

UNIVERSITY OF TARTU
DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE

USING PEER ASSESSMENT IN BASIC SCHOOL
EFL CLASSES
MA thesis

KRISTIN LAUR
SUPERVISOR: LECT. ÜLLE TÜRK

TARTU
2014

ABSTRACT

The purpose of the study was to determine the efficiency of using peer assessment in basic school EFL classes. The topic was prompted by the fact that the majority of research on using peer assessment in EFL classes has been carried out in higher educational setting; whereas very little data is available about implementing peer assessment with younger learners such as basic school students. Also the Estonian National Curriculum for Basic Schools emphasizes the importance of involving students into peer assessment in order to raise their study motivation and develop their ability to set objectives and analyse their own learning process and behaviour in the light of those objectives. Due to the reasons mention above, the thesis was aimed at answering the following three research questions:

1. What kind of attitudes do the students of form 8 have towards using peer-assessment in the form of marks as well as in the form of comments in EFL classes for assessing their classmates' oral presentations?
2. How do the marks students get for their presentations influence their attitudes towards peer assessment?
3. How well do the scores and marks given by the students correspond to the ones given by the teacher?

The thesis consists of two chapters, in the first of which background information about peer assessment as an assessment method is provided. To be more exact, the definitions and variations of peer assessment are presented; the benefits and weaknesses of implementing peer assessment in EFL classes are discussed and an overview of the guidelines for implementing peer assessment is provided. The second chapter focused on research which was carried out among the basic school students of form 8 in Rakvere Reaalgümnaasium. The sample consisted of 14 students who were asked to peer assess each other's EFL oral presentations. The students were provided some training on assessment after which they gave their peers summative feedback as well as feedback in the form of comments. The data elicitation instruments were questionnaires, which were distributed to the students before and after the peer assessment procedure, as well as peer feedback forms for summative assessment, teacher feedback forms for summative assessment and feedback forms for comments.

The results of the study showed that the basic school students were in general positively disposed to using peer assessment in their EFL classes both before and after putting it into practice as they found it beneficial to them for several reasons. Furthermore, the majority of them considered their peers as well as themselves as capable of assessing the others adequately and objectively. It was also noticed that the students who got the lowest marks from their peers were less positively disposed to peer assessment than those who got the highest marks. Finally, the study also demonstrated that as young students as the 8th-formers are able to assess their peers' EFL oral presentations similarly to the teachers, as the scores and marks given by the sample students corresponded closely to the ones given by the two teachers. However, some signs of friendship marking could be reported, and what concerns peer assessment in the form of comments, more training would have been necessary in order to raise the quality of the feedback. Despite the overall positive attitude of the large majority of the sample students and the fact that they assessed similarly to their teachers, half of them answered that they would still prefer only teacher assessment in their EFL classes.

TABLE OF CONTENTS

ABSTRACT	2
LIST OF ABBREVIATIONS	5
INTRODUCTION	6
1 PEER ASSESSMENT AS AN ASSESSMENT METHOD	17
1.1 Definition and Variations of Peer Assessment	17
1.2 Benefits of Peer Assessment	19
1.2.1 Benefits of Using Peer Assessment in Educational Setting	19
1.2.2 Benefits of Using Peer Assessment in EFL Context	21
1.3 Weaknesses of Peer Assessment – the Question of Reliability and Validity	25
1.3.1 Influence of Social Relations within Peer Group	27
1.3.2 Lack of Skills and Language Proficiency in EFL Context	28
1.3.3 Negative Attitudes of Students	31
1.4 Implementing Peer Assessment	32
1.4.1 Phases of Implementing Peer Assessment	33
1.4.2 Guidelines for Implementing Peer Assessment	35
2 RESEARCH ON USING PEER ASSESSMENT IN BASIC SCHOOL EFL CLASSES	38
2.1 The Aim of the Study and the Research Questions	38
2.2 Method	39
2.2.1 Research Setting and Sample	39
2.2.2 Instruments	39
2.2.3 Procedure	41
2.3 Results	48

2.3.1 Research Question One	49
2.3.2 Research Question Two.....	57
2.3.3 Research Question Three.....	61
2.4 Discussion of the Results.....	65
CONCLUSION	80
LIST OF REFERENCES	85
APPENDIX 1 Pre Peer Assessment Survey.....	89
APPENDIX 2 Post Peer Assessment Survey	91
APPENDIX 3 Training Materials	94
APPENDIX 4 Student Feedback Forms 1-3 for Comments	96
APPENDIX 5 Feedback for Summative Assessment	98
APPENDIX 6 Guidelines for Oral Presentation	99
APPENDIX 7 Rubric for Assessing EFL Oral Presentations	100
APPENDIX 8 Rubric for Assessing EFL Oral Presentations (in Estonian).....	101
APPENDIX 9 Table 13	102
RESÜMEE	103

LIST OF ABBREVIATIONS

EFL – English as a Foreign Language

PA – Peer Assessment

INTRODUCTION

Assessment has an important role to play in the context of education. It is a vital part of the everyday teaching and learning process at school. O'Donnell and Topping (1998: 259) define assessment as “the determination of the amount, level, value or worth of something”. *Longman Dictionary of Language Teaching and Applied Linguistics* offers a more detailed definition about assessment in educational context stating that assessment is “a systematic approach to collecting information and making inferences about the ability of a student or the quality or success of a teaching course on the basis of various sources of evidence” (Richards and Schmidt 2002a: 35). Traditionally, teachers have been the assessors who give feedback about the progress and abilities of the students. Nevertheless, since 1990s one has witnessed the increasing popularity of “alternative assessment” methods which underline the importance of involving students into assessment procedure. This type of assessment methods have been implemented in the context of a number of school subjects including EFL classes. One of these “alternative assessment” methods is peer assessment which is in the focus of the present study. However, before treating the concept of peer assessment in a more detailed way, the possible classifications of different assessment methods are briefly introduced, and their presentation in the Estonian National Curricula is touched upon. This is done with the purpose of showing the position of peer assessment in the context of different assessment methods, as well as with the purpose of establishing the position of peer assessment in the context of Estonian National Curricula.

To begin with, one can distinguish between “traditional” and “alternative assessment”. “Traditional assessment” relies mainly on standardized testing emphasising the importance of objectivity, standardization and reliability (Falchikov 2005: 60). “Alternative assessment”, on the other hand, arose as a reaction to “traditional assessment” in 1990s, and its proponents emphasised the importance of the following aspects:

- assessment should emanate from “normal classroom learning activities” that learners are familiar with;
- it should concentrate on higher level thinking skills;
- it should enable students to produce language in real-world context;
- scores and grades should be given by human beings rather than computers.

(Douglas 2010: 73)

Some of the methods that are considered as alternatives in assessment are as follows: self- and peer assessment, conferences, portfolios, performance assessment (Douglas 2010: 73). Falchikov (2005: 82) points out that the most important difference between “traditional” and “alternative assessment” is related to the role of the students and the extent of power given to them by teachers. In the case of “alternative assessment” students are involved in assessment process and, thus, also gain more control of the process of assessment as well as their own learning (ibid). In EFL context “traditional assessment” is still widely used (e.g. in the form of standardized language tests) but it is combined with different “alternative” assessment methods. As Douglas (2010: 73) points out, nowadays nobody argues that the values which the proponents of “alternative assessment” cherish are not important; vice versa, the methods that were formerly seen as alternatives *to* “traditional assessment” have now become alternatives *in* assessment, “part of the repertoire of techniques available for assessing the language abilities of learners” (Douglas 2010: 73).

Secondly, it is possible to classify assessment methods according to the type of feedback given. Here, one can differentiate between *formative* and *summative assessment*. The aim of formative assessment is to give students feedback on their progress as well as on their strengths and weaknesses of learning (Douglas 2010: 72). According to Butt (2010: 49), the main purpose of formative assessment is “to create a closer link, essentially a *relationship*, between assessment and learning“. Moreover, formative feedback also

provides students with information about how to eliminate the weaknesses and which steps to take in order to guarantee further development (Douglas 2010: 72). This type of feedback is also very important in EFL context in order to assure the improvement of students' language skills.

The aim of summative assessment, on the other hand, is to measure achievement and the results are given in the form of numbers: scores or marks (Douglas 2010: 72). In addition to scores or marks corrective feedback may also be provided but it is not obligatory (*ibid*).

To sum it up, "traditional" assessment methods, which can be characterised by keywords such as objectivity, standardisation and reliability, tend to use summative feedback more than formative one; whereas "alternative" assessment methods, which emphasise the importance of involving students into the assessment procedure, are more closely related to formative feedback. However, this is not a rule: one and the same assessment method (e.g. peer assessment) can be used for both formative and summative purposes. In EFL context these different assessment methods can be successfully combined.

As for the assessment in the context of Estonian educational situation, the different types of assessment, introduced above, have been given a green light by the Estonian National Curricula. Namely, the Estonian National Curriculum for Basic Schools as well as the Estonian National Curriculum for Secondary Schools contain the aspect of using both formative and summative assessment in order to support the students' development and give them feedback about their progress (Ministry of Education and Research 2011a, sections 20, 21; Ministry of Education and Research 2011b, sections 16, 17).

What concerns summative assessment, the Estonian National Curricula state that comparing "the pupil's subject-related knowledge and skills" with the expected learning

outcomes presented in the syllabus serves as the grounds for summarizing grades (ibid a, section 21 (1); ibid b, section 17 (1)). According to the curricula, these knowledge and skills “may be evaluated in the course of the studies as well as at the end of the study topic“ (Ministry of Education and Research 2011a, section 21 (1); ibid b, section 17 (1)). In addition to that, both curricula introduce the five-point grading scale for assessment (Ministry of Education and Research 2011a, section 21 (2), (3); ibid b, section 17 (2), (3)). Differently from upper secondary school, basic schools are allowed to use other marking scales instead of the five-point one; however, it must be made clear in the school curriculum how the grades of the alternative marking system can be converted to the five-point scale (ibid a, section 20 (4)).

Using summative assessment and the five-point grading scale is the most common way of carrying out assessment in the context of different subjects in Estonian comprehensive schools, among others in the context of EFL classes. In addition to that, summative assessment is also used in the case of high-stakes tests like national examinations.

What concerns the presentation of formative assessment in the Estonian National Curricula, the focus lies on providing the student with feedback about his/ her strengths and weaknesses together with suggestions for further steps to be taken in order to support the student’s development (Ministry of Education and Research 2011a, section 20 (1); Ministry of Education and Research 2011b, section 16 (1)).

In connection with formative assessment, the Estonian National Curricula also contain the aspect of using “alternative assessment” methods such as self-assessment and peer assessment. The curricula state that students should be involved into self-assessment as well as peer assessment in order to raise their study motivation and develop their ability to set objectives and analyse their own learning process and behaviour in the context of

those objectives (Ministry of Education and Research 2011a, section 20 (3); Ministry of Education and Research 2011b, section 16 (3)). It is important to note that using formative assessment as well as implementing self-assessment and peer assessment are relatively new phenomena in Estonian educational setting as this type of assessment methods were not incorporated into the previous Estonian National Curriculum which was in force until 2011.

Due to the abovementioned reason, using peer assessment has recently become an issue worth discussing in Estonia. As it was not part of the previous national curriculum, using peer assessment more extensively still seems to be a relatively novel idea for Estonian students as well as for the teachers of different subjects, among others for EFL teachers. Teachers seem to doubt whether students are capable of assessing their peers objectively and adequately, whether the feedback which comes from classmates is taken as seriously as the feedback from the teachers etc. Students seem to have similar concerns: some of them do not trust their classmates as objective assessors, others doubt their own abilities of being skilful assessors. Thus, the issue of peer assessment is topical in Estonia and worth being investigated.

Furthermore, most of the research on peer assessment in EFL context has examined peer assessment in higher education setting but very little research has been done regarding basic school students, leaving the impression that peer assessment in basic school is a secondary matter; which, in its turn, cannot be true, at least not in the context of Estonian educational situation, because, as demonstrated above, peer assessment has been incorporated into the Estonian National Curriculum for Basic Schools. Also Topping (2003: 68, 2009: 24) points out that peer assessment has been in the focus of a number of researches covering a wide range of different subject areas but in higher education settings. Saito and Fujita (2004: 32) echo that peer assessment of various skill areas has been a

widely researched topic in psychology and mainstream education. They add that the results of several studies (Falchikov and Goldfinch, 2000; O'Donnell and Topping, 1998; Topping, 1998) suggest that peer assessment is valid and reliable enough in order to be used as an assessment method at school (ibid). In the context of the current thesis, however, it is important to pay attention to the studies which have focused on peer assessment in EFL context examining the correspondence of teacher and student assessment and/or students' attitudes towards peer assessment on the example of oral presentations. Here, one can notice the same tendency as pointed out above: the large majority of the studies have been carried out in higher education settings. In the following, a short overview of the results of the studies relevant in the context of the present thesis is provided.

To begin with, students' attitudes towards peer assessment on the example of oral presentations in EFL classes in higher education context have been studied by Maiz Arévalo (2008), Grez *et al* (2012) and Peng (2010). In his study Peng (2010) also examined the correspondence of peer and teacher grades. Similarly to Peng, the agreement of teacher and peer assessments was investigated by Patri (2002) and Grez *et al* (2012). The results of these studies revealed that most of the students held an overall positive attitude towards peer assessment (Grez *et al* 2012, Peng 2010, Maiz Arevalo 2008), even if they admitted feeling uncomfortable to assess their classmates (Maiz Arevalo 2008: 132). Moreover, grades given by the peers corresponded to the ones given by their teachers (Grez *et al* 2012, Peng 2010, Maiz Arevalo 2008, Patri 2002, Shimura 2006). However, Grez *et al* (2012) concluded that in spite of the overall positive relationship between the teacher and student scores, peers and teachers still interpret the assessment criteria in a different way. Shimura's (2006: 99) study, in its turn, revealed that the correlation of

students' and teachers' marks was affected by the language level of the students and that the students rate their peers with a much narrower range than the teacher.

Similarly to Shimura, Yinjaroen and Chiramanee (2011) studied the influence of students' EFL proficiency on peer assessment. On the contrary to the findings of the studies described above, Yinjaroen and Chiramanee (2011: 15) did not find significant correlation between peer assessment of all three sample groups with different language level and teacher assessment for the criteria such as content, preparation, expression and vocabulary, and fluency. Significant correlation was found only in the case of the criterion voice and pronunciation (Yinjaroen and Chiramanee 2011: 15). However, it is important to note that the students had not had any training on the assessment of oral presentations beforehand, and the authors assumed that if the students had been given training in assessing oral proficiency, the results would have been different (Yinjaroen and Chiramanee 2011: 15).

The importance of providing assessment training to students has been emphasised by several researchers (De Grez *et al.* 2012: 139, Topping 2009: 25, Greenstein 2010: 117, Falchikov 2005:158). Also Saito (2008) investigated the effects of training on peer assessment in the context of oral presentations in Japanese university EFL classes. Furthermore, in addition to summative peer assessment Saito's study also looked at peer feedback in the form of comments. He concluded that training may enhance students' comments but, surprisingly, it did not have any significant difference to peer ratings.

Similarly to Peng (2010) and Maiz Arevalo (2008), whose findings were introduced above, Cheng and Warren (2005) were interested in students' attitudes but from a different angle. They used peer assessment with both writing tasks and oral presentations and compared the students' attitudes towards assessing the English language proficiency and the other aspects of performance of their peers. The results of the study showed agreement

between student and teacher assessments in both language and nonlanguage related criteria; however, the students tended to mark within a narrower range than their teachers (Cheng and Warren 2005: 109). Furthermore, the students also interpreted oral and written language proficiency differently from the teachers (ibid: 110). The study also revealed that students had less positive attitude towards assessing their peers' language skills than other aspects of performance (Cheng and Warren 2005: 109) but, nevertheless, peer assessment was considered useful by both teachers and students in the terms of developing one's metacognitive skills (Cheng and Warren 2005: 111).

Peer assessment of oral presentations in EFL classes has also been examined in the context of group work. Saito and Fujita (2009) studied the similarities and differences between teacher and peer assessments of EFL group presentations which were carried out by Japanese university students. Similarly to the findings of several other studies mentioned above, the results of Saito's and Fujita's study confirmed that peer and teacher assessments were generally similar (Saito and Fujita 2009: 162). Furthermore, in the context of group work, it was found that most group members were able to estimate each member's contribution to the group project (ibid: 159).

In addition to the studies introduced above, it has also been investigated how peer ratings influence students' attitudes towards peer assessment. This has been done in the context of EFL writing tasks again by Saito and Fujita (2004) who found that the peer feedback received by the students did not have any significant effect on their overall attitude towards peer assessment.

All the above presented studies have been carried out in higher education setting. The current study, on the other hand, focuses on peer assessment in basic school setting. Unfortunately, there is remarkably little data available about the effect of age on peer assessment. Gatfield (1999, cited in Falchikov 2005: 207) investigated the attitudes of

students from different ages, but found no differences between the satisfaction levels of older and younger students who had taken part in group peer assessment (Falchikov 2005: 207).

In EFL context the attitudes of younger students have been examined by Meletiadiou (2012). The study focused on the attitudes of adolescent students towards peer assessment in the context of EFL writing tasks, and it was concluded that when students had been given support and training, their attitudes, which had been negative before the training, turned out to be positive after being involved in peer assessment (Meletiadiou 2012: 240).

Based on the studies presented above, the following conclusions about using peer assessment in the context of EFL oral presentations can be drawn:

- Students have had overall positive attitudes towards using peer assessment in EFL classes (Grez *et al* 2012, Peng 2010, Maiz Arevalo 2008). Even if the students had less positive attitudes towards assessing their peers' language skills, they considered peer assessment to be useful in the terms of developing one's metacognitive skills (Cheng and Warren 2005).
- Peer ratings tend to correspond to teacher ratings (Grez *et al* 2012, Peng 2010, Maiz Arevalo 2008, Patri 2002, Saito and Fujita 2009). However, in spite of the overall correspondence of peer and teacher scores, assessment criteria may still be interpreted differently by students and teachers (Grez *et al* 2012, Cheng and Warren 2005).
- It is advised to provide the students with assessment training before implementing peer assessment (De Grez *et al.* 2012: 139, Topping 2009: 25, Greenstein 2010: 117, Falchikov 2005:158, Yinjaroen and Chiramanee 2011: 15, Meletiadiou 2012). Nevertheless, findings by Saito and Fujita (Saito and Fujita 2009) indicate that

training has a more significant effect on the quality of the students' feedback in the form of comments than on their feedback in the form of grades.

- Students' attitudes towards peer assessment are likely not influenced by the nature of the feedback they receive from their peers (Saito and Fujita 2004).
- Adolescent students' negative attitudes towards peer assessment can be changed by giving them support and training when implementing peer assessment in EFL classes (Meletiadiou 2012).

However, it is once again important to note that all the studies presented above (except Meletiadiou's (2012)) were carried out in higher education settings. As the author of the present thesis has been working mostly with basic school students, using peer assessment with this age group is of main interest. Thus, the novelty of using peer assessment in Estonia as well as the lack of research on using peer assessment with basic school EFL learners have been a stimulus for conducting research for the current thesis. It would be interesting to see whether the results of the present study, which concentrates on basic school students, correspond to the abovepresented results of the studies of peer assessment carried out among university students.

