participants make it extremely clear that there needs to be a caring, motivating and a professional teacher to help them.

**Niina:** "-- you have a teacher to correct it for you. And the corrections are not because you are bad but because you don't know it"

**Mika:** "If the teacher is not motivated then you are not motivated"

**Tiina:** "But I have had amazing teachers that helped me with my English so that I could get back on track with my English"

Participants gave critique to the teachers who were lazy about their teaching and did not care about their students or their work. Niina also said that she found it important that the teacher takes them (the pupils) and their questions seriously. She also said that it would increase the overall motivation towards the language and learning.

The participants considered a **normal** school to be boring and somewhat factory-like. Normal school was interpreted as something else than Østerskov Efterskole. They also stated that the teacher should use a variety of teaching methods to keep the lessons interesting and that it would help with school motivation.

**Tiina:** "I think normal schools are like hard to get through, coz it kind of boring so yeah"



**Liisa:** "I, the school for me seems lazy coz teachers don't always seem to care and they seem lazy about the way they teach"

**Researcher:** "Ah, that's ok"

**Martti:** "-- As a whole I consider the different areas and the things that the school does is making everything grey and everything a factory-like"

"I think school can be good if you are motivated it all comes down to that. If you don't have any good friends at school you are not motivated, If the teacher is not motivated then you are not motivated..." - **Mika**

"-- And especially teachers in schools must know that people learn different ways"  
- **Niina**

The participants said that the more they use the target language, the more their confidence on it increases, and thus, increases their willingness to use the language. As is stated during the interview, the participants have spoken very little English in their school before entering to Østerskov Efterskole. When they entered Østerskov Efterskole they started to use English more and their language confidence grew with it. Mika states that he could have not dreamt about using only English in a full weekend long convention before entering to Østerskov Efterskole, he says: "-- I would never in elementary school go to an international event and being told you cannot speak your native language you can only speak English..." (Mika)

In conclusion, it seems that the culture, teachers, and the teaching methods used in Østerskov Efterskole promote an inclusive usage of the English language. This increases the language use confidence of the students, makes the students feel appreciated and cared for, and removes the factory-like teaching. The interviewees stated, that although their ability in the English language has increased with informal learning, they still would need the help of a professional teacher to aid them. Therefore, the overall learning of English language is deemed to be a combination of both learning sources formal and informal. This brings up a question of which aspect of English is attributed to which learning source?

Furthermore, the participants feel like they have improved while being in the Østerskov Efterskole:

**Martti:** "My English definitely improved a lot in last year in exams I would get 7 or 8 but whit out a doubt that's the highest I have ever got in any test ever. I know I have improved, not even doubled but even more"

**Niina:** “I like humans learned don’t think I have been improving drastically but of course obviously practice makes perfect so because I use it more, I also get better at it. It’s not something that I would write home about so to say”

**Mika:** “Well as I have said before I have drastically improved in Østerskov. My English is way better now and I have more confidence speaking English I would never in elementary school go to an international event and being told you cannot speak your native language you can only speak English”

### **The Listening Aspect**

The listening aspect of the English language was considered mostly to be learned from informal sources of learning. For example, from the environment in Østerskov Efterskole, gaming, videos, and social media.

“I would say speaking and listening because I use English where I speak it more and listen more to it than reading it you know, I do also read it when I am on my phone sometimes umm, but mostly I speak and listen” - **Tiina**

“Without a doubt, I was told in all of those my private life that what I think is mostly reading and listening” - **Martti**

“Outside of school informal learning, I think I have learned most of my English from informal learning writing speaking listening, reading from informal learning” - **Mika**

Though some participants have received pronunciation help from teachers and have heard English in school, they stated that the biggest influence on their listening competence has been sources of informal learning. Still, the participants acknowledged that the proximity of their native language (Danish) and English.

**Liisa:** “-- Oh yeah, because a lot of the words are the same. Or are pronounced kind of the same or are spelled kind of the same so”

**Interviewer:** “The same when compared to?”

**Liisa:** “For example, [Danish word for window] when put against (the English word) each other sound kind of the same”

**Niina:** “That’s because of Vikings!”

[All agree]

### **Writing Aspect**

Learning to write English was attributed mostly to formal learning. The participants stated that during their elementary school they did little else apart from writing and grammar

exercises, these lessons were considered boring to the point of complete ignorance towards the subject.

