



Uit

NORGES
ARKTISKE
UNIVERSITET

Faculty of Humanities, Social Sciences and Education

Peer Assessment in EFL Classrooms

A Qualitative Approach to Implementing Peer Assessment in Oral Communicative Learning Activities in EFL Classrooms.

—

Hauk Andreas Perander

Master Thesis Eng-3980 English Language and Education. November 2017



Innhold

Acknowledgements	III
A Qualitative Approach to Implementing Peer Assessment in Communicative Learning Activities.....	5
1. Introduction	5
2. Theoretical Background and Terminology.....	8
2.1 What is Peer Assessment?.....	8
2.2 Background of Peer Assessment.	8
2.3 Formative assessment.....	8
2.4 The Role of Feedback.....	9
2.5 Metacognition.....	10
2.6 Providing Examples and Modelling	10
2.7 Implementing Peer Assessment.....	11
2.8 Importance of Group Dynamics.	13
2.9 Possible Problems Related to Peer Assessment.	14
2.10 But Why Peer Assessment?.....	14
2.11. But What Is Considered Oral Interaction?	15
3. Methodology.	17
3.1 Qualitative method	17
3.2 Selection of Test Subjects.	17
3.3 The Role of the Researcher.	18
3.4 Staging Oral Interaction with Formal Criteria.....	18
3.4.1 Selecting ways of communication.....	19
3.4.2 Connection Between Marine Radio Communication and Oral Interaction Skills.....	20
3.5 Spontaneous Interaction or Prepared Production?.....	21
3.6 The timeline of the project.	22
3.7 Data gathering.	25
3.8 Group Compositions.....	25
3.7 Limitations and Shortcomings.....	26
3.7.1 Lack of Control Group.	26
3.7.2 Possible Exposure to the Scenario Prior to Testing.....	27
3.7.3 Possibility of Feedback Received Orally.....	27
3.7.4 Different Evaluation Setting Compared to Training.	27
3.7.5 Limited Number of Test Subjects.....	28
3.7.6 Insufficient Prepared Assessors.....	28

4. The Study.....	29
4.1 Results.....	29
4.2 Conveyed Messages and Response of Group 1.....	30
4.2.1 Test subject 1.....	30
4.2.2 Test subject 2.....	32
4.2.3 Test Subject 3.....	34
4.2.4 Test Subject 4.....	36
4.3 Examination of group 1.....	38
4.3.1 Perceptual Mistakes.....	40
4.3.2 Inability to Concretise Feedback.....	40
4.3.3 Summary of Group 1.....	40
4.4 Conveyed Messages and Response of Group 4.....	41
4.4.1 Test Subject 12.....	41
4.4.2 Test Subject 14.....	43
4.4.3 Test Subject 13.....	45
4.5 Examination of Group 4.....	47
4.5.1 Summary of Group 4.....	51
4.6 Summary of the Results.....	52
4.8 Perception.....	53
4.9 Lack of Motivation and Miscalculated High Judgment of Learning.....	53
4.10 Inflating Social Power Structures.....	54
5. Conclusion.....	55
5.1 Aspects Worthy of Implementation in Future Studies.....	56
5.2 Closing Remarks.....	56
Bibliography.....	57
6.1 Appendix 1. Evaluation Sheet.....	59
6.2 Appendix 2. Scenarios Used During the First Test Round.....	60
6.3 Appendix 3. Scenarios Used During the Second Test Round.....	60

Acknowledgements

I would like to thank my supervisor, Annelise Brox Larsen, for constructive feedback and support. You have truly encouraged and motivated me throughout the making of this thesis, and I am grateful for that.

Furthermore, I would like to thank the participants who have made the entire process interesting to behold and be a part of. Thanks is also to be given to my dear colleagues who have contributed with their experience and thematic insight.

Lastly, I would like to thank all of my peers, who has contributed in numerous ways over the last five years. It is your help that has sparked my enthusiasm for peer assessment as a concept.

Tromsø 2017

A Qualitative Approach to Implementing Peer Assessment in Communicative Learning Activities.