The main aim of the present study is to determine the efficiency of using peer assessment with basic school students in EFL classes. The study focuses more specifically on using peer assessment with the students of form 8 in the context of EFL oral presentations, and is aimed to address the following research questions:

1. What kind of attitudes do the students of form 8 have towards using peer-assessment in the form of marks as well as in the form of comments in EFL classes for assessing their classmates' oral presentations?
2. How do the marks students get for their presentations influence their attitudes towards peer assessment?

3. How well do the scores and marks given by the students correspond to the ones given by the teacher?

The overall structure of the study takes the form of two main chapters. The first chapter provides background information about peer assessment as an assessment method; whereas the second chapter focuses on the empirical part of the thesis presenting the research questions, sample, setting, instruments, procedure and results followed by discussion.

1 PEER ASSESSMENT AS AN ASSESSMENT METHOD

The present chapter gives an overview of a number of aspects related to using peer assessment at school, especially in EFL classes, and is organised into four sections. Section 1.1 touches upon the definitions and variations of peer assessment. Following, section 1.2 focuses on the benefits of using peer assessment in educational context in general as well as in the context of EFL classes. Section 1.3, on the other hand, deals with the weaknesses of peer assessment, more specifically, with the issues related to validity and reliability. Finally, in section 1.4 the phases of implementing peer assessment at school are introduced and suggestions for successful implementation are provided.

1.1 Definition and Variations of Peer Assessment

In order to outline the essence of peer assessment, its definitions and possible variations are introduced below.

According to Topping (2009: 20), peer assessment is “an arrangement for learners to consider and specify the level, value or quality of a product or performance of other equal status learners”. Falchikov and Goldfinch (2000: 287) add that in the case of peer assessment students “are engaging with criteria and standards” which they rely on when making judgement of the works of their peers. The feedback, which the learners provide to each other, may be confirmatory, suggestive, or corrective (Topping 2009: 21). The variety of “products” which can be peer assessed is large; including writing, oral presentations, portfolios, test performance, or other skilled behaviours (ibid.).

Peer assessment has been inspired by several theories like social constructionism (e.g. Vygotsky), androgogy (e.g. Cross), and the theories of active learning (e.g. Piaget) (Falchikov & Goldfinch 2000: 287), and the influences of those different theories have played an important role in the formation of peer assessment into a versatile assessment

method rich in variations. Topping (2003: 65) has listed a number of ways how peer assessment activities can vary:

- peer assessment can be used in a wide range of different subjects and thus, the product or output can vary (e.g. portfolios, oral presentations, writing, test performance etc);
- peer assessment can be formative or summative;
- the organisation of the assessors can be different (e.g. individual assessors, assessors in pairs or groups);
- assessees and assessors may belong to the same or different year of study and be of the same or different ability;
- the directionality can vary as peer assessment can be one-way, mutual or reciprocal;
- place and time can vary as peer assessment can occur formally in class, or informally out of class;
- the objectives of using peer assessment may vary. (Topping 2003: 65)

O'Donnell and Topping (1998: 259) emphasize that peer assessment should not be confused with *peer monitoring*. The latter focuses on learning processes and procedures as students are expected to monitor whether their peers use effective and suitable procedures of learning. Peer assessment, on the other hand, concentrates on products and outcomes of learning which peers are asked to assess (ibid). At the same time, O'Donnell and Topping (1998: 259) point out that in everyday practice peer monitoring and peer assessment are closely linked to each other as effective study behaviours are likely to lead to good products; and constructive and adequate assessment of products is likely to have a positive impact on students' learning behaviours.

To sum it up, although peer assessment is rich in variations, the essence of peer assessment is that learners are the assessors who provide feedback on their peers achievements, skills etc. according to certain agreed criteria or standards.

1.2 Benefits of Peer Assessment

Below a number of benefits of using peer assessment at school are presented. Section 1.2.1 focuses on the general benefits of peer assessment which are relevant in the context of different subjects; whereas section 1.2.2 concentrates more specifically on the benefits of using peer assessment in EFL classes.

1.2.1 Benefits of Using Peer Assessment in Educational Setting

Peer assessment has been promoted by several researchers claiming that it has a number of benefits compared to other forms of assessment (Topping 2003, 2009, Falchikov 2005, Butt 2010, Douglas 2010, Sebba *et al* 2008).

The essential goal of peer assessment is “to provide feedback to learners” (Topping 2009: 22) but the students do not just get feedback, they can also give feedback themselves. When doing the latter, students have to keep in mind the objectives of the work that they assess as well as stick to the settled assessment criteria. Thus, as pointed out by Douglas (2010:73), peer assessment contributes to raising students’ awareness of learning goals and criteria for judging the quality of learning and knowledge, which is definitely one of the most important benefits of peer assessment. Also the review of research evidence of the impact that self- and peer assessment have on secondary students carried out by Sebba *et al* (2008: para.3) reported increased engagement with setting learning goals, clarifying objectives as well as taking responsibility for learning.

Moreover, in addition to developing students' ability to assess the work of others, peer assessment has also a positive impact on students' efforts of self-assessment (Butt 2010: 83). Similarly, Black *et al* (2003:50) state that peer assessment is an important complement to self-assessment. Furthermore, it can even be regarded as "a prior requirement for self-assessment" (Black *et al* 2003: 50) as it helps to develop the skills that are essential for assessing one's own work and knowledge (ibid: 51). Also Maiz Arevalo (2008: 128) concludes that through the process of assessing their peers, students actually learn how to "assess themselves in their future performances". Topping (2009: 21) has nicely formulated the importance of peer assessment by saying that "involvement in peer assessment at school can develop transferable skills for life" as in different times and situations of the everyday life one can experience both roles: the role of an assessor as well as the role of an assessee.

Peer assessment can also give students a better picture of institutional assessment processes (Topping 2003: 68). Falchikov (2005: 151) emphasises that assessment can be seen as an instrument of power and the more learners are involved, "the greater the potential of assessment to improve learning and encourage personal, academic and professional development" (Falchikov 2005: 151).

Moreover, the feeling of being involved and having some power in the process of assessment contributes to increasing the students' self-confidence. The aspect of the positive influence of peer assessment on the students' self-esteem was also brought out by Sebba *et al* (2008: para.3).

Another important aspect is that peer assessment is "plentiful" (Topping 2009: 22). Usually there is only one teacher but a number of students in a classroom and that is why "feedback from peers can be more immediate and individualized than can teacher feedback" (ibid). Cole (1991, cited in Topping 2009: 22) claims that there is a difference in

how students perceive feedback from adults and peers: the first one is considered to be “authoritative but ill-explained”, whereas the latter “gives richer feedback that is open to negotiation” (ibid). Also Black *et al* (2003: 50) echo that students often take the criticism from their peers more seriously than the same remarks from their teacher. Furthermore, Black *et al* (2003: 50) claim that when students comment on each other’s work they use “shared language forms” and can this way “provide tenable models” to those students who have had difficulties with understanding or coping with the task appointed by the teacher.

Finally, peer assessment does not contribute to developing only the students’ meta-cognitive skills but has meta-cognitive benefits for teachers, too. Namely, an intention to use peer assessment with students can work as an impulse for teachers to “scrutinize and clarify the assessment objectives and purposes, criteria and marking scales” (Topping 2003: 68).

1.2.2 Benefits of Using Peer Assessment in EFL Context

The previous section listed a number of benefits of peer assessment which are relevant in the context of different subjects, among others in EFL context. These benefits are all important facets of successful language learning. However, after reading the previous section, one might raise the following question: is there any direct evidence on the improvement of students’ performance in EFL classes due to the implementation of peer assessment? As presented in the Introduction, a number of studies on peer assessment have focused on students’ attitudes and the issues of validity and reliability comparing teacher and student assessments. Significantly fewer studies have focused on investigating the direct effect of peer assessment on students’ EFL performance. Have students’ EFL skills actually improved after they have been involved in peer assessment? As already mentioned, there are not many studies which have concentrated on this aspect of peer

assessment, but, nevertheless, there are some and the results of these studies indicate that one could give a positive answer to the question stated above. The current section gives an overview of the studies which have reported positive effects of peer assessment in EFL context.

To begin with, there are examples of successful implementation of peer assessment in the context of EFL writing tasks. In the case of writing tasks, peer assessment offers several options: it can be used for giving general feedback as well as more detailed feedback about possible improvements; it can concentrate on the whole written piece as well as on the elements of the writing process (e.g. planning, editing) (Topping 2003: 70). Topping (2003: 76) claims that there is substantial evidence on the positive effect of using peer assessment at school with writing tasks. He states that “peer assessment seems to be at least as effective in formative terms as teacher assessment, and sometimes more effective” (Topping 2003: 76).

Topping speaks about writing tasks in general not in relation to EFL classes. Nevertheless, there are studies which indicate that Topping’s statement fits also into the context of EFL writing tasks. Jahin (2012) studied the effect of peer reviewing on EFL essay writing and found that it had a positive effect on the essay writing ability of the students of the experimental group. Furthermore, the study also revealed that the students who were involved in peer reviewing experienced significantly less writing anxiety than the students of the control group (Jahin 2012: 72). Similarly, in their study Birjandi and Siyyari (2010) reported the positive effect of peer assessment on students’ EFL writing skills. They focused on university students’ paragraph writing skills during an EFL advanced writing course of 16 weeks (Birjandi and Siyyari 2010: 8). The participants of the study were divided into the following three groups: the control group, the peer assessment group, and the self assessment group (Birjandi and Siyyari 2010: 8-9). The

comparison of the three groups' writing performance showed that using peer assessment over a period of time was the most effective means for improving the students' writing performance; whereas the participants of the control group showed the least degree of improvement in comparison to the other two groups (Birjandi and Siyyari 2010: 15-16). A similar study to the previous one was carried out by Birjandi and Tamjid (2012). In order to examine the role of self- and peer assessment in promoting writing performance of EFL learners, four experimental groups and one control group were formed. The groups according to the assessment methods used were as follows: 1) journal writing as a self-assessment technique + teacher assessment; 2) self-assessment + teacher assessment; 3) peer assessment + teacher assessment; 4) self-assessment + peer assessment; 5) teacher assessment (Birjandi and Tamjid 2012: 517). The writing performance of the students of all the groups was pre-tested at the beginning of the semester and post-tested at the end of the semester (*ibid.*). The results revealed that the mean difference in the writing scores between pre-test and post-test was maximum for the second group (self-assessment + teacher assessment) and the third group (peer assessment + teacher assessment) (Birjandi and Tamjid 2012: 529). Thus, using self-assessment and peer assessment accompanied by teacher assessment led to the maximum improvement of the students' writing performance compared to the control group, where only teacher assessment was employed, and to the other two experimental groups.

The previous section gave an overview of the studies that have reported the positive effect of peer assessment on improving the learners' EFL writing performance. The current thesis, however, focuses on peer assessment of oral presentations. Unfortunately, in the context of EFL oral presentations, there is not such clear evidence available, as one could see in the context of EFL writing skills, which would prove that using peer assessment leads to the improvement of learners' EFL oral presentations. It is probably due to the fact

that one's EFL oral presentation skills are more difficult to pre-test and post-test than writing skills. Topping (2003: 76) claims that, except writing skills, in other areas such as oral presentations, group skills etc. evidence for the effects of using peer assessment are more dependent on softer data such as students' subjective perceptions. Due to that, any unambiguous conclusions about the direct effect of peer assessment on students' EFL oral presentation skills cannot be drawn, yet.

However, there is some indirect evidence available which indicate that peer assessment is likely to have a positive effect on learners' EFL oral skills. There are studies which have not attempted to "measure" students' EFL oral presentation skills before and after peer assessment but focused on the students' attitudes. According to the result of the studies by Nakamura (2002), Cheng and Warren (2005), and White (2009) the students have found peer assessment to be useful for improving their EFL oral presentation skills. Once again, there is no evidence which prove that the oral presentation skills of these students actually improved, but based on their experience of peer assessment the students claimed that being involved in peer assessment had been beneficial in terms of developing their EFL oral presentation skills. Also Falchikov (2005: 16) states that engaging students in the assessment of oral presentations is highly beneficial as having the responsibility of giving feedback requires more concentration and engagement than just listening (Falchikov 2005: 16).

To sum up the entire section of the benefits of peer assessment, a number of authors (Falchikov 2005, Topping 2003, Topping 2009, Black *et al* 2003, Butt 2010, Maiz Arevalo 2008, Douglas 2010, Sebba *et al* 2008) claim that peer assessment has a variety of advantages which can be summarised as follows:

- raising the students' awareness of learning goals and criteria for judging the quality of learning and knowledge;

- increasing learner responsibility by giving them more power to decide;
- contributing to increasing students' self-esteem;
- peer assessment is more plentiful than teacher assessment;
- students understand their friends' comments better than their teachers' as they use shared language;
- students take the comments from their peers more seriously;
- peer assessment has a positive impact on students' ability of self assessment.

There is also evidence on the positive effect of peer assessment in EFL context, mainly in relation to the improvement of students' EFL writing skills (Jahin 2012, Birjandi and Siyyari 2010, Birjandi and Tamjid 2012). Due to the benefits listed above, EFL teachers and students should be encouraged to incorporate peer assessment into their everyday teaching and learning practice.

1.3 Weaknesses of Peer Assessment – the Question of Reliability and Validity

The issue that the sceptics of peer assessment often point out is related to the validity and reliability of peer assessment. The larger is the extent to which an assessment can provide accurate and consistent measures of the abilities it is intended to measure, the more reliable it is (Douglas 2010: 10). Validity, on the other hand, is related to the inferences one makes on the basis of the results of assessment (ibid). One can speak about valid assessment if it actually measures what it is meant to measure (Richards and Schmidt 2002b: 575). It has been argued whether students are capable enough of assessing their classmates adequately, objectively and consistently. As Falchikov (2005: 190) points out, a number of teachers are concerned about how closely the results of peer assessment

correspond to their own. She (Falchikov 2005: 190) adds that as long as teachers have doubts about reliability or validity of peer assessment, they are unlikely to use it with their students.

The question of validity and reliability of peer assessment has been of the interest of a number of researchers in the context of different subject areas, among others in the context of EFL (see also Introduction), and several studies have been carried out where the correspondence between teacher and student assessment has been examined (Falchikov and Goldfinch, 2000; O'Donnell and Topping, 1998; Topping, 1998; Grez *et al* 2012, Peng 2010, Maiz Arevalo 2008, Patri 2002, Cheng and Warren 2005 etc.) (see also Introduction). Topping (2009: 24) suggests using the term “accuracy of peer assessment” in order to refer to the issue of correspondence between the assessments made by students and the assessments made by teachers who could be defined as professional “external experts”. However, here Topping (2009: 24) also draws one's attention to the fact that in this case it should be assumed that “expert assessments are, themselves, highly reliable and valid”, which in some contexts may actually be questionable. That is why he (Topping 2009: 24) claims it to be questionable whether the studies which focus on the correspondence of teacher and student assessments actually tell us something about “reliability or validity or both or neither”. Peng (2010: 90) explains that peer assessment should not be seen as a substitution for teacher assessment but it should be interpreted as “a supplementary assessment method for involving and empowering students”, and that is why most peer assessment proponents advise using it as a formative assessment tool rather than a summative one, and are sceptic about comparing the scores of teacher and student assessments.

Nevertheless, whether one uses formative or summative peer assessment, it is obvious that there are a number of factors which may influence the reliability and validity

of peer assessment. In the current section these factors have been divided into the following three categories:

- the influence of social relations within a peer group;
- lack of skills and EFL proficiency;
- students' negative attitudes towards peer assessment.

1.3.1 Influence of Social Relations within Peer Group

O'Donnell and Topping (1998: 267) point out that any group of students face social processes which may influence the reliability and validity of peer assessment. Pond *et al* (1995 cited in O'Donnell and Topping 1998: 266-267) list three possible ways of how social processes within a group can influence students' assessments and have, thus, a negative effect on the reliability and validity of peer assessment. Namely, it is possible to distinguish between friendship grading, collusive grading and decibel grading.

Friendship grading takes place when students assign higher marks to their friends than to other peers. The effect of collusive grading is lack of differentiation between peers and it is likely to occur in the case of high stakes assessment. Decibel grading means that the highest marks are given to the most active students.

In addition to the above listed three possible types of grading, Falchikov (2005: 154) points out that in rare occasions students may also assign lower marks to their peers than the teacher. According to the study by Tsai *et al* (Tsai *et al* 2002: 249, cited in Falchikov 2005: 154) the reason for doing this is the desire to leave an impression that one's own achievements are at a higher level than his/ her peers'.

Moreover, according to Falchikov (2005: 156) some studies (Lin *et al* 2001, Purchase 2000, Beaman 1998) have reported retaliatory behaviour of students. Namely,

there are students who are afraid of reprisals from their classmates' to whom they have assigned lower grades. Sometimes this fear seems to be justified as it has been reported that there are students who actually reduce the marks they have given to the others after they themselves have been assigned a lower grade than they expected (Lin *et al* 2001, cited in Falchikov 2005: 156).

As demonstrated in the present section, there are a number of ways how social relations within a peer group can become an obstacle for successful peer assessment. If the above presented types of grading occur within a group, the results of peer assessment can be considered neither valid nor reliable.

1.3.2 Lack of Skills and Language Proficiency in EFL Context

Another possible problematic issue related to peer assessment is the lack of knowledge or experience of the students to assess each other.

To begin with, peer assessment is challenging for the students in terms of its cognitive and meta-cognitive demands. Topping (2009: 25) demonstrates the complexity of peer assessment by pointing out that in order to give adequate feedback the assessor must have gained the understanding of the goals of the task as well as the criteria for success; and based on these goals and criteria, must be able to make judgements of the performance or product. Thus, peer assessment is cognitively demanding, but on the other hand, Falchikov (2005: 156) argues that "in this respect, students may be no different from a beginner teacher who also lacks experience"; both can be taught and trained to become more knowledgeable and experienced.

However, in the context of EFL classes students' language skills are also likely to influence the efficiency of peer assessment. One might assume that due to the lack of the

English language proficiency some students are unable to assess their peers adequately as they cannot identify the mistakes or weaknesses of peers' writings, presentations etc. As for giving peer feedback in the form of comments and suggestions in EFL classes, the feedback that peers are expected to give to each other may remain too simplistic as the students are not able to express more complex ideas in English and in this case the feedback would not fulfil its purpose. One possible solution here would be that the students are allowed to use their mother tongue in order to be able to provide more constructive and detailed feedback.

The results of several studies confirm the idea that the lack of foreign language proficiency can be an obstacle for successful peer assessment in foreign language classes. For example, in the case of peer reviews in foreign language, it has been claimed that because of their low proficiency of the foreign language students are not able to review their peers' writings appropriately which, in its turn, results in the distrust of their peers' reviews (Nelson and Carson 1998, cited in Shimura 2006: 100). Even the results of Cheng and Warren's (2005: 109) research, which actually confirmed the overall agreement between student and teacher assessments (see also Introduction), revealed that most of the students did not regard themselves as qualified enough to assess their peers' EFL proficiency due to their own insufficient EFL competence as well as due to the uncertainty of what constituted proficiency in EFL.