However, the participants said that the presence of a good teacher helps a lot with the motivational issues when it comes to writing. The participants said that informal learning has not really been a big influencer on their writing competence, but still they said that there has been some influence from social media.

### **Reading Aspect**

The reading aspect was left with little notice during the interview. It was mentioned several times, but it was connected to both sources evenly. This could mean that the pupils have learned most of their skills in reading English from a nonformal formal source or that they have learned their skills evenly from both sources.

“They (teacher is her former school) thought us English kind of reading the book, you know read the book” – **Martti**

“We got home and we read the pages we had to for the next day. Then the next day you come to school, you are ready you read it” – **Niina**

“I think that writing and reading, is taught for me outside” – **Liisa**

In conclusion, it seems that the pupils learn the different aspects of the English language from different sources. Writing seems to be mostly learned from a formal source of learning while the listening aspect was mostly learned from a non-formal source. The reading aspect measured somewhere between these two since the interviewees mentioned it in connection to both learning sources.

Considering the overall language learning, the interviewees mentioned that they did not like school or formal learning before attending Østerskov Efterskole, and after enrolling in the school they have seen and experienced an increase in their language competence, as well as in their confidence in using the language.

### 5.3 Connection of the Learning Source and Language Competence

The results of the mixed-method analysis can be seen in table 21. In this table, both data sets were merged together in order to produce a joint display model to better highlight the

link between the quantitative and qualitative results. In table 21 the rows represent the different aspects of the English language that were tested (Reading, listening and writing) and the final row represents the total points and overall findings.

The first two columns “Change in mean” and “Change in Median” show the quantitative data increase of the mean and median between the pre-test and post-test, where the first number shows the mean of the pre-test, the second the mean of the post-test, and the third column represents the quantitative results that highlight the personal experiences of the interviewees to the language learning, and the last column shows the mixed method interpretation of the data as they are merged together.

The joint display model Table 21 shows us that there has been a statistically significant increase in the scores of all aspects of the English language. This can be seen from columns two and three. Column four shows that the learning of the aspects can be attributed to different sources of learning, these sources are formal and informal. The final column explains the mixed method interpretation of the integration of the data.

A conclusion can be drawn from the joint display model (Table 21) that there is a connection between the culture and pedagogy used in Østerskov Efterskole and the learning results. According to the interviewees’ personal experience the informal style of teaching increases their motivation and involvement in learning. This would explain the increase between the pre-test and post-test. The culture that promotes and encourages the students to use the language outside the classroom seems to be the key factor when trying to explain the results gained from the quantitative part of the study.

**Table 21. The Joint Display of Mixed Methods Findings**

<b>1. Aspect of English</b>	<b>2. Change in mean</b>	<b>3. Change in Median</b>	<b>4. Qualitative findings</b>	<b>5. Interpretation</b>
<b>Reading</b>	17.433 to 22.093  p<0,00	18.75 to 23.75	A clear connection to both sources of learning. The findings were surprisingly evenly distributed between both sources. This means that there could be an impact created by non-formal formal learning.	Both sources aid learning as much. The learning is attributed to both sources. However, the effect seems to be positive.
<b>Listening</b>	17.594 to 20.381  p<0,00	18.75 to 21.25	Mostly learned from informal sources. Gaming, social media, videos. The culture in Østerskov Efterskole has a big impact on the participants and they feel like their ability has grown.	The culture in Østerskov Efterskole has helped the students to increase their scores, but still, the increase is mostly because of the informal learning and cannot be attributed to a formal source of learning.
<b>Writing</b>	11.358 to 15.632  p<0,05	11.87 to 18.13	Even though the teaching methods and classes were considered boring during elementary school, the participants said that most of their writing competence has come from a formal source.	An increase in the writing results can be attributed to the increased motivation and participation in classes, this is due to the teaching methods and the culture of Østerskov Efterskole. It seems that even a task that the interviewees considered boring is made interesting in Østerskov Efterskole.
<b>Total/Overall</b>	46.385 to 51.608  p<0,05	47.50 to 56.25	The culture of Østerskov Efterskole and a healthy student-teacher relationship indicate a good learning environment that promotes language learning and aids the students to learn and use the English language. Overall the participants said that a motivational and caring environment helps with learning in general. Also, the interviewees had experienced an increase in their skills and confidence in using the English language.	The culture of Østerskov Efterskole and a healthy learning environment increase motivation to learn the language, as well as the culture that promotes using the language outside the classroom. This partially explains the results from the quantitative part of the tests.