1. Introduction

In any given learning situation, the student or pupil needs to be able to identify the goal of the learning process. In many cases the learner is not able to pinpoint exactly what is required to improve his or her in a specific subject or skill. A method to achieve better performances from the learner is by implementing formative assessment, a situation where the learner gets feedback based on his current performance in order to polish the final product. Traditionally, this feedback is provided by teachers and supervisors as they are perceived as experts with formal competence in that given subject. However, Keith Topping, professor of Educational and Social Research at the University of Dundee, refers to a number of studies that suggest that the pupils themselves are able to provide reliable feedback to their peers (Topping 2009:24), if trained to do so. Furthermore, peer assessment is reported to give learners better insight into criteria required to enhance their performance in a given subject, as they act as both assessor and assessee. Trude Slemmen (2009), senior advisor at the Norwegian Directorate of Education, state this by highlighting that students need to gain insight into criteria relevant to the task at hand, which benefits the assessor in relation to his or her own performance. In addition, Slemmen emphasizes that by training students into assessing each other, response and feedback are more accessible. In learning a foreign language, a situation where the learners are at the same stage of the learning process and has somewhat the same level of competence, continuous feedback would be quite useful. The primary research question for this master thesis will be: Can peer assessment be implemented as a teaching method with aims to enhance performance in oral communication?

Many pupils and students in Norway will need to be able to communicate in English because of their occupation. This would certainly apply to students at maritime vocational studies at upper secondary school in Norway. In many cases, the staff and crewmembers these students will encounter in their future occupation do not originate from Norway, but comes from different parts of the world. As such, the language of usage on vessels is in many cases English. Additionally, future sailors might find themselves in international waters when occurrences would require them to contact nearby vessels or stations. In some of these potential situations, not being able to communicate with nearby seafaring agents could end fatally. Subsequently, it seems mandatory that the pupils attending upper secondary schooling

in marine seafaring need be adequately competent in spoken English, especially related to situations of distress.

Bjørke (2014:30) states that one of the most fundamental aspects of acquiring a new language is to be able to express oneself orally. Two essential terms that are integral to any language is production and perception. Perception could be reading or listening, while production could be spoken language or writing. Both perception and production needs to be accounted for if the communication should be defined as interaction. If a speaker receives a message from another person, and perceives the content of that message, he or she (hereby generalized to “he” or “him”) needs to evaluate the necessary parts that need to be conveyed further, or think of how he want to respond. As such, perceptual skills become the foundation of production. However, merely perceiving words and information is not guaranteeing relevant production based on the perceived input. A cognitive awareness is needed to understand the essential perceived input in order to produce a reasonable response to that specific input. Thus, it is clear that there is a connection between perception, cognition and production to successfully communicate via spoken language.

Much of the previous research on the implementation of peer assessment in teaching of English as a Foreign Language (EFL) is done with aims to improve writing such as Zhao (2014) and Burner (2015). Others have tried to use peer assessment to enhance learners’ oral performance in presentations such as Liu (2012). However, little research has been conducted on peer assessment as a method while aiming to enhance learners’ abilities for oral interaction while addressing both perception and production. Pupils attending maritime studies will need to be able to interact with fellow crewmembers, passengers, radio station staff and coastguard in certain situations, and in some of these situations there is no time to prepare a presentation or a manuscript. In this project the aim is to implement peer assessment in way that promotes the pupils’ attention and cognition towards oral communication with focus on reception, cognition and production. If the general idea that peer assessment can enhance learners’ metacognition, as well as performing better in later stages of testing, can be transferred from general peer assessment and EFL writing to EFL oral interaction performance it would suggest that pupils gain awareness to requirements of certain features of oral interaction.

Spontaneous interaction is in many ways more complicated than to present a given topic or subject. With spontaneous interaction, the speaker does not know the content of what he is about to perceive. It becomes equally important to perceive as to produce (Bjørke 2014:33). With focus on oral performance in the format of presentations the spoken language would be

prepared and rehearsed, and although it would help to develop pronunciation and vocabulary, the ability to listen and respond to an immediate message does not occur. In an emergency situation on board a vessel a crewmember would in many cases have to communicate with others in order to take care of the situation, often without substantial amount of time to their disposal. This way of communication is quite different to a rehearsed presentation at school.