Furthermore, in the context of foreign language classes, it has also been noticed that students often tend to concentrate on looking for mechanical mistakes in their peers' works and forget to pay attention to the other aspects such as content or organisation (Sengupta 1998, cited in Shimura 2006: 100).

Coming back to the influence of foreign language proficiency on peer assessment, presenting the results of Shimura's (2006) study (see also Introduction) in a more detailed

way is to the point. Namely, Shimura (2006) compared the agreement of teacher and peer assessment amongst Advanced, Upper Intermediate and Lower Intermediate students in Japanese university EFL classes in the context of oral presentations. Above, it was discussed that students with lower proficiency of the foreign language are unlikely to succeed in assessing their classmates appropriately. In the light of the previous discussion, it would be logical to assume that the results of Shimura's (2006) study indicated that the assessment by the Lower Intermediate EFL students corresponded the least to teacher assessment; whereas the assessment given by the Advanced EFL students corresponded the most. However, this was not the case. Surprisingly, the study revealed that the ratings of Upper Intermediate students correlated most closely with the ones given by their teacher, the ratings of Lower Intermediate the next, and Advanced the least (Shimura 2006: 104). Both the Lower Intermediate and the Advanced students over-marked their peers compared to the teacher ratings (ibid: 105). Shimura (2006: 105) suggested that Advanced students overmarked due to the fact that they were confident about the high language level of their peers and, thus, did not assess them severely enough. Based on the results of the study Shimura (2006: 106) also concluded that in the case of oral presentations Lower Intermediate students tend to focus more on assessing the aspects such as voice, gesturer, eye contact and visuals rather than language-related aspects. It was assumed that as their language proficiency gets higher, this tendency declines and they begin to assess more like the instructor as could be seen in the group of Upper Intermediate EFL students (ibid). However, as their proficiency gets as high as a TOEFL score of 550 (the language level of the Advanced group), they stop assessing and differentiating their peers as they once did (ibid).

Thus, the results of Shimura's (2006) study indicate that the level of EFL proficiency of the students has an effect on peer assessment, and Lower Intermediate

students do have difficulties with assessing their peers' EFL skills. Nevertheless, the results also revealed that high EFL proficiency *per se* cannot guarantee valid and reliable peer assessment.

In conclusion, obtaining valid and reliable peer assessments in school lessons may fail due to the lack of assessment experience and skills of the assessors. In addition to that, in EFL classes the validity and reliability of peer assessment can be also influenced by the EFL proficiency of the students. There are studies (e.g. Nelson and Carson 1998, Sengupta 1998, Shimura 2006), which have reported the negative influence of low EFL proficiency to valid and reliable peer assessment. However, it was demonstrated by Shimura (2006) that high EFL proficiency *per se* may not lead to adequate peer assessment.

1.3.3 Negative Attitudes of Students

The third factor which can influence successful implementation of peer assessment is related to students' attitudes and feelings towards this type of assessment. If students are negatively disposed to peer assessment, one can hardly expect it to be effective and fulfil its purpose; and, thus, one cannot speak about valid and reliable results either.

Falchikov (2005: 153-154) presents some possible negative attitudes from the students that teachers can face when implementing peer assessment. Accordingly, students may claim that it is not their job to carry out assessment as teachers are paid to do that; or they may feel that peer assessment is just for saving teacher's time (*ibid*). Students may also be reluctant to participate in peer assessment as they find it stressful to assess their peers (Falchikov 2005: 161).

Falchikov (2005: 168-169) suggests some steps to be taken in order to cope with the negative attitudes of the students. Most importantly, peer assessment activities have to be carefully prepared and executed by teachers. Enough time should be spent on explaining

the benefits of peer assessment and providing evidence of the positive influence. The roles and responsibilities of teachers and students should be made explicit. (Falchikov 2005: 168) To reduce the anxiety and stress level of the students, teachers should ensure that students “gain satisfaction from the increased responsibility and power peer assessment affords them” (Falchikov 2005: 169). In addition to that, it is the teachers’ responsibility to praise students for being honest when assessing their peers (Falchikov 2005: 169). In the next section more steps to be taken for successful implementation of peer assessment are presented in more detail.

In conclusion, the question of validity and reliability of peer assessment has raised concerns among teachers and students as the validity and reliability of peer assessment may be influenced by social relations, lack of skills, and negative attitudes.

1.4 Implementing Peer Assessment

As demonstrated in section 1.2, peer assessment has several benefits. However, involving students successfully in peer assessment is not an easy task and there are a number of factors which may decrease the efficiency of peer assessment.

In order to facilitate good quality of peer assessment, teachers have to take time for “organisation, training and monitoring” (Topping 2003: 68). Also Saito (2008: 554) emphasises the importance of training saying that „the benefits which peer assessment may bring into a language classroom cannot be guaranteed unless students are capable of implementing the assessment“. In order to help teachers with planning and organising peer assessment, several authors (Topping 2009: 25-26, Falchikov and Goldfinch 2000: 317-318, Falchikov 2005: 125, Weeden *et al* 2002, cited in Butt 2010: 81-82) have outlined the guidelines and the most important issues to be considered. This is what the current section focuses on.

1.4.1 Phases of Implementing Peer Assessment

As already demonstrated above, peer assessment is a complex assessment method which is related to a number of different procedures (e.g. training, discussing assessment criteria, giving feedback etc.). Falchikov (2005: 125) has suggested a generic pattern of how to carry out peer assessment. She has organised the different procedures which belong to the process of peer assessment into the following eight phases which all together form a cycle:

- Dissemination (staff development conferences, workshops etc.);
- Preparation and training of students (benefits, mechanisms, pre-measures e. g. attitudes);
- Decisions about criteria (teacher-student negotiation/ discussion);
- Methods of measurement (checklists, rating scales, model answers etc.);
- Implementation (feedback, formative/ summative assessment);
- Evaluation (by students as well as teachers);
- Outcomes (the correspondence of teacher and student marks, effects on learning, effects of practice, benefits to teachers/ students, problems, post-measures e.g. attitudes);
- Improvements and modifications to schemes (learning from experience, application of theory/ psychological principles etc.). (Falchikov 2005: 125)

Peng (2009: 61) modified Falchickv's cyclic scheme for his study and divided it into pre-implementation, implementation and post-implementation phase which contain the following procedures:

- **Pre-implementation phase**
 - preparation and technology;
 - student training;

- discussing assessment criteria with students;
- deciding about measurement methods with students (ibid: 62-63).
- **Implementation phase**
 - implementation
 - demonstrating how to carry out peer assessment;
 - monitoring the peer assessment process and making adjustments when necessary;
 - discussing concerns that emerge from the peer assessment process and providing solutions;
 - monitoring the quality of peer feedback.
 - evaluation
 - students' work is assessed by students and/or teachers
 - enabling peer discussion before evaluating the work of peers;
 - emphasising the importance of giving constructive feedback. (Peng 2009: 63).
- **Post-implementation phase**
 - outcomes and investigations;
 - identifying problems and making improvements and modifications for better implementation in the future. (Peng 2009: 64)

The phases presented above demonstrate in which order to organise and carry out the different procedures of peer assessment and may, thus, be useful for teachers to follow.

1.4.2 Guidelines for Implementing Peer Assessment

Topping (2009), Falchikov and Goldfinch (2000), Greenstein (2010) and Weeden *et al* (2002, cited in Butt 2010: 81-82) have outlined more detailed suggestions which nicely complement the cycle by Falchikov as well as the phases suggested by Peng.

To begin with, teachers should first of all identify their purpose for using peer assessment (Greenstein 2010:116). For example, the goal of using peer assessment may be to encourage students to reflect on what their peers have learned; to develop the students' ability to assess their peers objectively; to promote groupwork etc. By identifying the purpose for peer assessment, teachers can choose a suitable strategy for implementation (*ibid*).

Secondly, Maiz Arevalo (2008: 129) emphasizes the importance of secure classroom atmosphere. It is not advised to carry out peer assessment before the students are familiar and comfortable with their classmates as well as the teacher. Only then the atmosphere of the classroom is supportive enough for developing cooperation. Also Topping (2009: 25) points at the importance of building a comfortable and familiar atmosphere when he suggests teachers to introduce the idea of using peer assessment for a specific task or product very early, and then, over time, steadily ask for and take into consideration the students' advice in order to, finally, get their approval of the plan.

As for the phase of making decisions about criteria, it is very important to involve the students into the process of developing the criteria for assessment (Topping 2009: 25; Falchikov 2000: 317; Falchikov 2005: 148; Greenstein 2010: 116) as it enables to create the feeling of ownership which, in its turn, should decrease the students' anxiety (Topping 2009: 25). The students will probably not suggest anything completely different from their teacher's ideas of the criteria; nevertheless, the draft criteria suggested by the teacher should lead to some student initiated modifications (*ibid*).

In the training phase, students should be explained about the essence of the roles of assessors and assessees. They should know what kind of behaviour is expected from them (Topping 2009: 25). It is the teacher's task to show the students how to do peer assessment. Topping (ibid) suggests using a role play between two adults in order to demonstrate how peer assessment should be carried out, after which the students can themselves "practise peer assessment on a very short task selected for the purpose". While the students practise, the teacher moves around in the classroom monitoring the activity, giving feedback and helping when necessary. Finally, the students should be provided with short written or pictorial guidelines or reminders about the whole process (Topping 2009: 25).

Greenstein (2010: 117) emphasises the importance of teaching students that the feedback they give must be constructive, i.e. all their evaluative comments must be supported by evidence. She (Greenstein 2010: 117) suggests teachers to give examples of constructive feedback such as: "The layout of your poster helped me understand the sequence of photosynthesis," instead of "I like your poster".

What concerns the method of measurement, Falchikov and Goldfinch (2000: 317) found that in the case of summative assessment student assessors should not be expected to rate many individual dimensions, instead it would be better to use an overall global mark with well understood criteria. Also Peng (2009: 151) suggests using only a few assessment criteria.

During the next phase, the implementation phase, teachers are advised to monitor the process but keep a low profile, and give further instructions and feedback when necessary (Topping 2009: 26).

The seventh phase by Falchikov (2005: 125) is related to the outcomes of peer assessment. What to do with the feedback students have got from their peers? As Greenstein (2010: 126) has put it: "feedback from peers /.../ should show students their

weak areas and the steps they can take to progress towards goals.” It is teachers’ task to help students make use of the data (e.g. help them to make a learning plan, suggest alternative resources, approaches etc.) (ibid).

Both Falchikov (2005: 125) and Topping (2009: 26) suggest teachers to compare a portion of peer assessment against their own assessment of the product or task. In addition to that Topping (ibid) recommends teachers to monitor consistently the correspondence of various peer assessments (if more than one peer assesses the same piece of work). It is important that teachers would not assume that their assessments are more reliable than the peers’ (Topping 2009: 26).

In order to improve the quality of student feedback, teachers should provide students with information about their observations of their performance as peer assessors, as well as with information about the quality and reliability of their assessments (Topping 2009: 26). Only this way can students become more successful peer assessors and give feedback of higher quality.

In conclusion, as pointed out at the beginning of the section, peer assessment is a complex assessment method which consists of a number of phases and procedures, but the more students can practise peer assessment, the more skilful assessors they become, and this way the procedure of carrying out peer assessment with all its phases becomes smoother and less time-consuming. Furthermore, the quality, validity and reliability of peer assessment are also likely to improve.

2 RESEARCH ON USING PEER ASSESSMENT IN BASIC SCHOOL EFL CLASSES

The second chapter focuses on the research on peer assessment conducted in the scope of the present thesis. In the following sections the research questions, an overview of the method, the results of the study as well as the discussion of the results are provided.

2.1 The Aim of the Study and the Research Questions

The aim of the present study is to determine the efficiency of using peer assessment with basic school students in EFL classes. The thesis is aimed at answering the following research questions:

1. What kind of attitudes do the students of form 8 have towards using peer-assessment in the form of marks as well as in the form of comments in EFL classes for assessing their classmates' oral presentations?
2. How do the marks students get for their presentations influence their attitudes towards peer assessment?
3. How well do the scores and marks given by the students correspond to the ones given by the teacher?

It is expected that even if the students have negative attitudes towards peer assessment beforehand, then after being provided with training and an opportunity to put it into practice, they will be favourably inclined. Furthermore, it is assumed that getting lower marks from their peers do not foster negative attitudes towards peer assessment. Finally, it is hoped to find correspondence between summative peer assessments and summative teacher assessments.

2.2 Method

2.2.1 Research Setting and Sample

The research was carried out in Rakvere Reaalgümnaasium in form 8. The sample consisted of 14 students (6 girls and 8 boys) who belonged into the same English language learning group which had not been formed by the language ability of the students but by simply dividing the whole class into two halves. Thus, it was a mixed-ability language group with the average mark for the study year 2012/13 being 4.1.

The author of the study was familiar to the students as she had been occasionally teaching them in forms 3-4. During all the phases of preparing and carrying out peer assessment, a number of procedures (e.g. explaining the task, training) were carried out by the author of the study in the presence of the current English teacher of the sample group. Together with the author of the study, she took part in negotiating assessment criteria and assessing the students' presentations.

2.2.2 Instruments

The data for the empirical part of the thesis was collected from the responses of two questionnaires (Appendices 1 and 2). One was distributed to the students before the peer assessment procedure and the other one after the students had tried peer assessment and analysed the marks and feedback given to them by their peers.

As for the content of the two questionnaires, in order to compare students' attitudes before and after trying peer assessment, it was almost the same with small differences. Both questionnaires began with the section of personal data (age, form, sex) but in the second questionnaire the students were also asked to put down the average marks they were given by their peers for their presentations. This was aimed at finding out about the possible connections between the attitudes and marks.

Both questionnaires contained 13 Likert items. A number of Likert items were adapted from Peng's (2009) survey and their wording was changed in order to make them more appropriate for basic school students. In addition to that, some open questions of Patri's (2002) evaluation form were turned into Likert items for the present survey. The statements were the same in both questionnaires, only in some cases the tense was different (i.e. in the second questionnaire past tenses were used).

The first four statements of the questionnaires were aimed at finding out whether, in the opinion of the students, peer assessment helped to develop their learning skills, analytical skills and EFL skills. Statements 5 and 6 were related to students' motivation and aimed at finding out whether peer assessment motivated them to work harder on their own presentations as well as concentrate more carefully on the presentations of their peers. Statements 7-13 focused on the issues related to the assessment process: the objectivity of assessment, the usefulness of the feedback, the capability of the students of commenting on and marking their peers' presentations.

In addition to the section of personal data and the block of Likert items, both questionnaires contained some open and closed questions. In the pre peer assessment survey these questions were aimed at finding out about the students' previous experience of peer assessment; whereas the last questions of the post peer assessment survey focused on the students' preferences of different assessment methods and the reasons behind their preferences.

Both questionnaires were translated into Estonian in order to ensure that the students understand the Likert items and questions correctly. Furthermore, before the students started to fill in the questionnaires, the items were reviewed together with the author of the study and they had chance to ask if there was something incomprehensible or confusing. Most of the items and questions seemed to be understandable for the students,

an exception was item 4. Some of the students could not understand how peer assessment can be related to understanding teacher's requirements; and, thus, it was explained that if they assess their peers using the same criteria and standards as their teacher usually does, then maybe next time, when their teacher gives a similar task, they will know better what the teacher expects them to do, what kind of aspects should be kept in mind, what will be assessed etc.

Besides the questionnaires, data was also collected analysing the Student Feedback Forms. There were different types of feedback forms – one for giving feedback in the form of points (Appendix 5) and others for giving feedback in the form of comments (Appendix 4). As for the feedback forms for comments, there were three different versions which focused on various aspects to be assessed.

Also the English teacher of the group as well as the author of the study gave points and marks to the students using the same feedback form as the students did for summative assessment. These forms were used in order to compare the correspondence of the teachers' and students' marks.

2.2.3 Procedure

Following the example of Peng (2009: 62) the procedures of collecting and analysing data for the current research were organised into pre-implementation, implementation and post-implementation phase. However, due to the nature of the present study, the content of these phases is slightly different from Peng's. The phases with their corresponding procedures are presented in the table below.

Table 1. Phases and Procedures of Implementing Peer Assessment

PRE-IMPLEMENTATION PHASE
<ul style="list-style-type: none"> • Deciding on how and what type of peer assessment to carry out with the students of form 8. • Deciding on the topic and guidelines of the oral presentation together with the English teacher of the students of the sample. • Preparing training materials. • Informing students about the task of oral presentation. • Explaining what peer assessment is. • Revising the guidelines for making successful oral presentation. • Discussion. • Students fill in the first questionnaire about their attitudes towards peer assessment. • Negotiating and settling the assessment criteria together with the teacher and the students. • Student training. • Discussing the students' problems and concerns about peer assessment.
IMPLEMENTATION PHASE
<ul style="list-style-type: none"> • Students give their oral presentations. • Half of the group assesses the presentations by giving grades, whereas the other half provides written comments. When half of the presentations have been delivered, students swap the assessment methods. This way all the students can use both types of assessment as well as get both types of feedback. In the case of written comments each assessor focuses only on two or three aspects of the presentation. • The students are given grades by their teacher and the author of the study.
POST-IMPLEMENTATION PHASE
<ul style="list-style-type: none"> • Students' and the teachers' grades are compared. • The teacher and the author give feedback about their observations of the process. • Students are informed about the results of peer assessment and teacher assessment. • Students fill in the second questionnaire. • The data of both questionnaires are compared and analysed. • The author of the study informs the students about the results of the whole study.

In terms of 45-minute-long English classes, during which a number of the above presented procedures were carried out, the following schedule was followed:

Table 2. Procedures carried out in 45-minute-long EFL classes.

Lesson 1	<ul style="list-style-type: none"> • Students are informed that in a week they have to deliver 2-3-minute-long oral presentations about their dream travel destination which they are asked to peer assess. • It is explained what peer assessment is and what the role of an assessor is (Appendix 3). • Discussion – positive and negative aspects of peer assessment are discussed; students are encouraged to ask questions and share their views about peer assessment. • Guidelines for making a successful oral presentation are revised (e.g. structure, language, voice, body language). (Appendix 6) • Students fill in the first questionnaire about their attitudes towards peer assessment.
Lesson 2	<ul style="list-style-type: none"> • Together with the students assessment criteria are negotiated and agreed upon using the tool of http://rubistar.4teachers.org.
Lesson 3	<ul style="list-style-type: none"> • In groups analysing the assessment criteria settled in the previous class. • Student training – demonstrating how to give constructive feedback (Appendix 3). • Students are asked to assess two oral presentations shown from the videos. The first video is assessed by providing written comments, whereas the second one is assessed by giving marks. • Discussing the students' problems and concerns about peer assessment.
Lesson 4	<ul style="list-style-type: none"> • 7 students give their oral presentations. • In the case of each presentation, 7 students give summative feedback, whereas 6 students give written comments. • The EFL teacher of the group and the author of the study carry out summative assessment.
Lesson 5	<ul style="list-style-type: none"> • 7 students give their oral presentations.