## 5.4 Answers to Research questions

The answer to the first research question is a simple one, there is a statistically significant increase in the student's competence in different aspects of the English language between the pre-tests and the post-test, as seen in section 5.2.

The second research question requires a more complicated answer. It seems that the students have learned most of their English from different sources. The Reading portion was learned mostly from both sources (formal and informal sources of learning) or most likely from a source that was formal but used non-formal teaching methods. Listening was attributed mostly to non-formal sources, due to social media and the Internet. Writing, however, was attributed mostly to formal sources of learning. There is a clear connection towards the overall culture and environment that the Østerskov Efterskole has created to the students and it seems that the pupils of Østerskov Efterskole are motivated and are learning language skills better than in the schools they were before.

For the third research question, there seems to be a clear connection between the personal experience of learning English and the English language competency tests. Students said that they feel that their language competence had increased during their time on Østerskov Efterskole and that their confidence in using the language has also increased. On each tested aspect of English language, the increase of scores gained from the text can be attributed to the motivating factor of eduLARP as well as the environment and 'feeling' in the school itself.

To conclude seems that the eduLARP method helps in providing and improving a learning culture that promotes language learning and thus increases language competence. (see section 5.3)

## 5.5 Discussion

There is an increase in the quantitative test results over the time period which indicates that the school and its teaching methods work. Across the tested variables, there has been a statistically significant increase. However, interpreting the connection between the eduLARP and the results proved to be a more complicated matter.

The interview shows that the experiences of the students on how they learn best varies from the aspect of language to another aspect of language. Most of the listening is learned from the informal sources of learning and most of the writing and grammar from the formal sources. From the interview can also be derived a theme of the culture and environment, which is clearly the greatest influencer to learning according to the participants. The culture consists of aspects from both sources of learning and ties them into a schoolwide environment that is not present only during breaks or LARPing, but also during the classes and more formal teaching.

In the qualitative part of the research, the participants say that their competence and confidence in English have increased while they have been attending to the Østerskov Efterskole. When this is compared to the results gained from the quantitative part of the research a connection can be seen. Students feel that their abilities have increased during the time they have been in the school, and the data shows their feelings to be correct.

Therefore, the connection made in chapter 3.3 between the eduLARP and L2 learning is plausible and can be proven with scientific methods. However, more research is required to give further proof of this theory.

## 6. Conclusion and Ethics

This section of the thesis explains and discusses the conclusions and ethics of the research as well as some pitfalls the research encountered along the way.

### 6.1 Conclusion

In conclusion, the research has been a success. The research premise functioned as expected and all research tasks were finished within a reasonable time. Furthermore, all of the research questions were answered and provided new information about eduLARP.

A multitude of research choices in this study were made because of the time restraint, number of researchers, available funding, the scope of master's thesis, and other very practical reasons that limited the study. For example, the possibility of utilizing a control group from another school, creating and holding proper pilot tests and interviews, and a longer research period with multiple measure points of both qualitative and quantitative data, were not possible due to the reasons presented before.

The issues of a short longitudinal study, inability to hold proper pilot interviews due to time constraints and funding are the main constraints of this research. There were a few questions that were asked after the interview via e-mail, however because of time constraints these questions did not make it into the research. These issues might have been able to be countered by research team or a partner, and proper funding.

Another thing to take into consideration is the computer usage of Østerskov Efterskoles' students. They seem to use a computer in their daily learning process and in their daily lives, playing videogames and chatting with their friends. This claim is based on my own



observation of how the students behave in their free time. Therefore, I would dare to make the claim that the students would have scored higher results on the writing tests if they would have been able to complete the writing test with a computer, even without a spellcheck and grammar aids. However, the decision to use a pen and paper on this test was made on the grounds that the test provided by ECL was directed to be written by hand.