In this study, a group of Norwegian vocational students' oral skills were put to the test through tasks that challenged their perceptual skills, cognitive skills and production skills. The participants were going to convey a successful urgency- or distress call, based on input from peers who read a provided scenario which contained multiple pieces of information, with limited amount of time to respond. Their peers assessed these conveyed messages before all the participants did a second round of tests. Between these two tests, they all received feedback and response from their peers that aimed at improving their performance in the communicative oral format of distress calls.

2. Theoretical Background and Terminology

The following section aims to establish an overview of peer assessment and its traits. Peer assessment is a complex learning activity that requires numerous features to be accounted for. There are additional terms used in this section which includes to broach this complexity, namely metacognition, formative assessment, peer modelling and the role of feedback. Common pitfalls and issues related to peer assessment will be accounted for in this section.

2.1 What is Peer Assessment?

Topping defines 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” (2009:20). In essence, this will create a situation where the teacher’s feedback is not the only response received by the learners. The feedback can be both formative and summative, but the main focus of this project will be based on formative assessment, as the aim of the project is to identify learners’ ability to give constructive feedback and strengthen their ability to provide this sort of assessment.

2.2 Background of Peer Assessment.

Peer assessment and the ideas of Keith Topping are based on the perspectives of Leo Vygotskij in relation to teaching and learning. The general idea of Vygotskij’s social constructivist approach to teaching is that learning and development is primarily a result of social interaction (Imsen 2005:261). The learner does not implicitly develop their skills, but the learner has to receive explicit response in interaction in order to further increase his or hers skills. However, there are conditions that this responsive interaction require for the learning to take place. Vygotskij introduced the idea of the Zone of Proximal Development (Imsen 2005:258). The ZPD suggests that there are limitations to what a learner can learn by himself, and that external feedback and help is needed to further expand the competence of the learner. To reach the next level of performance the learner needs to have a partner who is capable of providing guiding questions and give constructive feedback. This requires the assessor or partner to have a set of certain abilities and competence. In the initial stages of life, this external response is provided by parents, as well as by teachers in classroom situations. The very essence of peer assessment is to enable peers to be able to provide constructive feedback to one another.

2.3 Formative assessment.

As mentioned, the aim for the assessment provided by the peer was for it to be formative. Formative assessment aims to strengthen certain aspects of the product or performance of a

learner (Topping 2009:20). Without feedback the learner would find it difficult to enhance his or her chances of achieving a better score in a given topic. Formative assessment aims to identify to the learner what the next step should be in order to improve his performance. Paul Black defines an assessment activity as formative if “it can help learning by providing information to be used as feedback, by teachers and by their students, in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged” (Black 2011:74). It is important to note that although the intention of the feedback is to modify the learning activities, it does not necessarily mean that it is formative. It requires the response to be understandable and concise to the learner, and relevant to the task at hand.

2.4 The Role of Feedback.

Hattie and Yates state that students express that feedback should be focusing on how they can improve their work so that they can do better in the future (2014:64-65). Hattie and Yates further emphasize that students have a rather negative attitude towards critique, especially if the students have put an effort in performing well. However, if the performances of the students are not without errors, some sort of correction and guiding will be needed in order to regulate the work of the students. Students’ reluctance to receiving critique and the need for corrections creates a balance that can be hard to account for in terms of providing feedback. Hattie and Yates (2014:65) also claim that negative feedback is stronger than positive feedback. In essence this would mean that feedback aiming to improve aspects of students performance tend to overshadow the positive feedback. This highlights the need for positive feedback, and that learners are emotionally involved in the learning process.

There are four different variants of feedback and response according to Hattie and Timperley (2007). The feedback can be based on the products itself, the strategies used, self-regulation and lastly the learner as an individual. The first category of feedback highlights aspects of a product that is good in addition to aspects that needs improvement. This category is the one that was mostly focused on in this study in terms of methodology.