	<ul style="list-style-type: none"> • In the case of each presentation 7 students give summative feedback, whereas 6 students give written comments. • The EFL teacher of the group and the author of the study carry out summative assessment.
Lesson 6	<ul style="list-style-type: none"> • The author of the study and the English teacher of the sample give feedback about their observations of the assessment process. • Each student is provided with all the feedback forms (points as well as comments) concerning his /her presentation. • The author of the study informs the students about the results of peer assessment and teacher assessment. • Students fill in the second questionnaire. • Students share their views and opinions about peer assessment and presentations.
When the study is completed	<ul style="list-style-type: none"> • The students and the teacher are informed about the results of the whole study.

The students had three English classes a week; thus, it took two weeks to carry out the whole procedure of peer assessment.

As can be seen from the table above, the first lesson served as an introduction to what was going to take place in the following classes. As suggested by Topping (Topping 2009: 25) (see also section 1.4.2) the basic aspects of peer assessment, such as the definition, different types, the aim of assessment, the role of an assessor, were introduced to the students. It was explained what is expected from them in the role of assessors as well as in the role of students delivering their oral presentations.

In the next lesson, a great effort was made in order to settle assessment criteria for oral presentations. Following the advice of several assessment experts (Topping 2009: 25; Falchikov 2000: 317; Falchikov 2005: 148; Greenstein 2010: 116) (see also section 1.4.2), the students were involved into compiling a suitable rubric. Of course, it is not feasible for the 8th-graders new to peer assessment to come up with the whole descriptions of the

different performance levels for each performance criteria; but what they could do, was to suggest what kind of aspects should the rubric contain; what should be assessed in the case of an oral presentation. Based on the experience of Falchikov and Goldfinch (2000: 317), and Peng (2009: 151), the English teacher of the group and the author of the study suggested the students choose maximum 5 different aspects in order not to make the assessment procedure too complicated for them. However, agreeing upon only five aspects turned out to be impossible and after a heated discussion the final version of the rubric contained seven different aspects to be assessed (Appendix 7). The author of the study had several ready-made rubrics as drafts from where the descriptions of the performance levels for the chosen aspects were introduced to the students, discussed with them and then adapted. Taking into consideration the students' suggestions, the final version of the rubric was compiled by the author of the study and the English teacher of the sample group. It was delivered to the students in the next lesson with an Estonian translation on the other side (Appendix 8). In the groups of four the students were asked to underline the key words of each descriptor in order to make the rubric more familiar to them. At the end of the second lesson, it was also agreed that the final mark the students get for their presentations would be the average of the following three marks: the average mark given by the peers, the mark given by their English teacher, and the mark given by the author of the study. It might seem somewhat unfair, but it was the proposal of the students to have three distinct marks – two from the teachers and one from the peers.

After working with the criteria and rubric, it was time for training. Although the results of the first questionnaire revealed that all the students had had previous experience with peer assessment in various subjects, they had not tried it in their EFL classes, and thus the importance of training could not be underestimated. To begin with, following the suggestions of Greenstein (2010: 117) (see also section 1.4.2) the students were taught how

to give constructive feedback (Appendix 3), and only then they finally got the chance to put all their knowledge on assessment into practice. To be more exact, the 8th-formers were shown two videos of EFL students' oral presentations, the first of which they had to assess by giving written comments and the second of which by giving points. It is important to note that during this lesson the sample group ran a little bit out of time which means that by the time, the students compared the points they had given to the presentation seen from the video, the lesson had already ended and they did it during the break. Due to that, it is probable that some of the students did not concentrate on their task as carefully as they would have done during the lesson. For the same reason, also the part of discussing students' thoughts and concerns related to peer assessment turned out to be less thorough than planned beforehand.

As one can notice from the tables above, students had an opportunity to give their peers summative feedback as well as feedback in the form of written comments. In order to make that possible, students were divided into two groups of seven. During the 4th lesson the students of the first group gave their presentations and carried out assessment in the form of written comments; whereas the students of the second group were engaged in summative assessment. In the 5th lesson the students of the second group gave their presentations and now it was their turn to provide peer feedback in the form of written comments; whereas the students of the first group could try summative assessment. This way each student got points and marks from seven peers and comments from six peers. It was decided that seven students are engaged into summative assessment rather than six, in order to have more data for the statistical analysis of the scores.

As for the summative peer assessment, in the case of each presentation the students assessed all the seven aspects (Appendix 5); but what concerns providing written comments, the students were not asked to comment on all the seven aspects at the same

time. Instead, during each presentation various students focused only on 2-3 different aspects at a time. There were three different types of feedback forms (Appendix 4) for providing written comments which focused on the following aspects:

- Student Feedback Form 1 – content; organization; visual props
- Student Feedback Form 2 – grammar, vocabulary and sentence structure; body language
- Student Feedback Form 3 – pronunciation; voice control and fluency.

It was taken care of that in the case of each presentation two students filled in Student Feedback Form 1, other two Student Feedback Form 2, and the final two Student Feedback Form 3. During the next presentation each student got a different feedback form than his/her previous one. This way it was guaranteed that by the end of the lesson each student had had an opportunity to comment on each aspect twice. The procedure of providing different types of peer feedback, as well as the distribution order of different student feedback forms for comments is demonstrated in Table 3 below.

Table 3. The division of the groups giving different types of assessment and the distribution order of different feedback forms for comments.

Lesson 4		The distribution order of different student feedback forms (1-3)	Lesson 5	The distribution order of different student feedback forms
Give feedback in the form of comments. Give presentations	Student A	x 2 1 3 2 1 3	Student H	x 2 1 3 2 1 3
	Student B	1 x 3 2 1 3 2	Student I	1 x 3 2 1 3 2
	Student C	2 3 x 1 2 3 1	Student J	2 3 x 1 2 3 1
	Student D	3 1 2 x 3 1 2	Student K	3 1 2 x 3 1 2
	Student E	1 2 3 1 x 2 3	Student L	1 2 3 1 x 2 3
	Student F	2 3 1 2 3 x 1	Student M	2 3 1 2 3 x 1
	Student G	3 1 2 3 1 2 x	Student N	3 1 2 3 1 2 x
Give summative feedback.	Student H Student I Student J Student K Student L Student M Student N		Student A Student B Student C Student D Student E Student F Student G	

The students were allowed to use the Estonian language, in case they run short of vocabulary in English or find it difficult to express an idea in English. After each performer, the students were given some time to put down their comments.

At the beginning of the 6th lesson, each student got into his/her possession all the peer feedback forms concerning his/her presentation. The students were given time to read the comments and have a look at the points and marks given to them by their peers. Next, each student got a personal feedback sheet which contained the following information:

- the mean score given by the peers for his/her presentation and the mark;
- the number of points and the mark given by their English teacher;
- the number of points and the mark given by the author of the study;
- the final mark (the average of the three marks listed above).

When the students had familiarised themselves with the results listed above, the author of the study also informed them about the level of correspondence of teachers' and students' scores and marks in general. Finally, the students filled in the second questionnaire where they could express their views on using peer assessment in EFL classes, now that they had had a real chance to try it out and experience it themselves.

2.3 Results

The present section gives an overview of the results of the two questionnaires (pre peer assessment survey and post peer assessment survey) as well as the results of the analysis of the student feedback forms and of the comparison of teachers' and students' scores and marks.

2.3.1 Research Question One

What kind of attitudes do the students of form 8 have towards using peer-assessment in the form of marks as well as in the form of comments in EFL classes for assessing their classmates' oral presentations?

As explained above, in order to investigate the students' attitudes towards peer assessment the sample group was asked to fill in a questionnaire before experiencing peer assessment in their EFL classes and another questionnaire after they had finished with peer assessment and familiarised themselves with the results of peer assessment.

Both questionnaires contained 13 similar Likert items as well some open and closed questions (Appendices 1 and 2). The comparison of the results of the answers to all the Likert items before and after peer assessment is given in Table 13 in Appendix 9. In the present section the results of the Likert items are demonstrated in the sets of two to four items.

Table 4. The comparison of the mean scores of Likert items 1-4 in pre- and post peer assessment survey.

Item	1 Strongly disagree % (no.of students)	2 Disagree % (no.of students)	3 Neutral % (no.of students)	4 Agree % (no.of students)	5 Strongly agree % (no.of students)	Mean	Paired t-test	
							T	p
1 pre	0	14.3% (2)	21.4% (3)	64.3% (9)	0	3.5	2.2804	0.0401 Sig. dif.
post	0	0	21.4% (3)	78.6% (11)	0	3.79		
2 pre	0	0	14.3% (2)	64.3% (9)	21.4% (3)	4.1	1.4720	0.1648 Not sig.
post	0	0	21.4 (3)	64.3% (9)	14.3% (2)	3.93		
3 pre	0	0	35.7% (5)	35.7% (5)	28.6% (4)	3.9	0.3661	0.7202 Not sig.
post	0	0	21.4% (3)	71.5% (10)	7.1% (1)	3.86		
4 pre	0	0	14.3% (2)	57.1% (8)	28.6% (4)	4.1	3.6056	0.0032 Sig. dif.
post	0	0	50% (7)	35.7% (5)	14.3% (2)	3.64		

To begin with, the statements 1-4 focused on the influence of peer assessment on developing the students' learning skills, analytical skills as well as their EFL skills. As demonstrated in Table 4 above, the mean scores, together with the standard deviations, of all the first four items of the pre-peer assessment survey indicate that most of the students agreed that peer assessment has a positive influence on their EFL skills, their skills of making oral presentations, their skills of analysing the strengths and weaknesses of the presentations delivered by others, as well as on their ability of understanding teacher's requirements better. In the case of all the items 1-4, there were some students who marked their answer as neutral, whereas the majority of the students either agreed or strongly agreed to the statements given. Only in the case of the first item, there were 2 students (14.3%) who disagreed with the statement that peer assessment is helpful in order to understand how to make a good oral presentation.

As for the items 1-4 in the post peer assessment survey, none of the students disagreed to the statements anymore, being, thus, either neutral, agreeing or strongly agreeing. In order to find out whether the differences between the mean scores of the Likert items were significant or not, relevant t-tests were carried out. The t-tests revealed that, although the majority of the students responded positively to the statements 1-4 in the context of both pre- and post peer assessment surveys, in the case of items 1 and 4, there was a significant difference in the attitudes of the sample students before and after trying out peer assessment in their EFL classes. What concerns item 1, the results of the t-test ($t=2.28$; $p=0.04$) showed that after putting peer assessment into practice in their EFL classes, the students were more convinced than before that peer assessment helps to improve their skills of making good oral presentations. In the case of item 4, on the other hand, after the whole process of peer assessment, the students were not as firmly convinced as before that peer assessment actually helps them to understand teacher's requirements

better. None of the students disagreed to the statement; however, 50% of them remained neutral compared to the 14.3% in the pre peer assessment survey. Thus, a number of students who had agreed or strongly agreed to the statement 4 in the pre-survey, changed their opinion to neutral, and according to the results of the t-test ($t=3.61$, $p=0.003$), the decrease of the mean score in the post peer assessment survey was significant.

Secondly, items 5-6 were related to motivation. Statement 5 was aimed at finding out whether, in the opinion of the students, peer assessment motivated them to work harder on their presentations; whereas statement 6 focused on peer assessment as a motivator for the students to concentrate more carefully on the presentations of their peers.

Table 5. The comparison of the mean scores of Likert items 5-6 in pre- and post peer assessment survey.

Item	1 Strongly disagree % (no.of students)	2 Disagree % (no.of students)	3 Neutral % (no.of students)	4 Agree % (no.of students)	5 Strongly agree % (no.of students)	Mean	Standard deviation	Paired t-test	
								T	p
5 pre	0	14.3% (2)	28.6% (4)	35.7% (5)	21.4% (3)	3.6	1.0	2.6874	0.0186
post	0	21.4% (3)	35.7% (5)	35.7% (5)	7.1% (1)	3.29	0.91		Sig. dif.
6 pre	0	0	28.6% (4)	28.6% (4)	42.9% (6)	4.1	0.86	2.2804	0.0401
post	0	0	28.6% (4)	57.1% (8)	14.3 (2)	3.86	0.66		Sig. dif.

Looking at the mean scores of items 5 and 6 in the context of both pre- and post peer assessment surveys in Table 5 above, reveals that, similarly to the items 1-4, the majority of the students were either neutral or agreed to the statements. In the case of item 6 none of the students disagreed or strongly disagreed to the statement; whereas in the case of item 5, 2 students out of 14 (14.3%) disagreed in the pre peer assessment survey and 3 students (21.4%) in the post survey. Yet, the results of the t-test revealed a significant difference in the attitudes towards these two statements before and after carrying out peer assessment in the EFL classes. Namely, in the post peer assessment survey, there were

fewer students to strongly agree to statement 5 and more students who disagreed and remained neutral. Thus, the number of the students who were convinced that peer assessment motivated them to work harder on their presentations had decreased after their actual peer assessment experience ($t=2.69$, $p=0.02$). T-test showed a significant difference also in the case of item 6 ($t=2.28$, $p=0.04$); here, however, the change of the attitudes was not as striking as in the previous example. Namely, in the pre-peer assessment survey 28.6% of the students agreed and 42.9% of the students strongly agreed to the statement that peer assessment motivated them to concentrate more carefully on their classmates' presentations; whereas in the post peer assessment survey, the percentage of those who strongly agreed had decreased (14.3%), and the percentage of those who agreed had increased (57.1%). The percentage of those who were neutral remained unchanged (28.6%). Thus, the large majority of the students still agreed to the statement but not as firmly as before.

Thirdly, the seven final items dealt with the issues related to the assessment process. More specifically, statements 7-9 were aimed at investigating whether the respondents saw their peers as capable of giving useful comments, capable of assessing according to the marking scale, as well as capable of being objective in their decisions.

As demonstrated in Table 6 below, in the case of all the three items, at least half of the students agreed and strongly agreed to the statements in both pre- and post peer assessment surveys. Another common tendency regarding the answers to the statements 7-9 is, that in the pre peer assessment survey, none of the respondents strongly disagreed; whereas in the post peer assessment survey one student (7.1%) out of 14 strongly disagreed to statements 7 and 8, and two students (14.3%) strongly disagreed to statement 9. What is more, the students' reaction to item 8 is different from their reactions to the other items in the sense that none of the respondents were neutral towards the issue in the

pre peer assessment survey and only one student was neutral in the post peer assessment survey. In the case of all the other twelve items the number of the students who were neutral was bigger. What concerns the t-tests of the items 7-9 in pre- and post peer assessment surveys, the difference cannot be considered statistically significant.

Table 6. The comparison of the mean scores of Likert items 7-9 in pre- and post peer assessment survey.

Item	1 Strongly disagree % (no.of students)	2 Disagree % (no.of students)	3 Neutral % (no.of students)	4 Agree % (no.of students)	5 Strongly agree % (no.of students)	Mean	Standard deviation	Paired t-test	
								T	P
7 pre	0	14.3%	35.7%	42.9%	7.1%	3.43	0.85	1.0000	0.3356
post	7.1% (1)	14.3% (2)	21.4% (3)	57.1% (8)	0	3.29	0.99		Not sig.
8 pre	0	35.7%	0	57.1%	7.1%	3.36	1.08	0.5631	0.5830
post	7.1% (1)	28.6% (4)	7.1% (1)	42.9% (6)	14.3% (2)	3.29	1.27		Not sig.
9 pre	0	14.3%	35.7%	50%	0	3.36	0.74	0.6939	0.5000
post	14.3% (2)	21.4% (3)	14.3% (2)	28.6% (4)	21.4% (3)	3.21	1.42		Not sig.

In the case of the last four Likert items the students had to evaluate their own assessment behaviour. As seen in Table 7 below, T-test showed a significant difference regarding the answers given to item 10 in pre- and post peer assessment surveys ($t=2.69$, $p=0.02$). Namely, compared to the pre peer assessment survey, there were more students in the post-survey, who agreed (59% against 35.7%) and strongly agreed (28.6% against 21.4%) to the statement according to which they were able to assess their peers objectively; whereas fewer students were neutral (14.3% against 28.6%) and disagreed (7.1% against 14.3%).

Table 7. The comparison of the mean scores of Likert items 10-13 in pre- and post peer assessment survey.

Item	1 Strongly disagree % (no.of students)	2 Disagree % (no.of students)	3 Neutral % (no.of students)	4 Agree % (no.of students)	5 Strongly agree % (no.of students)	Mean	Standard deviation	Paired t-test	
								T	p
10pre	0	14.3% (2)	28.6% (4)	35.7% (5)	21.4% (3)	3.64	1.0	2.6874	0.0186 Sig. dif
Post	0	7.1% (1)	14.3% (2)	50% (7)	28.6% (4)	4.0	0.88		
11pre	14.3% (2)	21.4% (3)	42.9% (6)	0	21.4% (3)	2.93	1.33	0.8062	0.4346 Not sig.
Post	7.1% (1)	42.9% (6)	21.4% (3)	21.4% (3)	7.1% (1)	2.79	1.12		
12pre	14.3% (2)	28.6% (4)	42.9% (6)	14.3% (2)	0	2.57	0.94	1.8829	0.0823 Not sig.
post	0	42.9% (6)	35.7% (5)	21.4% (3)	0	2.79	0.8		
13pre	14.3% (2)	28.6% (4)	42.9% (6)	14.3% (2)	0	2.57	0.94	0.0000	1.000 Not. sig.
Post	0	50% (7)	42.9% (6)	7.1% (1)	0	2.57	0.65		

As for the statements 11-13, the t-tests did not show a significant difference between the answers given in the pre- and post peer assessment survey; however it is noteworthy that only a small minority of the students agreed or strongly agreed to these statements before as well as after their peer assessment experience in the EFL classes. For example, after trying out peer assessment in their EFL classes, 50% of the students found that giving negative comments to their peers was not uncomfortable for them; 21.4% of the students marked their opinion as neutral, whereas only 28.6% admitted feeling uncomfortable when giving their classmates negative feedback. What concerns encountering difficulties in commenting their peers' presentations (item 12), only 14.3% of the students agreed in the pre-survey that it might not be an easy task to decide how to comment the presentations of their classmates; the results of the post peer assessment survey revealed that after the peer assessment procedure the percentage of those students had risen to 21.4%, but the rest of the respondents remained neutral (35.7%) or disagreed (42.9%) to the statement. The similar tendency also applies to item 13. Namely, in the pre

peer assessment survey only 14.3% of the students agreed to the statement that deciding which marks to give to their classmates for their presentations was difficult; and in the post peer assessment survey the percentage was even lower – 7.1%. What is more, a relatively large number of the respondents (42.9% in both pre- and post peer assessment surveys) marked their answer as neutral in this matter.

In addition to the above presented Likert items, the post peer assessment questionnaire contained some multiple choice questions (see Appendix 2) which were also aimed at finding out about the students' attitudes towards peer assessment. Furthermore, what concerns the last two questions, the students were also asked to explain the reasons behind their choices of the answers. These explanations are presented and discussed in the section *2.4 Discussion of the Results*; whereas the following passages give an overview of the results of the multiple choice questions.