The research process itself begun with a slight misunderstanding between the researcher and the Østerskov Efterskole. The research was first meant to encompass only the students who are in Østerskov Efterskole for the first year, this misunderstanding doubled the number of research participants in one night and resulted in a far greater data set. Furthermore, the inclusion of older students might have affected the result some way, but their competence also increased during the test period.

For future research, the possibility of a control group from another school should be explored, that could provide a reliable source of peer data to compare to. A more specific research could investigate how the eduLARP affects the special needs students. Another possibility for future research is to repeat this study with proper funding and follow more than one school subject for a longer period of time, with the corrections stated above. Another interesting opportunity would be to measure the level of involvement in the student or how the teachers see or feel about the eduLARP.

## 6.2 Ethics

The study is conducted and planned to be as ethical as possible for everyone involved. Consequentialism derives itself from the notion that the rightness of an action is only assessable by the outcome of it (Harrison & Rooney 2012, 32). With this ideal, ethical problems seem almost trivial to overcome. however, there has been some critique of the consequentialist approach to science, especially the Utilitarian part. According to Harrison and Rooney (2012) the given explanation to the critique towards utilitarianism is mostly majoritarian. This means that utilitarianism bases its interest on the majority of the populus. Therefore, the result of a test that follows this ideal, is clearly towards a popular result more than a true and correct one. (2012, 32.)

Therefore, this study employs deontology, where the concepts of right and wrong arise from the individual's duty to oneself and others (Harrison & Rooney 2012, 33). Harrison and Rooney introduce the main ideas behind deontology, Immanuel Kant and W. D. Ross. According to Harrison and Rooney two principles are essential when applying deontology to research:

“first, the duty of the researcher to treat the subject as the researcher would wish to be treated; and second, the historically more recent development of the notion of rights. This second principle means that the research subject has rights, and in particular, the right to be properly and adequately informed about the nature, impact, and outcomes of the research and to consent to participation in the research.” (Harrison & Rooney 2012, 38.)

This study aims to be as explicit and thorough as possible. All the participants are made aware of the study, and all participants are asked their permission to take part in this study. In the case of underage students, their parents/caretakers are asked the permission to participate in the study. According to Edwards and Hillyard (2012), the researcher should leave the field with the knowledge that their presence did not have any deleterious effect. However, they list three main effects the researcher will affect school life:

1. affecting the behavior of staff and pupils
2. detracting from time teachers or pupils spend in, or preparing for, classroom activities
3. become a logistical problem or a hindrance to the schools' finite resources.

(Edwards & Hillyard 2012, 135-136.)

With these points considered, the study itself aims to be as ethical as it possibly can and follow the guidelines set by the Finnish Advisory Board on Research Integrity (TENK). A complete list of responsible conduct of the research can be found in “Responsible conduct of research and procedures for handling allegations of misconduct in Finland” (Finnish Advisory Board on Research Integrity 2012, 30-31). The research aims to comply with these regulations and is revised before the execution of the field study and publishing so that no guidelines are broken. The researcher also vouches to make any alteration needed to fulfill the TENK's guidelines.

The actual concrete actions that are taken involve providing interviewed students with pseudonyms, keeping the amount of information the researcher has on the students minimal. The names of the students will not be recorded on the final paper. However, naming the students during the research period is vital to track individual progress between the language competence pre-test and the post-test. Other possible solutions to this would be to assign the students by number, but that would require the students to remember the number over a lengthy period. These names will not be revealed to any other parties apart from the researcher. The names are assigned to a randomized number during the analyzing period to assure the anonymity of the students. The research also requires a group interview. The interviewees are all treated equally, and all the interview follows a semi-structured plan. In the final paper, the students who take part in the interviews are given them pseudonyms to provide the interviewees with anonymity.

## 7. References

- Barnes, O. Flack, K. & Mason, M. (2008). *Warhammer 40000: Dark Heresy*. Nottingham, England: Black industries.
- Bowman, S.L. (2014). Educational Live-action Role-playing Games: A Secondary Literature Review. In S.L. Bowman ed., *The Wyrd Con Companion Book*. Wyrd Con, pp. 112-132.
- Bowman, S.L. and Standiford A. (2015). Educational LARP in the Middle School Classroom: A Mixed-Method Case Study. *International Journal of Role-playing*. 5(1).
- Caillois, R. (1961). *Man, Play and Games*. Trans. B Meyer. Glencoe: The free press of Glencoe.
- Carter, M., Sims, C., Grey, S.F. and Perkins, C. (eds). (2014). *Dungeons & Dragons: Fifth Edition: Players Handbook*. Renton, California: Wizards of the Coast.
- European Centre for the Development of Vocational Training. (2008). *Terminology of European Education and Training Policy: A Selection of 100 Key Terms*. Office for Official Publ. of the European Communities. Available at: [http://www.cedefop.europa.eu/en/files/4064\\_en.pdf](http://www.cedefop.europa.eu/en/files/4064_en.pdf) [Accessed 16 Sep. 2019].
- Cohen, L., Manion, L. and Morrison, K. (2018). *Research Methods in Education*. 8th ed. Routledge, 2018.
- Council of Europe, Education Department Education Policy Division Language Policy Programme. (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment: companion volume with new Descriptors*.
- Crawford, K. (2011). *Stars Without Number*. Sine Nomine publishing.
- Creswell, J.W. and Plano Clark V.L. (2017). *Designing and Conducting Mixed Methods Research*. 3rd Ed. Thousand Oaks, California; London, United Kingdom; New Delhi, India; Singapore: SAGE.
- Curtis, W., Murphy, M. and Shields, S. (2013). *Research and Education*, Routledge, London. [ebook] Available at: ProQuest Ebook Central. [Accessed 12 Sep. 2019].
- Daniau, S. (2016). The Transformative Potential of Role-Playing Games: From Play Skills to Human Skills. In *Simulation & Gaming*. 47(4), pp. 423-444.

- Denscombe, M. (2010). *The Good Research Guide: for Small-Scale Social Research Projects*, McGraw-Hill Education, United Kingdom.
- Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011a). 'From Game Design Elements to Gamefulness: Defining Gamification', In *Proceedings of the 15th international Academic MindTrek Conference: Envisioning future media environments*, pp. 9-15. ACM.
- Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011b). Gamification: Toward a Definition. [Online] Available at: [gamification-research.org/wp-content/uploads/2011/04/02-Deterding-Khaled-Nacke-Dixon.pdf](http://gamification-research.org/wp-content/uploads/2011/04/02-Deterding-Khaled-Nacke-Dixon.pdf) [Viewed 19 Sep. 2019].
- Edwards, M. and Hillyard, S. (2012). Improvisation, Ethical Heuristics and the Dialogical Reality of Ethics in the Field, in K Love ed. *Ethics in Social Research*. [ebook] Available at: ProQuest Ebook Central. Bingley: Emerald Group Publishing Limited. [Accessed 16 Feb. 2018].
- Ellis, R. (1997). *Second language acquisition*. Oxford, England, Oxford University Press.
- Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford, England, Oxford University Press.
- European Consortium for the Certificate of Attainment in Modern Languages. (2018a). Objectives. Available at: <http://eclexam.eu/objectives/> [Accessed 14 Feb. 2018].
- European Consortium for the Certificate of Attainment in Modern Languages. (2018b) Task types. Available at: [http://eclexam.eu/files/task\\_types.pdf](http://eclexam.eu/files/task_types.pdf). [Accessed 14 Feb. 2018].
- European Consortium for the Certificate of Attainment in Modern Languages. (2018c). Quality Assurance. Available at: <http://eclexam.eu/quality-assurance/> [Accessed 14 Feb. 2018].
- Feilzer, M.Y. (2010). Doing Mixed Methods Research Pragmatically: Implications for the Rediscovery of Pragmatism as a Research Paradigm. *Journal of Mixed Methods Research*, 4(1), pp. 6–16.
- Figueras, N. (2012). The impact of the CEFR. *ELT Journal*, [online] 66(4), p. 477–485. Available at: <https://doi.org/10.1093/elt/ccs037> [Accessed 18 Sep. 2019].
- Fontana, A. and Prokos, A.H. (2007). *The Interview: From Formal to Postmodern*. Walnut Creek, California, Left Coast Press.
- Fox, J., Bailenson, J. and Binney, J. (2009). Virtual experiences, Physical Behaviors: The Effect of Presence on Imitation of an Eating Avatar. *Presence: Teleoperators and Virtual Environments*. 18(4), pp. 294-303.