To begin with, after the whole peer assessment procedure, the sample students were asked about their preferences regarding different types of peer assessment and the results were as follows: 42.9% (6 students) of the respondents chose summative peer assessment, whereas only 7.1% (1 student) decided upon peer assessment in the form of comments; 28.6% (4 students) preferred both summative peer assessment as well as peer assessment in the form of comments; and 21.4% (3 students) claimed that they did not prefer neither peer assessment in the form of marks nor in the form of comments. As the percentages indicate, the majority of the students were in favour of using summative peer assessment rather than commenting on their peers works.

Secondly, the students were asked whether they were interested in using peer assessment in their EFL classes also in the future. In this case, 35.7% (5 students) gave a positive answer; the whole 42.9% (6 students) could not make up their mind regarding this

issue and, thus, chose the answer “I do not know”; whereas, similarly to the results of the previous question, only 21.4% (3 students) gave a negative answer.

Finally, in order to sum up the whole procedure of peer assessment in their EFL classes, the students’ were asked to express their preferences concerning the use of various types of assessment in their future EFL classes. Here, half of the students (7) preferred teacher assessment; 21.4% (3 students) expressed their liking towards using peer assessment; and 28.6% (4 students) chose the combination of peer and teacher assessment as their preference. The further examination of the data showed that some of those students who could not make up their minds in the case of the previous question, decided here upon teacher assessment. Thus, according to the results of the last question, after their experience of peer assessment, the sample students became divided into two equal groups, the one of which still preferred the traditional teacher assessment; whereas the members of the other group were interested in using peer assessment either individually or in the combination with teacher assessment.

All in all, in the light of the first research question it can be concluded that most of the sample students were in general positively disposed towards peer assessment both before and after trying it out in their EFL classes by assessing their peers’ oral presentations. In some cases, the t-tests revealed statistically significant differences between the answers given before and after the peer assessment process; yet, the majority of the students had not changed their attitudes from one side of the Likert scale to the other. However, despite the fact that the majority of the students were positively disposed to peer assessment both before and after trying it out in their EFL classes by assessing their peers’ oral presentations, exactly half of them would still prefer only teacher assessment in their future EFL classes.

2.3.2 Research Question Two

How do the marks students get for their presentations influence their attitudes towards peer assessment?

In order to find out whether there is a relation between the marks the students get from their peers and their attitudes towards peer assessment, the mean peer assessment marks were analysed in the light of the students' answers to the last question of the post peer assessment questionnaire as well as in the light of their choices regarding Likert scale items of the post peer assessment questionnaire. It was expected that the students who had got lower marks from their peers were not more negatively disposed to peer assessment than their peers who got higher marks.

The total mean score of peer assessment was 23.44 which in the terms of marks corresponded to a high "4". The final peer assessment marks for each presentation were distributed as follows: four students got "5"; eight students "4" and only two students received "3". Thus, based on these peer assessment marks, the students were divided into three groups in order to investigate the possible connection between the marks and the attitude.

To begin with, Table 8 gives an overview of the students' answers to the question which type of assessment they prefer in their future EFL classes.

Table 8. Statistics regarding mean peer assessment scores/marks and students' answers to the question which type of assessment they prefer in their EFL classes in the future.

	n	Prefer teacher assessment (no. of students)	Prefer peer assessment (no. of students)	Prefer the combination of teacher and peer assessment (no. of students)
Group 1: students whose mean peer assessment score was less than 21 points (mark 3).	2	2	0	0
Group 2: students whose mean peer assessment score was 21-24 points (mark 4)	8	4	2	2
Group 3: students whose mean peer assessment score was 25-28 points (mark 5)	4	1	1	2

As can be seen from the table above, both students who got the lowest peer assessment marks claimed that in the future they preferred teacher assessment to peer assessment.

Looking at the second and the largest group of students whose presentations deserved “4” from their peers, it can be seen that 50% (4 students) of these students would like to use only teacher assessment in their future EFL classes; whereas the members of the other half of the group prefer either peer assessment (2 students) or the combination of teacher and peer assessment (2 students).

What concerns the third group of the sample students, whose peer assessment mark was “5”, the majority of them are in favour of using the combination of teacher and peer assessment (2 students) or peer assessment (1 student); whereas only one student prefers teacher assessment.

Secondly, with the help of the Likert scale answers of the post peer assessment survey it was investigated whether there can be seen a connection between the marks students got from their peers and their attitudes. These results are presented in Table 9 below.

Table 9. Statistics regarding the mean scores of Likert items of the post peer assessment survey in relation to the marks the students got from their peers.

	n	Items 1-6		Items 7-10		Items 11-13	
		Mean score	SD	Mean score	SD	Mean score	SD
Group 1: students whose mean peer assessment score was less than 21 points (mark 3).	2	3.67	0.78	2.25	0.89	3	0.89
Group 2: students whose mean peer assessment score was 21-24 points (mark 4)	8	3.65	0.53	3.44	1.1	2.67	0.82
Group 3: students whose mean peer assessment score was 25-28 points (mark 5)	4	3.92	0.88	4.13	1.0	2.67	0.98

In Table 9 the Likert items have been divided into three groups and the mean scores of the students' answers have been calculated. Based on the mean scores of those three categories the following conclusion can be drawn:

- the higher the mean score of items 1-6 is, the more the students agreed that peer assessment is beneficial to them for several reasons (e.g. improves their EFL skills, skills of making oral presentations, skills of analysing the presentations of others, motivates to concentrate more carefully on their classmates presentations etc.).
- the higher the mean score of items 7-10 is, the more the students were convinced that their peers could give useful comments, assess according to the rating scale and that both the respondent himself/ herself as well as the classmates were able to assess objectively.
- the lower the mean score of items 11-13 is, the more the students disagreed to the statements according to which giving negative comments was uncomfortable and deciding which comments to write and which marks to give difficult.

Thus, the results presented in Table 9 above demonstrate that as for the items 1-6, the mean scores of all the three groups are higher than 3, which indicates that regardless of the mark the students had got from their peers, the majority of them found peer assessment to be useful and beneficial for several reasons listed above. The mean score of the group of the students who got the highest marks, is considerably higher (3.92) than the mean scores of groups 1 and 2 which are very similar (respective 3.67 and 3.65). However, the standard deviations of groups 1 and 3 are much higher than the standard deviation of group 2, although group two had the biggest number of students. This indicates that in groups 1 and 3 there was a variety of different opinions regarding the statements 1-6.

What concerns items 7-10, the differences between the mean scores of the three groups are considerably larger than in the case of the previous items. The mean score given

to items 7-10 by the students who got the lowest marks is only 2.25, which indicates that these students did not consider their peers as competent and objective assessors. Their mean score is significantly lower than the mean scores of groups 2 and 3, respective 3.44 and 4.13. The fact, that the students of group 1 were disappointed in their peers as competent and objective assessors, might be a reason, why they were also in favour of teacher assessment in their future EFL classes (see Table 8). The students who got the highest mark were, on the other hand, determined that the comments and marks given by their peers as well as by themselves had been adequate and objective.

Finally, as for the items 11-13, the mean scores of the three groups do not differ as greatly as in the case of the previous items presented above. The students of all the three groups seem to be rather neutral towards the statements according to which commenting on and marking their peers presentation had been difficult, and giving negative comments uncomfortable. However, the mean score of group 1 is higher (3.0) than the mean scores of the other groups (2.67 for both), which indicates that the students who got the highest marks found commenting on and marking their peers' presentations less complicated and uncomfortable than the students of group 1.

To sum it up, as the sample of the students is small (14 students), the conclusions one can draw about the relation between the attitudes towards peer assessment and the marks the students got from their peers are limited. For example, there were only two students who got the lowest mark and, thus, the mean scores of group 1 were based on the answers of these students alone. However, according to the limited results of the current study, the students who got the lowest marks were less convinced in the benefits of peer assessment than the students who got higher marks. Furthermore, the students who got the lowest marks found that their peers had not been competent and objective assessors, whereas the students who got the highest marks were convinced that their peers had given

adequate marks and comments as well as been objective in their decisions. The answers of the post peer assessment survey also revealed that the students who got the lowest marks were in favour of using teacher assessment in their future EFL classes, but the majority of the students who got the highest marks preferred either peer assessment or the combination of peer and teacher assessment.

2.3.3 Research Question Three

How well do the marks given by the students correspond to the marks given by the teacher?

To answer the third research question about the level of correspondence of the marks given by the students and the teachers, the results of the summative assessment feedback forms (Appendix 5) were analysed. In order to increase the reliability of teacher assessment, summative assessment was carried out by two EFL teachers – by the current teacher of the group and by the author of the study. In the case of each presentation the students as well as the teachers assessed seven aspects according to the marking scale. Altogether there were 126 summative feedback forms, 98 of which were filled in by the peers (7 forms per each performer) and 28 filled in by the two teachers.

To begin with, Table 10 demonstrates the total mean scores given to the oral presentations by the teachers as well as the peers.

Table 10. Statistics regarding the total mean scores given to the presentations. Maximum score that could be obtained for each presentation was $4 \times 7 = 28$ points.

Total	Teacher 1*	Teacher 2**	Teacher 1 + Teacher 2	Peers
Mean	22.86	23.21	23.04	23.44
Standard deviation	2.9	2.97	2.89	2.67

*Teacher 1 is the current English teacher of the sample group.

**Teacher 2 is the author of the study

The statistics presented above reveal that the total mean score of peer assessment (23.44) does not differ significantly from the mean score of the teachers' assessment (23.04). In the terms of marks, both scores would have given a "4".

In order to get a more detailed picture about the correspondence of teachers' and students' marks, the results of the summative feedback forms should be looked at in the context of each performer individually. The results are presented in Table 11 below.

Table 11. Statistics of the mean scores given by the teachers and students to each performer.

	Teachers' mean (max 28 p.)	Mark based on points given by the teachers	Peers' mean (max 28 p.)	Mark based on points given by the peers
Student A	26 p.	5-	26.1 p.	5-
Student B	23 p.	4	23 p.	4
Student C	23 p.	4	24.1 p.	4+
Student D	28 p.	5	26.9 p.	5-
Student E	25.5 p.	5-	27.3 p.	5-
Student F	23 p.	4	21.1 p.	4-
Student G	25 p.	5-	23.6 p.	4+
Student H	23.5 p.	4+	23.9 p.	4+
Student I	22 p.	4	25 p.	5-
Student J	24.5 p.	5-	24 p.	4+
Student K	18 p.	3	19.4 p.	3
Student L	22 p.	4	23 p.	4
Student M	22 p.	4	23.1 p.	4
Student N	17 p.	3	17.7 p.	3

As seen from Table 11, in the case of 3 students (21.4%) (Students A, B, H) the scores given by the peers and the teachers are identical or the difference is smaller than 0.5 points. What concerns the scores given to the rest of the performers, the differences between the teachers' and students' scores are larger than 0.5 points. A total of 7 students (50%) have obtained higher scores from their peers than from the teachers (Students C, E, I, K, L, M, N) and 4 students (28.6%) (Students D, F, G, J), on the other hand, have been assessed more severely by their peers than by the teachers. In the case of three students (G, I, J), the different scores given by the teachers and the peers lead also to different

marks. Based on the scores given by the teachers Students G and J would have got a “5-“, whereas according to the points from their peers the mark would have been “4+”. With Student I it is the other way round; the teachers found that his/her presentation was worth a “4”, but the score given by the peers would have resulted in “5-“. The largest difference between the points given by the peers and the teachers is 3 points and it occurred in the case of the previously mentioned Student I.

In addition to the data introduced above, the similarities and differences of the scores given by the teachers and the peers can also be analysed in the context of the seven aspects that the assessors had to assess in the case of each oral presentation. When the analysis of the scores in the context of each student’s presentation revealed that in the case of some performers, the scores given by the teachers and the peers varied enough in order to result in a different mark for the presentation, then the analysis of the scores in the light of different aspects does not reveal such a significant difference between teachers’ and students’ points.

Table 12. Comparison of teachers’ and students’ mean scores given to different aspects of oral presentations

	Teachers’ mean (max 4p)	Teachers’ Standard deviation	Students’ mean (max 4p)	Students’ Standard deviation	Paired t-test	
					t	P
Content	3.75	0.44	3.72	0.49	0.00	1.0 Not sig.
Organization	3.64	0.49	3.66	0.50	0.25	0.80 Not sig.
Grammar, vocabulary, sentence structure	3.25	0.7	3.14	0.69	0.54	0.59 Not sig.
Pronunciation	3.21	0.63	3.21	0.52	0.63	0.54 Not sig.
Voice control and fluency	2.89	0.74	3.05	0.68	1.16	0.26 Not sig.
Visual props	3.79	0.42	3.68	0.51	1.31	0.2 Not sig.
Body language	2.5	0.75	2.94	0.84	2.52	0.0179 Significant

As demonstrated in Table 12 above, in the case of content, organization and pronunciation the mean scores of teachers and students are very close to each other. What concerns the aspects such as grammar, vocabulary and sentence structure; voice control and fluency; and visual props, a slight difference can be noticed. For instance, when assessing grammar, vocabulary and sentence structure as well as visual props, the students have been a little bit more severe in their ratings than the teachers as the bias in both cases is 0.11. As for the aspect of voice control and fluency, it is the other way round, the mean score of the teachers is higher giving the bias of 0.16. Nevertheless, these differences can be considered minor ones. Furthermore, even the standard deviations of teachers' and students' scores are very similar in most cases.

The aspect of body language has not been touched upon yet, and this appears to be the only aspect in the case of which the students' and teachers' scores can be considered significantly different. Namely, the mean score of the students is 2.94, whereas the mean score of the teachers is 2.5. Also the results of the t-test confirm the significance of the difference ($t=2.52$, $p=0.02$).

In the light of the third research question, it is also relevant to introduce the result regarding the students' answers to the question which aspects had been the most difficult to assess. They could choose several aspects and according to the results the aspect of grammar, vocabulary and sentence structure was considered the most difficult to assess by 50% (7 students) of the respondents. The next aspects which were seen as complicated to assess were content (42.9%, 6 students) and pronunciation (35.7%, 5 students). Body language, the aspect in the case of which there was the largest difference between the scores given by the teachers and the peers, was mentioned by 28.6 % (4 students) of the respondents, whereas organization and visual props were both mentioned twice, and voice

control and fluency once. Thus, some aspects were mentioned more frequently some less, but all of them were mentioned at least once.

To sum up the section, the analysis of the data revealed no significant difference between the total mean score given by the teachers and the total mean score given by the students. Similarly, when analysing the scores in the context of all the seven aspects the assessors had to assess in the case of each oral presentation, a significant difference was reported only in the case of the aspect of body language. However, the comparison of the scores and marks of the teachers and peers given to each performer individual, revealed that half of the students got somewhat higher scores from their peers than from the teachers, whereas four performers (28.6%) were assessed more severely by their peers than by the teachers.

2.4 Discussion of the Results

In the current section the most interesting results of the ones presented above are discussed. In order to explain and illustrate the reasons behind the results, some examples of the students' answers to open questions are provided. Furthermore, examples of the students' comments regarding their peers' oral presentations are also presented in the section. In addition to that, the results of the present study are compared with the results of similar former studies which were introduced in the first part of the thesis.

To begin with, the results of the results of the questionnaires revealed that the majority of the 14 adolescent sample students were positively disposed to peer assessment both before putting it into practice in their EFL classes as well as after the peer assessment procedure when they had become aware of the scores and marks of peer and teacher assessment. Most of the students agreed to the statements according to which peer assessment was beneficial in terms of developing their EFL skills, their skills of making

oral presentations, their skills of analysing the presentations of their peers, as well as their ability of understanding teacher's requirements better (items 1-4). Furthermore, the majority also found that due to peer assessment they concentrated more carefully on their peers' presentations as well as paid more attention to preparing their own presentations (items 5-6). These implications are based on the fact that the mean score given to each item 1-6 was over 3.0 both in the pre- as well as in the post peer assessment survey. However, the majority of the mean scores of these 6 items were somewhat lower in the post peer assessment survey than in the pre peer assessment survey. In some cases the difference cannot be considered significant, but what concerns the issues such as understanding teacher's requirements better thanks to peer assessment, putting more effort into preparing one's own presentation, as well as peer assessment as a motivator to focus more intensely on the classmates' presentations, the difference can be considered significant based on the results of the t-tests. Only, in the case of the first item which stated that peer assessment helped to understand how to make a better oral presentation, the mean score of the post peer assessment survey was significantly higher than the mean score of the pre-survey. The latter is consistent with the previous studies of Nakamura (2002), Cheng and Warren (2005), and White (2009) according to the results of which the sample students also found that peer assessment had been beneficial in the terms of improving their EFL oral presentation skills (see also section 1.2.2). Thus, the results of the first 6 items indicate that basic school students held positive attitudes towards using peer assessment in their EFL classes both before and after the whole peer assessment procedure as they found this type of assessment beneficial in several ways; yet it is important to keep in mind that the majority of the mean scores of the post peer assessment survey were more or less lower than the previous ones.

The same tendency could be noticed also in the case of items 7-9 which focused on the quality and objectivity of peer assessment. The majority of the students marked their opinion as neutral or agreed to the statements according to which their peers could write useful comments, rate according to the rating scale as well as remain objective in their decisions, however, although the t-tests did not show a significant difference between the mean scores of pre peer assessment survey and post peer assessment survey, the mean scores of items 7-9 were lower in the post peer assessment survey. In the pre peer assessment survey, there were not any students who had strongly disagreed to these three statements, but in the post peer assessment survey there were some, which indicates that some students had become disappointed in their peers' assessment skills. Also the comments that the respondents wrote in relation to question 4 in the post peer assessment questionnaire demonstrate the disappointment of some of them. Some examples of the reasons of why they would not like to practise peer assessment in their future EFL classes were as follows:

- *Some marks were unfair.*
- *Some students got marks that they didn't deserve.*
- *Too complicated and not very fair.*

On the other hand, the number of students who strongly agreed to statements 8 and 9 had also increased, which proves that although a couple of students had become more negatively disposed to their classmates assessment skills, some students had at the same time become more certain in their classmates as skilful and objective assessors.

Interestingly, an exception is item 10, in the case of which each student had to evaluate his/her own ability of being an objective assessor. Unlike the mean scores of all the other items presented above, except item 1, the mean score given to item 10 was significantly higher after the peer assessment procedure than it had been before. Moreover, the post peer assessment mean score was 4.0 which was the highest of all the mean scores

given to items 1-10 by the students after the peer assessment procedure. Thus, when in the case of item 9 there were altogether 5 students (35.7%) who held the opinion that their peers had not been objective assessors, then what concerns one's own objectivity, only one student doubted in his/her ability to remain objective while assessing the peers' presentations, whereas the large majority of the students (78.6%) either agreed or strongly agreed to the statement according to which they had been able to assess their peers objectively. This indicates that despite the fact that the students did not doubt in their own abilities of being fair assessors, some of them still did not consider their classmates as reliable assessors as they themselves. This demonstrates nicely the complexity of self-assessment and the close connection between self- and peer assessment which was discussed in section 1.2.1.