- Fox, J. and Bailenson, J.N. (2009). Virtual Self-modeling: The Effects of Vicarious Reinforcement and Identification on Exercise Behaviors. *Media psychology*. 12 (1), pp. 1-25.
- Frey, J.H. and Fontana, A. (1991). The group interview in social research. *The Social science journal*. 28(2), pp. 175-187.
- Gjedde, L. (2013). Role Game Playing as a Platform for Creative and Collaborative Learning. In *European conference on game-based learning*. Vol. 1, p. 190. Academic Conferences International Limited.
- Gjedde, L. (2014) October. Potentials of a Narrative Game-based Curriculum Framework for Enhancing Motivation and Collaboration. In *European conference on games based learning*. Vol. 1, p. 151. Academic Conferences International Limited.
- Gray, S.F., Carter, M., Sims, C. and Wilkes, J.C. (eds.) (2014). *Dungeons & Dragons: Fifth Edition: Dungeon masters guide*. Renton, California: Wizards of the coast.
- Halinen, I. (2015). *OPS2016 Uudistuva perus- ja lisäopetus*, Opetushallitus 2015.
- Hamari, J. (2015). *Gamification motivations & effects*. Aalto university. Doctoral dissertations 11/2015.
- Harder, S. (2007). Confessions of a Schoolteacher: Experiences with Roleplaying in Education. *Lifelike*. pp.229-235.
- Harrison, J. and Rooney, D. (2012). Knowledge and Its Integrity within a Knowledge Economy: Ethics and Wise Social Research. In K. Love (ed.) *Ethics in social research*. [ebook] Bingley: Emerald Group Publishing Limited. Available at: ProQuest Ebook Central. [Accessed 14 Feb. 2018].
- Harviainen, J.T., Lainema, T. and Saarinen, E. (2014). Player-reported impediments to game-based learning. In *Transactions of the digital games research association*. 1(2), pp.55-83.
- Harviainen, J.T. and Savonsaari, R. (2013). Larps in High Schools. In A. Moseley and N. Whitton (Eds.) *New traditional games for learning*. [ebook] pp 134-145. London: Routledge.
- Harviainen, J.T. (2016). Physical Presence in Simulation: A Scratch at the Surface of Complexity. *International journal of role-playing*. 6, pp.32-38.
- Heliö, S. (2004). Role-playing: A narrative Experience and a Mindset. In M. Montola & J. Stenros (eds.) *beyond role and play tools, toys and theory for harnessing the imagination*. Helsinki: Ropecon. pp.65-74.

- Henriksen, T. (2004). On the Transmutation of Educational Role-play. In M Montola & J Stenros (eds.) *Beyond role and play tools, toys and theory for harnessing the imagination*. Helsinki: Ropecon, pp.107- 130.
- Hjalmarsson, S. (2011). Live-action Role-play as a Scenario-based Training Tool for Security and Emergency Services. In *European conference on cyber warfare and security*. pp. 132. Available at: Academic Conferences International Limited.
- Hulstin, J.H. (2007). The Shaky Ground Beneath the CEFR: Quantitative and Qualitative Dimensions of Language Proficiency. In *The Modern Language Journal*, 91(4), pp.663-667.
- Huotari, K. and Hamari, J. (2012). Defining Gamification: a Service Marketing Perspective. In *The Proceedings of the 16th International Academic MindTrek Conference*, Tampere, Finland, October 3-5, 2012, ACM. pp. 17-22.
- Howatt, A.P.R. (1985). *A history of English language teaching*. 2. corr. impr. Oxford: Oxford U. P.
- Hyltoft, M. (2008). The role-players' school: Østerskov Efterskole. In *Playground worlds Creating and Evaluating Experiences of Role-playing Games*. Vol. 14, pp. 12.
- Hyltoft, M (2014). Full-time edu-larpers. In M. Lunau (ed) *Læringsrollespil undervisningsmetoden på østerskov efterskole i Hobro*.
- Ivanova, B. and Trofimova, N. (2016). Gaming at the Lessons of Foreign Languages. In *The system of higher professional education in russia*. SHARING PRACTICES—THIS IS MY STORY, WHAT IS YOURS International week 10.-14.10. 2016. p.67.
- Johnson, R.B. and Onwuegbuzie A.J. (2004). Mixed Methods Research: A Research Paradigm Whose Time has come. In *Educational researcher*. 33(7), pp. 14-26.
- Kirk, R.E. (2013). *Experimental design: Procedures for the behavioral sciences*. Thousand Oaks, CA: SAGE Publications, Inc.
- Kolb, D.A. (1984). *Experimental learning: experience as the source of learning and development*. Prentice- New Jersey: Hall, Inc. Englewood Cliffs.
- Kolb, A.Y. and Kolb, D.A. (2005). Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education. *Academy of management learning & education*. 4(2), pp. 193-212.
- Kramsch, C. (2017). Applied linguistic theory and second/foreign language education. In S. May and N. V. Deussen-Scholl, eds., *Encyclopedia of Language and Education: Second and Foreign Language Education*. 3rd ed. Cham, Switzerland: Springer.

- Lave, J. and Wenger E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Leppälahti, M. (2009). *Roolipelaaminen: Eläytymistä ja fantasia harrastusverkkoja*, Helsinki: Nuorisotutkimusverkosto/ Nuorisotutkimusseura.
- Lunau, M. (2014). Læringsrollespil på Østerskov efterskole. In M Lunau ed., *Læringsrollespil Undervisningsmetoden på Østerskov Efterskole i Hobro. Østerskov Efterskole, Østerskov, Denmark*.
- Mark, M.M. (2015). Mixed and Multimethods in Predominantly Quantitative Studies, Especially Experiments and Quasi-Experiments. In *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. University Press, Oxford England. [ebook] Available at: [http://www.oxfordhandbooks.com.ezproxy.uef.fi:2048/view/10.1093/oxfordhb/978\\_0199933624.001.0001/oxfordhb-9780199933624-e-15](http://www.oxfordhandbooks.com.ezproxy.uef.fi:2048/view/10.1093/oxfordhb/978_0199933624.001.0001/oxfordhb-9780199933624-e-15) [Accessed 18 S Sep. 2019].
- Martin Lee, A. and McKinney, J. (2013). *Understanding and Applying Research Design*, [ebook] John Wiley & Sons, Incorporated, ProQuest Ebook Central. Available at: <https://ebookcentral-proquest-com.ezproxy.uef.fi:2443/lib/uef-ebooks/detail.action?docID=1120726> [Accessed 18 Sep. 2019].
- Merriam-Webster. (2019). *Merriam-Webster Dictionary; Paradigm*. Available at: <https://www.merriam-webster.com/dictionary/paradigm> [Accessed 18 Sep. 2019].
- Merriam-Webster. (2019). *Merriam-Webster Dictionary; Pragmatism*. Available at: <https://www.merriam-webster.com/dictionary/pragmatism> [Accessed 18 Sep. 2019].
- Mertens, D.M. and Hesse-Biber, S. (2013). Mixed methods and credibility of evidence in evaluation. In D. M. Mertens and S. Hesse-Biber eds., *New Directions for Evaluation*. Somerset: John Wiley & Sons, Incorporated, number 138, pp. 5-13.
- Ministriet for Børn, Undervisning og Ligestilling. (2016). *Engelsk, Fagformål for faget engelsk*. Ministriet for Børn, Undervisning og Ligestilling.
- Morton B. (2007). Larps and their cousins through the ages. In, J Donniss, M Gade & L Thorup eds., *Lifelike*. Arco Grafisk A/S.
- Opetushallitus. (2014). *Perusopetuksen Opetussuunnitelman perusteet 2014*. Helsinki: Opetushallitus.
- Opetushallitus. (2016). *National core curriculum for basic education 201*, Helsinki: Finnish National Board of Education.