Coming once again back to item 7, in the light of the analysis of the feedback forms with the students' comments, it was somewhat surprising that after the students had read the comments their peers had written to them, the majority of the respondents stated that these comments had been useful. In the opinion of the author of the study as well as the English teacher of the group there was a plenty of room for improvement in what concerned the comments of the sample students. Although the students had been taught how to give constructive feedback (Appendix 3) and they had had a chance to practise it on the example of a video before they actually started to comment their peers' oral presentations, the analysis of the written comments revealed the following most common shortcomings which occurred in the case of several feedback forms:

- lack of examples or reasoning (e.g. *Your pronunciation wasn't very good and there were several mistakes*);

- writing comments in the 3rd person, as if they were not writing to the performer but rather to the teacher (e.g. *He usually stands straight, but he also looked away from the audience sometimes.*);
- copying exact sentences from the rubric (e.g. *Stands up straight, looks relaxed and confident. Has eye contact with the listeners.*)

Fortunately there were also a number of examples of constructive feedback as well as nice encouraging comments for the future performances. Some examples are as follows:

- *You were bit in hurry, try to slow down next time.*
- *Eiffel tower is high not long.*
- *You spoke Estonian sometimes. Try to avoid that!*
- *Don't be nervous, we won't bite you:)*
- *You seemed a bit insecure, don't be afraid of us!*
- *You showed a picture about Stonehenge, but didn't talk anything about it.*
- *You had brought out very interesting pictures and I enjoyed them.*

All in all, based on the analysis of the written comments, it can be concluded that commenting on their classmates EFL presentations is not an easy task for basic school students and more time and effort should be spent on training in order to improve the quality and efficiency of the students' feedback.

What concerns items 11-13, the lower the mean scores of these items were, the less uncomfortable the students had felt when giving negative feedback to their peers, and the less difficult it had been for them to decide which comments to write and which marks to give to their classmates. Similarly to item 10, evaluating these items required self-assessment skills. In section 1.3.3 it was discussed that one possible reason why students may not be willing to take part in peer assessment is that they find it stressful to assess their classmates (Falchikov 2005: 161). Based on the scores the sample students of the current study gave to items 11-13, it can be concluded that this was not an issue for the majority of the students here. After the peer assessment procedure as much as half of the students disagreed that giving negative comments had been uncomfortable and some of

them remained neutral towards this aspect, whereas only 4 students admitted that they had felt uncomfortable. These results differ from the ones revealed in the study by Maiz Arevalo (2008: 132) according to which most of her sample students admitted feeling uncomfortable when assessing their peers. One possible reason, why the sample students of the present study found that giving negative feedback had not been a problem to them, is that the quality of the presentations was quite good, as can be seen from the marks and scores presented in Table 8 in section 2.3.3. Thus, in the case of a number of presentations the portion of positive feedback exceeded the portion of negative feedback and that is why the assessors did not feel uncomfortable. It is possible that if there had been more presentations with lower quality, which would have required more negative feedback, the attitudes of the sample students towards this issue would have been different, too.

Thus, based on the analysis of the Likert scale answers, the majority of the sample students held a positive attitude towards using peer assessment in their EFL classes. Similarly, also Peng (2009) reported the positive reaction of the students towards using peer assessment in their EFL classes; however, according to the results of Peng's (2009: 145) study, after experiencing peer assessment, the students' attitudes had become significantly more favourable than before. In the current study, on the other hand, this kind of tendency was noticed only in the case of items 1 and 10, which were demonstrated above.

One possible reason behind the decrease of the mean scores of the post peer assessment survey of the present study can be that the sample students were much younger than the sample students of Peng's study. The 8th-formers of the current study had not experienced this type of assessment in their EFL classes before, they were excited in the beginning, however as the whole process of peer assessment was quite long and detailed together with the training phase and filling in different types of feedback forms, the

students simply got tired by the time they finally had to fill in the second questionnaire and the feeling of boredom might also reflect in the slight decrease of the mean scores of the second questionnaire. This assumption is also supported by the following reasons the respondents brought out in relation to the question why they would not like to use peer assessment in their future EFL classes:

- *I'm too lazy, actually.*
- *It can be helpful, but I didn't like it very much.*
- *It was stressful but useful.*
- *It was a bit difficult.*
- *Too complicated and not very fair.*
- *I don't want so many marks.*

However, in addition to the above listed ones, the respondents also wrote a number of positive comments in relation to the question why they would like to continue using peer assessment in their EFL classes. These comments were as follows:

- *Because it helps me to understand my classmates.*
- *It was pretty interesting. I like assessing the others and this type of assessment is fairer.*
- *Because it makes me listen to my classmates presentations more carefully and assess more objectively.*
- *Because it's honest.*
- *Because classmates are softer on you than the teacher and understand you.*

The latter listing nicely illustrates and complements the results of the Likert items presented above.

The issues discussed in the passages above were related to the first research question which was aimed at investigating the students' attitudes towards peer assessment before and after the peer assessment process. The second research question, on the other hand, focused on examining how the results of peer assessment influence the basic school students' attitudes towards this type of assessment. As stated above, due to the small size of the sample, broad generalisations could not be made about the relation of peer assessment marks and students' attitudes. However, the limited results of the study

revealed that the students who got the lowest mark “3” were more negatively disposed to peer assessment than the students who got the highest mark “5”. The further examination of the questionnaires in the light of the second research question revealed that those students who had become more critical towards their peers’ assessment skills after the whole peer assessment procedure, had deserved from their classmates a “3” or a “4” as a mark for their presentation; whereas none of the students who had deserved a “5” doubted in his/her classmates’ objectivity and assessment skills. Thus, the results of the present study, indicate that the nature of peer feedback is likely to influence basic school students’ attitudes towards peer assessment. A reverse tendency was, on the other hand, noticed in the case of college students by Saito and Fujita (2004) who studied students’ attitudes towards peer assessment in the context of EFL writing tasks and concluded that the nature of peer feedback was not likely to influence the students’ attitudes, but as noted above the age difference between Saito and Fujita’s sample students and the students of the current study is remarkable.

Returning to the results of the present study, which has a sample consisting of basic school students, an interesting finding in relation to the attitudes of the students who got the lowest mark will be presented. Namely, these two students who got the lowest mark “3” for their presentations stated in the questionnaire that they would not like to practise peer assessment in their future EFL classes because “*some of the marks were unfair*” and preferred only teacher assessment because it was “*the fairest*” type of assessment and “*the teacher is more experienced in assessment*”. What is striking, is that although they both disagreed to the statements according to which their classmates had been able to assess their presentations according to the rating scale and claimed that teacher assessment was more objective, the comparison of teachers’ and peers’ marks and scores given to these two students (see Table 8 in section 2.3.3) reveals that they actually got the same mark

from the teachers, too. Furthermore, the mean score given by the two teachers was even a little bit lower than the mean score given by their peers. In the case of Student K it was 18 points from the teachers against 19.4 points from his/her peers, and Student N got 17 points from the teachers and 17.7 points from the peers. Yet, they were disappointed in their peers' assessment skills and claimed that teachers were more objective. This controversy can be related to the issue, treated in section 1.2.1, according to which students perceive feedback from peers and adults differently. Black *et al* (2003: 50) and Cole (1991, cited in Topping 2009: 22) stated that students tend to take the criticism from their peers more seriously than from their teachers. What is more, the sample of the present study consisted of adolescent 8th-formers who are in the age when the opinion of classmates is of great importance. Also in the case of Students K and N, it can be noticed that they were not disappointed in the marks and scores the teachers had given to them, but they were dissatisfied with the marks and scores of peer assessment, although the results of teacher and peer assessment were actually very similar. Thus, a "3" from the teachers was not perceived as seriously as a "3" from the classmates.

Interestingly, the average peer assessment mark of the whole sample was 4.1, which is identical to the average mark for the study year 2012/13. Thus, the quality of the oral presentations of the whole group corresponded to their overall English language level. A closer analysis of the peer assessment marks revealed that one student actually did better compared to the mark for the previous study year, as he/she got a "4" for the presentation, whereas the final mark of the year 2012/13 had been "3". Another student, on the other hand, obtained a lower mark for the presentation than his/her final mark for the previous year – namely, a "4" instead a "5". In the case of all the other students, the peer assessment marks for their presentations corresponded to the final marks they had received for the previous study year. This also indicates that the disappointment of some of the students in

their classmates' ability to provide adequate and objective assessment, may not be based on sound arguments, but are of emotional nature instead.

For a teacher who wants to use peer assessment with his/her basic school students it is, of course, worrisome that negative peer feedback may foster negative attitudes towards peer assessment. One solution here could be using peer assessment in the form of comments rather than in the form of marks, as the latter seem to cause more tension among adolescent students. Developing the students' skills of generating comments of good quality, on the other hand, requires more time and energy for training, as demonstrated by the results of the present study. Suggestions and steps to be taken in order to cope with the negative attitudes of the students were also suggested in sections 1.3.3 and 1.4.2. It requires time and energy, but it is in the teachers' power to help the students understand that they can actually make use of the negative feedback and turn it to their account.

Furthermore, when delivering oral presentations the students cannot observe their own performances in the way the audience can, and that is why their own judgements on how they coped with the task may not be adequate enough. That is why it could be beneficial to video, or at least record, the presentations in order to be able to analyse and discuss them later with each student individually. This way some of the students can be helped to understand that they did have some shortcomings in their presentations, and the peers have not been unfair to him/her as a person, they have just assessed the presentation, noted down the mistakes as well as their suggestions which the performer can now take advantage of in order to develop his/her presentation skills as well as EFL skills.

Finally, what concerns the third research question and the correspondence of teachers' and students' marks, the results of the present study demonstrated that as young students as the 8th-formers are able to assess their peers' EFL oral presentations very similarly to the teachers. This finding is consistent with the results of a number of former

studies that have investigated the correspondence of teacher and peer assessment in EFL classes in the context of higher education (e.g. Grez *et al* 2012, Peng 2010, Maiz Arevalo 2008, Patri 2002, Shimura 2006). As could be seen from Table 7 (section 2.3.3), the total mean scores of teacher and peer assessment were very close to each other, and also in the terms of the individual aspects that had to be assessed in the case of each oral presentation the differences were not significant (Table 9 in section 2.3.3). Interestingly, the only aspect, in the case of which a significant difference between students' and teachers' scores was revealed by the t-test, was the aspect of body language. The teachers appeared to be more severe in what concerned the body language of the performers than the students. This finding is contradictory to that of Shimura (2006: 105), who found that eye contact and gestures were the items in the case of which the students' and the teacher's ratings correlated most closely as these were the aspects that students could easily evaluate regardless of their EFL proficiency. One possible explanation to the significant difference between the teachers' and peers' scores given to the aspect of body language in the current study, may be that the students considered some other aspects more important and did not pay enough attention to the performers' body language, as it is something that is usually not focused on in their everyday EFL classes. This tendency was also touched upon in section 1.3.2. Namely, it has been noticed that during the peer assessment procedure in foreign language classes, the students tend to focus on finding mechanical mistakes in their peers' works and, thus, forget to pay attention to other important components (Sengupta 1998, cited in Shimura 2006: 100). Indeed, according to the results of the post peer assessment questionnaire the most difficult aspects to assess in the opinion of the sample students were language-related items such as grammar, vocabulary and sentence structure, followed by content and pronunciation. These results are comparable to the ones of Cheng

and Warren's (2013) study which also revealed that the sample EFL students were less confident about assessing language related criteria than non-language related criteria.

However, although the total mean scores of teacher and peer assessment as well as the mean scores given to different aspects by teachers and students were close to each other, it was noticed that in the case of some oral presentations the scores and marks given by the teachers disagreed the ones given by the students. As presented in the previous section half of the students got somewhat higher scores from their peers than from the teachers, whereby the biggest difference between the teachers' and students' scores was 3 points and it concerned the oral presentation of Student I (see Table 8 in section 2.3.3). In the opinion of the English teacher of the group, the higher peer assessment score for Student I may be an example of decibel grading (see section 1.3.1), as Student I always actively participates in the lessons, is eager to pronounce on a subject as well as willing to perform and organize. What concerns the case of Students F and G, who got somewhat lower scores from their peers than from the teachers, then according to the observations of their EFL teacher, these two students are the most modest ones in the group and not as socially active as the other members of the group, and this may be also a reason why they got a little bit lower scores from their peers. Thus, although the scores and marks given by the teachers and the students generally corresponded very well, some signs of friendship marking could be noticed. The assumption of the occurrence of friendship marking in the case of basic school students is also supported by some of the students' comments presented above which suggested that some students had obtained from their peers marks that they had not deserved.

In the light of the positive results of the correspondence of teachers' and students' marks, it is important to emphasize that it is not an intention to promote using summative peer assessment extensively with adolescent students. The aim of comparing teachers' and

students' scores and marks was exploratory and it was hoped to demonstrate to the sceptics who doubt in the validity and reliability of this type of assessment that adolescent students are able to assess similarly to their teacher. As pointed out in section 1.3, the main concern of a number of teachers is how well the results of peer assessment correspond to their own (Falchikov 2005: 190), and the results of the present study showed a close correspondence. Despite the fact that the students' scores and marks were similar to the ones given by the teachers, it is advised using peer assessment with adolescent students in the form of comments and as a formative assessment tool rather than a summative one.

Finally, it is noteworthy, that although the large majority of the students found peer assessment useful and saw that their scores were similar to the ones given by the teachers, as much as half of them would still prefer only teacher assessment in their EFL classes. The students were also asked to explain the reasons behind their choices and a number of students answered that teacher assessment was more objective and fair than peer assessment. One student, though, preferred teacher assessment due to the reason that it was simply *"easier and faster this way"*. What concerns the reasons of those students who marked peer assessment as their preference the following explanations were given:

- *Because students are kinder and understand better how terrifying it is to be in front of the class.*
- *Classmates are kinder.*
- *Because it's so much cooler.*

As can be seen from the examples above, these students do not value the aspects of objectivity and fairness as much as the students of who were in favour of teacher assessment, but emphasise the kindness and sympathy of their peers. However, similarly to the students who preferred only teacher assessment, the bunch of students who marked the combination of teacher and peer assessment as their preference, argued that this type of combined assessment method was the fairest and the most objective. One respondent also added that this kind of assessment is *"cooler"* because it enables the students to feel

themselves as teachers. The latter comment illustrates nicely one of the merits of peer assessment, which was also touched upon in section 1.2.1. Namely, peer assessment contributes to the feeling of being involved and having some power in the process of assessment.

In conclusion, the results of the present study lead to several implications. Firstly, EFL teachers should not hesitate to incorporate peer assessment into their everyday teaching practices with basic school students as the results demonstrated that the adolescent students of the sample were in general positively disposed to using peer assessment in their EFL classes as they found it beneficial to them for several reasons. Furthermore, the majority of them considered their peers as well as themselves as capable of assessing the others adequately and objectively. Secondly, the present study also revealed that what concerns summative assessment, as young students as the 8th-formers are able to assess their peers' EFL oral presentations similarly to the teachers. Namely, the scores and marks given by the sample students corresponded closely to the ones given by the teachers. However, some signs of friendship marking could be reported, and what concerns peer assessment in the form of comments, more training would have been needed in order to raise the quality of the feedback. Thirdly, it was noticed that the students who got the lowest mark from their peers were less positively disposed to peer assessment than those who got the highest mark. Despite the overall positive attitude of the large majority of the sample students and the fact that they assessed similarly to their teachers, half of them would still prefer only teacher assessment in their EFL classes as some of them found it less complicated and others fairer. The other half of the sample, on the other hand, was in favour of using peer assessment or the combination of peer and teacher assessment as they considered this type of assessment methods more objective, useful and "*cooler*". As the sample students experienced peer assessment in their EFL classes for the first time, it is

likely that if they are involved into this type of assessment also in the future, their assessment skills as well as their understandings of the benefits of peer assessment will improve.

CONCLUSION

The present thesis has investigated the use of peer assessment in basic school EFL classes. Peer assessment belongs to the category of “alternative assessment methods” which arose as a reaction to more “traditional assessment” (e.g. tests) in 1990s (Douglas 2010: 73). Topping defines (2009: 20) peer assessment as “an arrangement for learners to consider and specify the level, value or quality of a product or performance of other equal status learners”. Falchikov and Goldfinch (2000: 287) emphasize that in the case of peer assessment students “are engaging with criteria and standards” which they rely on when making judgement of the works of their peers. According to a number of researchers (Falchikov 2005, Topping 2003, Topping 2009, Black *et al* 2003, Butt 2010, Maiz Arevalo 2008, Douglas 2010, Sebba *et al* 2008) using peer assessment in classes has a variety of benefits such as raising the students’ awareness of learning goals and criteria for judging the quality of learning and knowledge; increasing learner responsibility; contributing to increasing students’ self-esteem etc.

However, despite the benefits of peer assessment, the question of validity and reliability of peer assessment has raised concerns among teachers as well as students. Validity and reliability of peer assessment can be influenced by social relations, lack of skills, and negative attitudes of the students. In order to assist teachers how to overcome these problems as well as help them with planning and organising peer assessment, several authors (Topping 2009: 25-26, Falchikov and Goldfinch 2000: 317-318, Falchikov 2005: 125, Weeden *et al* 2002, cited in Butt 2010: 81-82) have outlined guidelines and suggestions for successful implementation of peer assessment (e.g. involving students into the process of developing and negotiating the criteria for assessment; explaining the roles of assessors; training how to give constructive feedback etc.).

The purpose of the present study was to determine the efficiency of using peer assessment with basic school students in EFL classes. The topic was inspiring in two reasons. Firstly, the majority of the studies which have investigated peer assessment in the context of EFL classes have been carried out in higher educational settings; whereas very little data can be found about using peer assessment in EFL classes of basic school students. Secondly, during the last three years, the topic of peer assessment has become an issue worth discussing in Estonia as the Estonian National Curriculum for Basic School, which came into force in 2011, states that students should be involved into self assessment as well as peer assessment in order to raise their study motivation and develop their ability to set objectives and analyse their own learning process and behaviour in the context of those objectives (Ministry of Education and Research 2011a, section 20 (3)). As peer assessment had not been incorporated into the previous national curriculum, using this type of assessment more extensively still seems to be a relatively novel idea for Estonian students as well as for teachers. That is why the thesis was aimed at answering the following research questions:

1. What kind of attitudes do the students of form 8 have towards using peer-assessment in the form of marks as well as in the form of comments in EFL classes for assessing their classmates' oral presentations?
2. How do the marks students get for their presentations influence their attitudes towards peer assessment?
3. How well do the scores and marks given by the students correspond to the ones given by the teacher?

The research was carried out in Rakvere Reaalgümnaasium and the sample consisted of 14 students of form 8 who belonged into the same EFL learning group. During the first three 45-minute-long classes it was explained to the students what peer assessment

was, criteria for assessing oral presentations were negotiated and settled, and training on giving summative peer feedback as well as feedback in the form of comments was carried out. The students of the sample were divided into two groups. In the following two classes they gave their presentations and had chance to peer assess their classmates by giving them scores and marks as well as written comments. In order to find out about the attitudes of the students before and after the whole peer assessment procedure, they were asked to fill in two questionnaires the answers of which were compared. With the aim of investigating the correspondence of teachers' and students' assessment the results of the summative feedback forms were analysed and compared.

Returning to the questions posed above, it is now possible to state that most of the sample students were in general positively disposed towards peer assessment both before and after trying it out in their EFL classes by assessing their peers' oral presentations. In some cases, the t-tests revealed statistically significant differences between the answers given before and after the peer assessment process; yet, the majority of the students had not changed their attitudes from one side of the Likert scale to the other. The majority of them agreed that peer assessment had a positive influence on their EFL skills, their skills of making oral presentations, their skills of analysing the presentations of their peers, as well as on their ability of understanding teacher's requirements better. Furthermore, most of the students found that peer assessment served as a motivator to concentrate more carefully on their peers' presentations as well as on preparing their own presentations. What concerns the issues related to giving marks and comments to their peers and being objective in their decisions, the majority of the students saw their peers as capable of giving useful comments and assessing according to the rating scale; half of the respondents were convinced in their peers' objectivity, whereas the large majority was convinced of their own objectivity. As for the difficulties related to giving marks and comments to one's

peers, the minority of the students found it hard to make up their minds on the adequate comments and marks, whereas the rest of the students found it not difficult or remained neutral towards the issue. However, despite the fact that the majority of the students were positively disposed to peer assessment both before and after trying it out in their EFL classes by assessing their peers' oral presentations, exactly half of them would still prefer only teacher assessment in their future EFL classes.

What concerns the second research question, the results of the study suggest that in the case of basic school students, the nature of peer feedback is likely to influence their attitudes towards peer assessment. Namely, it was noticed that the students who got the lowest marks from their peers found peer assessment less beneficial than the students who got the highest marks. Furthermore, the students who got the lowest marks claimed that their peers had not been competent and objective assessors, whereas the students who got the highest marks stated that their peers had given adequate marks and comments as well as been objective in their decisions.

As for the third research question, the findings of the study indicate that basic school students are able to assess their peers' EFL oral presentations similarly to the teachers. Namely, the scores and marks given by the sample students corresponded closely to the ones given by the teachers.

As very little research has been done on using peer assessment with basic school students in the context of EFL classes, the findings presented above add to our understanding of the efficiency of using this type of assessment with adolescent students. As the sample students were positively disposed to peer assessment as well as able to assess their peers' oral presentations very similarly to their teachers, the results should be encouraging for EFL teachers to use peer assessment regularly with their basic school students.

The findings of the study also revealed some drawbacks related to using peer assessment with basic school students. Namely, the quality of the comments was not good enough, some signs of friendship marking could be noticed, and it was found that negative peer feedback fostered negative attitudes among the adolescent students. These issues could now be kept in mind by the teachers who start implementing peer assessment in their basic school EFL classes, as with the help of sufficient and adequate explaining and training these drawbacks can be prevented.

However, due to the small sample size, the findings of the present study should be treated with caution and broad generalizations cannot be made. Further research might explore and compare the attitudes that different age groups of basic school students have towards peer assessment. More data would also be needed to determine the connection between the nature of peer feedback and the attitudes of adolescent students. In addition to that, more information about the influence of training and using peer assessment with the same group over time would be appreciated.

All in all, peer assessment is a complex procedure, which requires time for preparation and organization; however, it has a number of benefits, and the results of the present thesis demonstrate that even adolescent students could understand and appreciate these. Furthermore, the results also indicate that basic school students can provide reliable summative assessment. It is hoped that EFL teachers can make use of the information presented in the thesis in order to incorporate peer assessment more successfully into their classes.

LIST OF REFERENCES

- Beaman, Ronda. 1998. The unquiet . . . even loud, andragogy! Alternative assessments for adult learners, *Innovative Higher Education*, 23:1, 47–59.
- Birjandi, Parviz, Masood Siyyari. 2010. Self-assessment and Peer-assessment: A Comparative Study of Their Effect on Writing Performance and Rating Accuracy. *Iranian Journal of Applied Linguistics*, 13(1), 23 - 45. Available at http://www.academia.edu/464232/Selfassessment_and_Peerassessment_A_Comparative_Study_of_Their_Effect_on_Writing_Performance_and_Rating_Accuracy , accessed March 13, 2013.
- Birjandi, Parviz, Nasrin Hadidi Tamjid. 2012. The role of self-, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education*, 37:5, 513-533.
- Black, Paul, Cristine Harrisson, Clare Lee, Bethan Marshall, Dylan Wiliam. 2003. *Assessment for Learning: Putting It Into Practice*. Midenhead: Open University Press.
- Butt, Graham. 2010. *Making Assessment Matter*. London: Continuum International Publishing Group.
- Cheng, Winnie, Martin Warren. 2005. Peer assessment of language proficiency. *Language Testing*, 22:1, 93-121.
- Cole, David A. 1991. Change in self-perceived competence as a function of peer and teacherevaluation. *Developmental Psychology*, 27, 682–688.
- De Grez, Luc, Martin Valcke and Irene Roozen. 2012. How effective are self- and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education*, 13:2, 129-142.
- Douglas, Dan. 2010. *Understanding Language Testing*. London: Hodder Education.
- Falchikov, Nancy, Judy Goldfinch. 2000. Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70:3, 287-322.
- Falchikov, Nancy. 2005. *Improving Assessment through Student Involvement: Practical Solutions for Aiding Learning in Higher and Further Education*. Oxon: Routledge Falmer.
- Gatfield, T. 1999. Examining Student Satisfaction with Group Projects and Peer Assessment. *Assessment and Evaluation in Higher Education*, 24:4, 365–77.
- Greenstein, Laura. 2010. *What Teacers Really Need to Know About Formative Assessment*. Alexandria, Virginia USA: ASCD.

- Jahin, Jamal Hamed. 2012. The Effect of Peer Reviewing on Writing Apprehension and Essay Writing Ability of Prospective EFL Teachers. *Australian Journal of Teacher Education*, 37:11, 59-84. Available at: <http://ro.ecu.edu.au/ajte/vol37/iss11/4>, accessed March 27, 2013.
- Lin, Sunny S.-J., Eric Zhi-Feng Liu and Shyan-Ming Yuan. 2001. Web based peer assessment: attitude and achievement. *IEEE Transactions on Education*, 44:2, 244-251.
- Maiz Arévalo, Carmen. 2008. Peer-assessment in the ESL Classroom: a Practical Project. *Porta Linguarum*, 9, 127-138. Available at http://www.ugr.es/~portalin/articulos/PL_numero9/9%20Carmen%20Maiz.pdf , accessed June 28, 2012.
- Meletiadiou, Elena. 2012. The impact of training adolescent EFL learners on their perceptions of peer assessment of writing. *Research Papers in Language Teaching and Learning*. 3:1, 240-251. Available at <http://rpltl.eap.gr> , accessed March 1, 2013.
- Ministry of Education and Research. 2011a. National Curriculum for Basic Schools 2011. Riigi Teataja I, 14.01.2011, 1.
- Ministry of Education and Research. 2011b. National Curriculum for Upper Secondary Schools 2011. Riigi Teataja I, 14.01.2011, 2.
- Nakamura, Yuji. 2002. Teacher assessment and peer assessment in practice. *Educational Studies*, 44, 203-215.
- Nelson, Gayle, Joan Carson. 1998. ESL students' perceptions of effectiveness in peer response groups. *Journal of Second Language Writing*, 7, 113-131.
- O'Donnell, Angela M, Keith Topping. 1998. Peers assessing peers: possibilities and problems. In Keith Topping, Stewart Ehly (eds) *Peer Assisted Learning*. Mahwah, NJ, & London: Lawrence Erlbaum Associates.
- Patri, Mrdula. 2002. The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing* 19, 119-131.
- Peng, Jui-ching. 2009. *Peer assessment of oral presentation in an EFL context*. Submitted to the faculty of the University Graduate School in partial fulfilment of the requirements for the degree Doctor of Philosophy in the Department of Literacy, Culture, and Language Education Indiana University. Available at <http://gradworks.umi.com/3380148.pdf> , accessed July 3, 2012.
- Peng, Jui-ching. 2010. Peer Assessment in an EFL Context: Attitudes and Correlations. In Matthew T. Prior, Yukiko Watanabe, Sang-Ki Lee (eds). *89 Selected Proceedings of the 2008 Second Language Research Forum*, 89-107. Somerville, MA: Cascadilla Proceedings Project. Available at <http://www.lingref.com/cpp/slrf/2008/paper2387.pdf> accessed July 3, 2012.

- Pond, Keith, Rehan Ul-Haq, Winnie Wade. 1995 Peer Review: a Precursor to Peer Assessment, *Innovation in Education and Training International*, 32:4, 314-323.
- Purchase, Helen C. 2000. Learning about interface design through peer assessment. *Assessment and Evaluation in Higher Education*, 25:4, 341–52.
- Richards, Jack C, Richard Schmidt. 2002a. Assessment. *Longman Dictionary of Language Teaching and Applied Linguistics*. 3rd ed, 35. Harlow and London: Pearson Education Limited.
- Richards, Jack C, Richard Schmidt. 2002b. Validity. *Longman Dictionary of Language Teaching and Applied Linguistics*. 3rd ed, 575. Harlow and London: Pearson Education Limited.
- Saito, Hidetoshi. 2008. EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing*, 25: 4, 553-581.
- Saito, Hidetoshi, Tomoko Fujita. 2004. Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching Research*, 8:1, 31–54.
- Saito, Hidetoshi, Tomoko Fujita. 2009. Peer-Assessing Peers' Contribution to EFL Group Presentations. *RELC Journal*, 40: 2, 149-171.
- Sebba, Judy, Deakin Crick, Yu, G., Lawson, H., Harlen, W., & Durant, K. 2008. *Systematic review of research evidence of the impact on students in secondary schools of self and peer assessment*. I Research Evidence in Education Library series. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Available at <http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2415> , accessed December 20, 2012.
- Sengupta, Sima. 1998. Peer evaluation: I am not the teacher. *ELT Journal*, 52, 19-28.
- Shimura, Mika. 2006. Peer- and Instructor Assessment of Oral Presentations in Japanese University EFL classrooms: A Pilot Study. *Waseda Global Forum No. 3*, 99–107.
- Topping, Keith. 1998. Peer assessment between students in colleges and universities. *Review of Educational Research* 68:3, 249–76.
- Topping, Keith. 2003. Self and Peer Assessment in School and University: Reliability, Validity, Utility. In Mien Segers, Filip J. R. C. Dochy, Eduardo Cascallar (eds). *Optimising New Modes of Assessment: In Search of Qualities and Standards*. Kluwer Academic Publishers.
- Topping, Keith J. 2009. Peer Assessment. *Theory Into Practice*, 48:1, 20–27.
- Tsai, Chin-Chung, Sunny S.J. Lin, and Shyan Ming Yuan. 2002. Developing science activities through a networked peer assessment system. *Computers and Education*, 38, 241–52.

- Weeden, Paul, Jan Winter, and Patricia Broadfoot. 2002. *Assessment: what's in it for schools?* London: RoutledgeFalmer.
- White, Eddy. 2009. Student Perspectives of Peer Assessment for Learning in a Public Speaking Course. *Asian EFL Journal – Professional Teaching Articles*. 33:1, 1-55.
- Yinjaroen, Pilantana , Thanyapa Chiramanee. 2011. Peer Assessment of Oral English Proficiency. *The 3rd International Conference on Humanities and Social Sciences April 2, 2011 Faculty of Liberal Arts, Prince of Songkla University Proceedings-Factors Affecting English Language Teaching and Learning*. Available at <http://sv.libarts.psu.ac.th/conference5/proceedings/Proceedings3/article/2pdf/001.pdf> accessed March 20, 2012.

APPENDIX 1 Pre Peer Assessment Survey

1. Personal data:

Boy ☐ Girl ☐ Age _____ Form _____

2. Have you ever experienced peer assessment in your English classes before?

Kas oled varem inglise keele tundides klassikaaslaste vahelise vastastikuse hindamisega tegeleenud?

- a) Yes
- b) No
- c) I don't know

3. Have you ever participated in peer assessment in other subjects?

Kas oled tegeleenud vastastikuse hindamisega teistes õppeainetes?

- a) Yes
- b) No
- c) I don't know

If your answer was „Yes“, name the subjects, please. (*Kui vastasid jaatavalt, siis palun nimeta need õppeained.*) _____

4. Indicate your opinion about peer assessment on the scale from 1 to 5.

Mil määral nõustud antud väidetega vastastikuse hindamise kohta. Märki oma arvamus skaalal 1-5 (1 – ei nõustu üldse; 5 – nõustun täielikult).

	1	2	3	4	5
	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
1. PA helps me to understand how to make a good oral presentation. <i>VH (Vastastikune hindamine) aitab mul mõista, kuidas teha head suulist ettekannet.</i>					
2. PA helps to improve my ability to identify strengths and weaknesses in my classmates' presentations. <i>VH aitab parandada minu oskust välja tuua klassikaaslaste ettekannete tugevusi ja nõrkusi.</i>					
3. PA helps to improve my English skills. <i>VH aitab kaasa minu inglise keele oskuse paranemisele.</i>					
4. PA makes me understand my teacher's requirements better. <i>VH aitab mul paremini mõista õpetaja nõudmisi.</i>					
5. The fact that my classmates assess me motivates me to work harder on my presentation. <i>Teadmine, et klassikaaslased mind hindavad, motiveerib mind oma ettekande kallal rohkem tööd tegema.</i>					

	1	2	3	4	5
6. PA stimulates me to concentrate more carefully on my classmates' presentations. <i>VH paneb mind hoolikamalt klassikaaslaste ettekannetele keskenduma.</i>					
7. I believe that the majority of my classmates' comments will be useful and I can benefit from them. <i>Ma usun, et enamus minu klassikaaslaste kommentaare on kasulikud ja ma saan nendest õppida.</i>					
8. I believe that the majority of my classmates are able to assess my presentation according to the rating scale and give me the mark that I have deserved. <i>Ma usun, et enamus minu klassikaaslaste on võimelised hindama minu ettekannet vastavalt hindamisskaalale ja panevad mulle hinde, mille olen ära teeninud.</i>					
9. Most of my classmates are able to assess their peers objectively. <i>Enamus minu klassikaaslaste on võimelised hindama oma kaaslaste erapooletult.</i>					
10. I am able to assess my classmates objectively. <i>Mina olen võimaline hindama oma klassikaaslaste erapooletult.</i>					
11. I feel uncomfortable to give negative comments to my classmates. <i>Ma tunnen ennast ebamugavalt, kui pean oma klassikaaslastele negatiivseid kommentaare kirjutama.</i>					
12. It is difficult for me to decide how to comment on my classmates' presentations. <i>Mul on raske otsustada, kuidas oma klassikaaslaste ettekandeid kommenteerida.</i>					
13. It is difficult for me to decide which marks to give to my classmates for their presentations. <i>Mul on raske otsustada, milliseid hindeid oma klassikaaslastele nende ettekannete eest panna.</i>					

APPENDIX 2 Post Peer Assessment Survey

1. Personal data:

Boy ☐ Girl ☐ Age _____ Form _____

The average mark that your classmates gave for your presentation _____

Keskmine hinne, mille klassikaaslased sinu ettekandele andsid.

2. Indicate your opinion on the scale from 1 to 5. (*Mil määral nõustud antud väidetega vastastikuse hindamise kohta. Märgi oma arvamus skaalal 1-5 (1 – ei nõustu üldse; 5 – nõustun täielikult).*)

	1	2	3	4	5
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
2. PA helped me to understand how to make a good oral presentation. <i>VH (Vastastikune hindamine) aitas mul mõista, kuidas teha head suulist ettekannet.</i>					
2. PA helped to improve my ability to identify strengths and weaknesses in my classmates' presentations. <i>VH arendas minu oskust välja tuua klassikaaslaste ettekannete tugevusi ja nõrkusi.</i>					
3. PA made me understand my teacher's requirements better. <i>VH aitab mul paremini mõista õpetaja nõudmisi.</i>					
4. PA helped to improve my English skills. <i>VH aitab kaasa minu inglise keele oskuse paranemisele.</i>					
5. The fact that my classmates assessed me motivated me to work more hard on my presentation. <i>Teadmine, et klassikaaslased mind hindasid, motiveeris mind oma ettekande kallal rohkem tööd tegema.</i>					
6. PA stimulated me to concentrate more carefully on my classmates' presentations. <i>VH pani mind hoolikamalt klassikaaslaste ettekannetele keskenduma.</i>					
7. The majority of my classmates' comments were useful and I benefitted from them. <i>Enamus minu klassikaaslaste kommentaare olid kasulikud ja ma sain nendest õppida.</i>					

	1	2	3	4	5
8. The majority of my classmates were able to assess my presentation according to the rating scale and gave me the mark that I deserved. <i>Enamus minu klassikaaslasti onlidvõimelised hindama minu ettekannet vastavalt hindamisskaalale ja panid mulle hinde, mille olin ära teeninud.</i>					
9. Most of my classmates were able to assess their peers objectively. <i>Enamus minu klassikaaslasti olid võimelised hindama oma kaaslasti erapooletult.</i>					
10. I was able to assess my classmates objectively. <i>Mina olin võimaline hindama oma klassikaaslasti erapooletult.</i>					
11. I felt uncomfortable to give negative comments to my classmates. <i>Ma tundsin ennast ebamugavalt, kui pidin oma klassikaaslastele negatiivseid kommentaare kirjutama.</i>					
12. It was difficult for me to decide how to comment on my classmates' presentations. <i>Mul oli raske otsustada, kuidas oma klassikaaslaste ettekandeid kommenteerida.</i>					
13. It was difficult for me to decide which marks to give to my classmates for their presentations. <i>Mul oli raske otsustada, milliseid hindeid oma klassikaaslastele nende ettekannete eest panna.</i>					

3. Which type of peer assessment do you prefer:

Millist vastastikuse hindamise viisi eelistad:

- a) giving marks to my classmates (*klassikaaslastele hinnete panemine*)
- b) writing comments to my classmates (*klassikaaslastele kommentaaride kirjutamine*)
- c) both of them (*mõlemaid neist*)
- d) neither of them (*mitte kumbagi neist*)

4. Which aspects were the most difficult to assess:

You can circle several aspects.

Milliseid aspekte oli kõige raskem hinnata. Võid mitmele variandile ringi umber tõmmata.

- a) content (*sisu*)
- b) organization (*ülesehitus*)
- c) grammar, vocabulary, sentence structure (*grammatika, sõnavara, lauseehitus*)
- d) pronunciation (*hääldus*)
- e) voice control and fluency (*hääle valitsemine ja soravus*)
- f) visual props (*visuaalsed abivahendid*)
- g) body language (*kehakeel*)

5. Would you like to practise peer assessment in your English classes in the future, too?

Kas sa sooviksid ka edaspidi kasutada inglise keele tundides vastastikust hindamist?

- a) Yes
- b) No
- c) I don't know

Why? (You can answer in Estonian)

6. Which of these assessment methods do you prefer in your English classes:

Milliseid hindamisviise sa oma inglisekeele tundides eelistad:

- a) teacher assessment (*õpetajapoolne hindamine*)
- b) peer assessment (*klassikaaslaste vastastikune hindamine*)
- c) the combination of teacher and peer assessment (*õpetajapoolse ja vastastikuse hindamise kombinatsiooni*)?

Why? (You can explain in Estonian) _____

APPENDIX 3 Training Materials

PEER ASSESSMENT IN ENGLISH CLASSES

WHAT IS PEER ASSESSMENT?



Usually teacher is the one who comments on and gives marks for the students' works, but peer assessment means that, instead of the teacher, the students themselves assess each other's work.