- Werquin, P. (2007). *Terms, Concepts, and Models for Analysing the Value of Recognition Programmes*. OECD. Available at: <http://www.oecd.org/education/skills-beyond-school/41834711.pdf> [Accessed 18 Sep. 2019].
- Oxford Dictionaries. (2018) *Oxford dictionaries; Actant*. Available at: <https://en.oxforddictionaries.com/definition/actant> [Accessed 18 Sep. 2019].
- Pettersson, J. (2005). *Roolipelimanifesti*. Helsinki: Like.
- Pihkala-Posti, L. (2015). Pelillisyyttä kielenoppimiseen!. In *Kieli, koulutus ja yhteiskunta*. 6(5).
- Richards, J.C. and Rodgers, T.S. (1986). *Approaches and Methods in Language Teaching A Description and Analysis*. Cambridge University Press. Cambridge, United Kingdom 1986.
- Rubin, H.J. and Rubin I.S. (2005). *Qualitative interviewing the art of hearing data*. 2<sup>nd</sup> ed. Sage publications.
- Ryan, M-L. (1991). *Possible worlds, artificial intelligence, and narrative theory*. Bloomington, Indiana University Press.
- Sarajärvi, A. and Tuomi, J. (2004). *Laadullinen tutkimus ja sisällönanalyysi*. Tammi.
- Syväluoma, M. and Turpeinen A-K. (2003). *Roolipeli oppilaanohjaajan Opetusmenetelmänä – Kehittämistutkimus*, Joensuu: Itä-Suomen yliopisto.
- Vanek, A. and Peterson, A. (2016). Live-action role-playing (LARP): Insight into an underutilized educational tool. In *Learning, Education and Games*. pp. 219-240.
- Wang, J. and Yan, Y. (2012). The Interview Question. In J. F. Gubrium ed., *The SAGE Handbook of Interview Research : The Complexity of the Craft*, 2nd ed. Thousand Oaks, Calif: SAGE Publications, Inc.
- Verhoeven, L. (1997). Sociolinguistics and Education. In Coulmas, F ed., *The handbook of sociolinguistics*. Oxford: Blackwell.
- White, L. (1989). Universal Grammar and Second Language Acquisition, *Language Acquisition & Language Disorders*. [ebook] John Benjamins Publishing Company, Amsterdam. Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=385654&site=ehost-live> [Accessed 3 Sep. 2019].
- Whitton, N. (2009). *Learning with digital games: A practical guide to engaging students in higher education*. New York: Routledge.

Østerskov Efterskole. (2019). *Østerskov in English*. Available at: <http://osterskov.dk/kontakt/in-english/> [Accessed 18 Sep. 2019].

## Appendix:

### Appendix 1.

Interview questions and the framework of the interview:

- 1. Mean of English evaluation throughout their school years. (self-evaluation on scale 1-7)**
  - a. How do you feel about school?
  - b. How do you feel about learning?
  - c. How about teaching in general?
  - d. Describe your ideal class/learning situation?
  - e. What has been your most effective class
  - f. And what has been your least effective.
  - g. How do you feel about the English language?
    - i. As a means of communication
    - ii. As a school subject
  
- 2. English language (informal learning)**
  - a. How much do you use English?
  - b. Where do you use it?
  - c. Is there any specific aspect of English you feel that informal learning helped you?
    - i. In what way
  - d. How /Where do you learn English?
  
- 3. Impact of formal learning**
  - a. In what way do you feel the School teaching has affected you...
    - i. English competence
    - ii. Learning
  - b. Is there any specific aspect of English you feel that formal learning helped you?
    - i. In what way
  - c. What kind of teaching...
    - i. You Like
    - ii. Helps you learn
  - d. 4 options: pic your favorite exercise type:
    - i. Listening
    - ii. Reading
    - iii. Writing (computer, hand)
    - iv. Talking
  
- 4. Impact of edu-LARP**
  - a. Why did you join Østerskov Efterskole?
  - b. How do you feel about the teaching method in Østerskov Efterskole?
    - i. English competence
    - ii. Learning
  - c. Does the teaching method support you... and why or why not?
    - i. English competence

- ii. Learning
    - iii. Is there any specific area that the teaching method helps you?
  - d. How do you feel like your English competence has changed while you have been in Østerskov Efterskole?
    - i. Why?
    - ii. In which specific aspects of language.
  - e. Do you feel that there has been a change in your English competence in the last 6 months?
    - i. What kind of
    - ii. In which aspects of language.