You can give grades to your classmates for their presentations, writings, projects etc. as well as provide comments on what was good about their work and what should have been done differently.

HOW TO ASSESS?

Together with your teacher you are going to create a rubric which tells you which aspects to assess and how to decide which mark or how many points to give. You have also a chance to practise peer assessment before assessing your classmates' presentations as your teacher is going to show you two videos in which students give their presentations.

HOW TO GIVE CONSTRUCTIVE FEEDBACK?

When you criticize your classmate's work your comments have to be **constructive**. It means that if you say that one's presentation was not good, you have to give examples of why you think so; what was wrong with the presentation and what should have been done differently. If you think that your classmate's presentation was good, you have to explain what was good about it.

EXAMPLES:

– Unconstructive comment: *"I did not understand."*

+ Constructive comment: *"I did not understand your presentation very well because you were reading your presentation from the paper too fast and too quietly. You should not read but have an eye contact with the audience. Next time speak in a louder voice."*

PRACTISE! Mark (+) behind the comments that are constructive and (-) behind the comments that are *unconstructive*.

1. I liked your presentation and I think that it was very interesting. ()
2. Your presentation was good because you spoke clearly and your sentences were logical and understandable. ()
3. You made a lot of mistakes. For example: My mother have...; I playing... ()
4. I did not understand some of your sentences. Words were not in the right order and you used incorrect tense forms. ()
5. Your presentation was not very good because you made a lot of mistakes but it was interesting, though. ()
6. It was good that you did not look at your notes too much and spoke by heart. Some of your sentences were weird and not correct, but I could understand what you wanted to say. ()

APPENDIX 4 Student Feedback Forms 1-3 for Comments**Student Feedback Form 1**

Name of the student assessed _____

Remember to give **constructive** feedback! Add **examples**! Comment on the following aspects:

Content

Organization of the presentation

Visual props

Other comments or suggestions _____

Student Feedback Form 2

Name of the student assessed _____

Remember to give **constructive** feedback! Add **examples**! Comment on the following aspects:**Grammar, vocabulary and sentence structure**

Body language

Other comments or suggestions _____

Student Feedback Form 3

Name of the student assessed _____

Remember to give **constructive** feedback! Add **examples**!

Comment on the following aspects.

Pronunciation

Voice control and fluency

Other comments or suggestions _____

APPENDIX 5 Feedback for Summative Assessment

Name of the student assessed _____

Content 4 3 2 1

Organization 4 3 2 1

Grammar, vocabulary and sentence structure 4 3 2 1

Pronunciation 4 3 2 1

Fluency and Voice control 4 3 2 1

Visual props 4 3 2 1

Body language 4 3 2 1

Total score _____ / 28

Mark _____

90-100% = 25-28p. = „5“

75-89% = 21-24p. = „4“

50-74% = 14-20p. = „3“

< 49% = < 13p. - = „2“

APPENDIX 6 Guidelines for Oral Presentation

Guidelines for Oral Presentation

Topic of the presentation: “My Dream Travel Destination”

Length: **2-3 minutes**

Choose a country or a place/ city/ town/ village etc. that you would like to visit and introduce it to your classmates.



Your presentation should consist of the following parts:

1. **A greeting and an opening sentence.** (*Dear classmates/ listeners....; Good afternoon, I would like to speak about Paris which is my dream travel destination; I am going to speak about...; I would like to introduce you; I have chosen to speak about Paris as my dream travel destination.)*
2. **An introduction**, which contains some general information about the country/city (*location, size, population etc.*)
3. **Main body** where you name and **explain shortly** at least **three reasons** why you would like to visit the place you have chosen to present. (*I would like to visit Paris due to the following reasons.../ There are several reasons why I would like to travel to Paris.../ Firstly.../ Secondly.../ Thirdly.../ Last but not least...;*)
4. **A conclusion**, where you shortly once again name all the aspects you spoke about. (*In conclusion.../ To sum it up...; x, y, z are the reasons why I would like to visit Paris; Paris is my dream travel destination because (of)*)
5. **Thanking the audience.**
6. As you can carry out your presentation in the computer class, you should show at least three **photos/ slides** about the places, sights that you are speaking about.

Remember!

- Do not read from the paper but try to talk by heart.
- Have an eye contact with the audience.
- Do not speak either too fast or too slowly.
- Speak loud enough.
- Don't worry, be happy! ☺

Good luck!

APPENDIX 7 Rubric for Assessing EFL Oral Presentations

Aspects to be assessed	4	3	2	1
Content	All the aspects given in the instruction are covered.	One of the aspects just mentioned but not discussed thoroughly enough.	Two of the aspects just mentioned but not discussed thoroughly enough.	None of the aspects discussed thoroughly enough, just mentioned OR one or more aspects missing entirely.
Organization	Has an appropriate introduction, main body, and conclusion. Different parts of the main body are logically organized.	Introduction OR conclusion not appropriate.	Both introduction AND conclusion not appropriate OR different parts of the main body not logical.	Separate sentences not a connected text.
Grammar, vocabulary and sentence structure	Grammar, vocabulary and sentence structure correctly used throughout the entire presentation. A few minor mistakes allowed.	Grammar, vocabulary and sentence structure mostly correct, although there are a number of minor grammar mistakes, as well some mistakes in sentence structuring.	Numerous mistakes in grammar and sentence structure. Some mistakes in the use of vocabulary. Nevertheless, it can be understood what the speaker wants to say.	So many mistakes in grammar, vocabulary and sentence structure that it is often impossible to understand what the speaker wants to say.
Pronunciation	Pronunciation, intonation, word and sentence stress are correct.	Pronunciation, intonation, word and sentence stress are generally correct. Some pronunciation mistakes and/or unnatural intonation are allowed.	Numerous mistakes in pronunciation, intonation and/or stress but the speaker can still be understood.	Numerous mistakes in pronunciation, intonation and/or stress make it very hard to understand the speaker.
Voice control and fluency	Volume is loud enough throughout the presentation. Speaks clearly all the time. Speech is fluent, no long pauses of thought. Does not read from the paper.	Volume is loud enough, only once or twice a word or phrase is not heard. Speaks clearly most of the time, only 1-2 phrases/ words cannot be understood. Speech is not completely fluent; there are some pauses of thought. Sometimes looks at the notes but does not read aloud from the paper.	A number of times the volume is too soft. Partly speaks clearly and distinctly but partly mumbles and cannot be understood. Many long pauses OR reads more from the paper than speaks to the audience.	Most of the time volume too soft to be heard. Often mumbles or cannot be understood. Many long pauses OR reads the whole presentation from the paper.
Visual props	Uses slides, photos, etc. which go together with the topic and are not distracting.	Uses slides, photos, etc. which go together with the topic, but may be slightly confusing or shown in the wrong order.	Includes too few visual props.	Uses no visual props such as slides, photos, etc. OR props do not go together with the topic.
Body language	Stands up straight, looks relaxed and confident. Has eye contact with the listeners.	Stands up straight and has eye contact with the listeners	Sometimes stands up straight and has eye contact.	Slouches and/or does not look at people during the presentation

APPENDIX 8 Rubric for Assessing EFL Oral Presentations (in Estonian)

Aspects to be assessed	4	3	2	1
Content <i>Sisu</i>	Käsitleb ettekandes kõiki tööjuhendis väljatoodud teema alapunkte.	Üks alapunktidest lihtsalt mainitud, kuid mitte käsitletud piisavalt põhjalikult.	Kaks alapunkti lihtsalt mainitud, kuid mitte käsitletud piisavalt põhjalikult.	Ühtegi alapunkti pole käsitletud piisavalt põhjalikult, lihtsalt mainitud VÕI üks või rohkem alapunkte puudub täielikult
Organization <i>Ülesehitus</i>	On asjakohane sissejuhatus, põhiosa ja kokkuvõte. Ettekande ülesehitus on loogiline.	Sissejuhatus VÕI kokkuvõte ei ole asjakohane.	Mõlemad, nii sissejuhatus kui kokkuvõte, ei ole asjakohased VÕI põhiosa erinevad alaosad ei ole loogiliselt organiseeritud.	Üksikud laused, mitte seotud tekst.
Grammar , vocabulary and sentence structure <i>Grammatika, sõnavara, lauseehitus</i>	Korrektne grammatika, sõnavara ja lauseehituse kasutamine läbi kogu ettekande. Mõned väiksed vead lubatud.	Grammatika, sõnavara ja lauseehitus enamjaolt korrektsed, siiski esineb mitmeid väikseid grammatika ja/või lauseehituse vigu.	Sagedased vead grammatikas ja lauseehituses. Vääratused sõnavara kasutamises. Vaatamata sellele on siiski võimalik aru saada, mida kõneleja öelda tahab.	Nii palju vigu grammatikas, sõnakasutuses ja lauseehituses, et on sageli võimatu aru saada, mida kõneleja öelda tahab.
Pronunciation <i>Hääldus</i>	Hääldus, intonatsioon, sõna- ja lauserõhud on korrektsed.	Hääldus, intonatsioon, sõna- ja lauserõhud on üldiselt korrektsed. Mõned hääldusvead ja /või ebaloosulik intonatsioon ei sega arusaamist.	Palju hääldus-, intonatsiooni ja/või rõhuvigu, aga kõnelejast on siiski võimalik aru saada.	Arvukad hääldus-, intonatsiooni ja rõhuvead muudavad kõnelajast aru saamise väga raskeks.
Voice control and fluency <i>Hääle valitsemine ja soravus</i>	Hääl on piisavalt tugev kogu ettekande vältel. Räägib selgelt. Kõne on sorav, ei ole pikki mõttepause. Ei loe paberilt maha.	Hääl on piisavalt tugev, ainult ühel või kahel korral ei ole sõna või fraasi kuulda. Räägib enamjaolt selgelt, kuid 1-2 fraasist/ sõnast ei ole arusaada. Kõne ei ole täiesti sorav, esinevad mõttepausid. Vahel vaatab märkmeid aga ei loe maha.	Hääl on korduvalt liiga vaikne. Osaliselt räägib selgelt, kuid osaliselt pobiseb ja kõnest ei ole võimalik aru saada. Palju pikki pause VÕI loeb pigem paberilt maha kui kõneleb kuulajatele.	Enamik ajast liiga vaikne hääl. Tihti pomiseb ja kõnest ei ole võimalik aru saada. Palju pikki pause VÕI loeb kogu ettekande paberilt maha.
Visual props <i>Visuaalsed abivahendid</i>	Kasutab pilte, slaide, mis sobivad teemaga kokku ja ei ole häirivad.	Kasutab pilte, slaide, mis sobivad teemaga kokku, kuid on vales järjekorras või tekitavad kuulajates segadust.	Kasutab liiga vähe pilte, slaide.	Ei kasuta üldse pilte, slaide VÕI kasutab pilte, slaide, mis ei ole teemaga seotud
Body language <i>Kehakeel</i>	Seisab sirgelt, näib rahulik ja enesekindel. Omab silmsidet kuulajatega.	Seisab sirgelt ja omab silmsidet kuulajatega.	Aeg-ajalt seisab sirgelt ja omab silmsidet kuulajatega.	Kehahoiak on lodev ja ei vaata ettekande ajal kuulajate poole.

APPENDIX 9 Table 13

The comparison of the mean scores of Likert items in pre- and post peer assessment survey.

Item	1 Strongly disagree % (no.of students)	2 Disagree % (no.of students)	3 Neutral % (no.of students)	4 Agree % (no.of students)	5 Strongly agree % (no.of students)	Mean	Standard deviation	Paired t-test	
								T	p
1 pre	0	14.3%	21.4%	64.3%	0	3.5	0.76	2.2804	0.0401 Sig.
post	0	(2) 0	(3) 21.4% (3)	(9) 78.6% (11)	0	3.79	0.43		
2 pre	0	0	14.3%	64.3%	21.4%	4.1	0.62	1.4720	0.1648 Not sig.
post	0	0	(2) 21.4 (3)	(9) 64.3% (9)	(3) 14.3% (2)	3.93	0.62		
3 pre	0	0	35.7%	35.7%	28.6%	3.9	0.83	0.3661	0.7202 Not sig.
post	0	0	(5) 21.4% (3)	(5) 71.5% (10)	(4) 7.1% (1)	3.86	0.53		
4 pre	0	0	14.3%	57.1%	28.6%	4.1	0.66	3.6056	0.0032 Sig. Dif.
post	0	0	(2) 50% (7)	(8) 35.7% (5)	(4) 14.3% (2)	3.64	0.74		
5 pre	0	14.3%	28.6%	35.7%	21.4%	3.6	1.0	2.6874	0.0186 Sig dif.
post	0	(2) 21.4% (3)	(4) 35.7% (5)	(5) 35.7% (5)	(3) 7.1% (1)	3.29	0.91		
6 pre	0	0	28.6%	28.6%	42.9%	4.1	0.86	2.2804	0.0401 Sig. dif
post	0	0	(4) 28.6% (4)	(4) 57.1% (8)	(6) 14.3 (2)	3.86	0.66		
7 pre	0	14.3%	35.7%	42.9%	7.1%	3.43	0.85	1.0000	0.3356 Not sig.
post	7.1% (1)	(2) 14.3% (2)	(5) 21.4% (3)	(6) 57.1% (8)	(1) 0	3.29	0.99		
8 pre	0	35.7%	0	57.1%	7.1%	3.36	1.08	0.5631	0.5830 Not sig.
post	7.1% (1)	(5) 28.6% (4)	(1) 7.1% (1)	(8) 42.9% (6)	(1) 14.3% (2)	3.29	1.27		
9 pre	0	14.3%	35.7%	50%	0	3.36	0,74	0.6939	0.5000 Not sig.
post	14.3% (2)	(2) 21.4% (3)	(5) 14.3% (2)	(7) 28.6% (4)	(3) 21.4% (3)	3.21	1,42		
10pre	0	14.3%	28.6%	35.7%	21.4%	3.64	1.0	2.6874	0.0186 Sig. dif
post	0	(2) 7.1% (1)	(4) 14.3% (2)	(5) 50% (7)	(3) 28.6% (4)	4.0	0.88		
11pre	14.3%	21.4%	42.9%	0	21.4%	2.93	1.33	0.8062	0.4346 Not sig.
post	(2) 7.1% (1)	(3) 42.9% (6)	(6) 21.4% (3)	(3) 21.4% (3)	(3) 7.1% (1)	2.79	1.12		
12pre	14.3%	28.6%	42.9%	14.3%	0	2.57	0.94	1.8829	0.0823 Not sig.
post	(2) 0	(4) 42.9% (6)	(6) 35.7% (5)	(2) 21.4% (3)	0	2.79	0.8		
13pre	14.3%	28.6%	42.9%	14.3%	0	2.57	0.94	0,0000	1.000 Not. Sig.
post	(2) 0	(4) 50% (7)	(6) 42.9% (6)	(2) 7.1% (1)	0	2.57	0.65		

RESÜMEE

TARTU ÜLIKOOL

INGLISE FILOLOOGIA OSAKOND

Kristin Laur

Using Peer Assessment in Basic School EFL Classes

Vastastikuse hindamise kasutamine põhikooli inglise keele kui võõrkeele tundides

Magistritöö

2014

Lehekülgede arv: 103

Annotatsioon:

Magistritöö keskendus vastastikuse hindamise rakendamise uurimisele põhikooli inglise keele kui võõrkeele tundides. Eesmärgiks oli leida vastused järgmistele uurimisküsimustele:

1. Kuidas suhtuvad 8. klassi õpilased vastastikuse hindamise kasutamisesse inglise keele tundides selleks, et hinnata ja kommenteerida üksteise ettekandeid?
2. Kuidas hinded, mida õpilased oma klassikaaslastelt saavad, mõjutavad nende suhtumist vastastikuse hindamise kasutamisesse?
3. Mil määral vastavad õpilaste poolt üksteisele pandud hinded õpetajate pandud hinnete?

Töö koosneb kahest osast. Esimeses peatükis tutvustatakse vastastikuse hindamise definitsioone ja võimalikke variatsioone, käsitletakse vastastikuse hindamise rakendamise plusse ja miinuseid ning antakse ülevaade juhtnööridest, kuidas vastastikust hindamist koolitunnis rakendada. Teine peatükk keskendub magistritöö raames läbi viidud uurimuse protsessi ja tulemuste kirjeldamisele ning analüüsimisele.

Uurimuse valim koosnes neljateistkümnest Rakvere Reaalgümnaasiumi õpilasest, kellel paluti hinnata üksteise suulisi ettekandeid inglise keele kui võõrkeele tunnis. Õpilased andsid oma kaaslastele tagasisidet nii punktide ja hinnete kui ka kirjalike kommentaaride kujul. Enne, kui õpilased üksteist hindama asusid, läbisid nad 3 x 45-minutilise treeningperioodi.

Andmekogumismeetoditena kasutati küsitlusi, mida õpilased täitsid enne ja pärast vastastikuse hindamise protsessi, ja õpilaste ning õpetajate poolt täidetud vastastikuse hindamise tagasiside lehti kõigi ettekannete kohta.

Toetudes uurimuse tulemustele on võimalik väita, et 8. klassi õpilased suhtusid üldjoontes positiivselt vastastikuse hindamise kasutamisesse inglise keele kui võõrkeele tundides ning leidsid, et sellisel hindamismeetodil on mitmeid plusse. Suurem osa õpilastest pidas oma klassikaaslast ja ka endid objektiivseteks hindajateks. Uurimuse tulemused viitasid ka sellele, et oma klassikaaslastelt madalaima hinde saanud õpilased olid vastastikuse hindamise suhtes negatiivsemalt meelestatud, kui need, kes said kõrgeima hinde. Lisaks sellele ilmnes õpilaste ja õpetajate antud hinnete ja punktide võrdlemisel, et 8. klassi õpilased on võimelised hindama oma kaaslast nii nagu õpetajad, sest uurimuses osalenud õpilaste ja kahe õpetaja antud punktid ja hinded olid väga sarnased. Samuti jõuti õpilaste kommentaare analüüsides järeldusele, et üksteisele konstruktiivse tagasiside andmist oleks pidanud põhjalikumalt õpetama ja harjutama, kui seda tehti antud uurimuse käigus, sest paljud kommentaarid ja soovitusid jäid liiga lihtsakoelisteks. Vaatamata sellele, et enamik õpilasi oli vastastikuse hindamise suhtes positiivselt meelestatud ning veendusid, et nende poolt antud hinded olid väga sarnased õpetajate hinnangutega, vastasid pooled valimi õpilastest, et tulevikus eelistaksid nad inglise keele tundides siiski ainult õpetajakeskset hindamist.

Märksõnad: inglise keele didaktika, inglise keele õpetamine, vastastikune hindamine

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina _____ Kristin Laur _____

(*autori nimi*)

(isikukood: _____ 48211170253 _____)

annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose _____

_____ *Using Peer Assessment in Basic School EFL Classes* _____

(*lõputöö pealkiri*)

mille juhendaja on _____ Ülle Türk _____,

(*juhendaja nimi*)

reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni; üldsusele kättesaadavaks tegemiseks ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus _____ 10.05.2014 _____ (*kuupäev*)

_____ Kristin Laur _____

(*allkiri*)