It is clear that feedback and response have an important role in assisting learners. However, Hattie and Yates (2014:66) claim that there are significant variations in the effectiveness of feedback. They state that “Feedback works because the goal is known and accurately defined through realistic assessment” (Hattie and Yates 2014: 66). This suggests that learners need to have knowledge of the goal in order to be able to achieve that goal, and that feedback is a vital part of reaching it. This means that learners needs to know what is expected of them before the learning process takes place. If not, the feedback provided will not be constructive

as it has to be based on the learner's current achievement of the goal. This notion coincides with Keith Topping's emphasize on the value of preparation (Topping 2009: 25), especially in terms of involving the learners in the process of creating criteria of assessment. Ultimately, the feedback needs to be based on the students' insight into the goal of the project. Trude Slemmen (2009) concretize this by providing three questions that feedback needs to answer. The first question "Where am I going?" which refers to identifying the goal. The second question, "Where am I", aims to locate to what extent the learner has reached his or her goal at the moment of the assessment. The last and final question is "What's the next step?" which aims to identify the feature that would stand to improve the product the most (Slemmen 2009:188). To have insight into desired goal of a project is of primary interest in peer assessment, as the peers will take the role of assessors. Without insight into the goal, giving constructive feedback that aims to guide the learner becomes hard, if not impossible.

2.5 Metacognition.

An important aspect of self- and peer assessment is to increase the level of self-awareness and self-regulation. Metacognition is in essence the ability to think about your own thinking (Krumsvik and Säljö 2013:120). Metcalfe and Kornell (2004) branched the idea of judgements of learning (JOL). JOL refers to the learner's own ability to evaluate whether or not his or her performance is up the standards of the desired goal. If the JOL is high, the learner could experience a lack of motivation to continue working with the project, as he feels adequate competent in that specific area. However, if the learners experience that their performance is not up to the desired standards of the goal, this should indicate to the learner that they need to improve certain aspects of their performance. If this knowledge gap seems too large, it could make learner feel disillusioned.

Each participant took the role of both assesse and assessor in this project. In order to successfully provide constructive feedback to a peer, it is necessary to gain knowledge of how to guide that peer in the right direction in terms of achieving the goal. With the added knowledge of how to approach the task, the idea is that the participants have a degree of self-awareness to their own performances. In essence, this would mean that the performances of the participants should benefit from the work they conducted as assessors.

2.6 Providing Examples and Modelling.

An important method that can be implemented in order for students to gain increased insight in how the goal of a product should look like is providing examples. Fjørtoft (2014:18), who researches on assessment, describes modelling as a method in which the teacher does not only

tell the learners how they should solve a problem or task, but solves the task in the presence of the students. When introducing a new goal to learners, it is of utmost importance that they gain knowledge of what they are going to learn, as insight into the goal of the project is needed. Simply put, you cannot hope to reach a certain place if you do not know where that place is or what it would look like.

The teacher is not always responsible for providing examples in terms of modelling. Topping and Ehly (2009:117) claim that students could benefit from watching examples provided by peers, rather than teachers. In many instances, students could experience that the exemplifications provided by the teacher could seem unattainable. With peer modelling on the other hand, the students would experience first-hand that the perceived level of competence is achievable to those at their level. Topping and Ehly further emphasize that peer models in most cases are not perfect, but could be quite competent. However, with being exposed to multiple answers to the same task, students could use aspects of other students' products of performances to enhance their own ones. Peer assessment is based on students' assessment of peers' performances or products. It does not necessarily mean that they copy the direct performance of their peers, but with an awareness to the required criteria it could cause the observers to assess features of the observed performance to be worthwhile to implement into their own performances.

2.7 Implementing Peer Assessment.

Implementation of peer assessment requires substantial planning and preparation. According to Topping (2009:25-26) there are 11 aspects that needs to be accounted for if peer assessment is to be successfully implemented. The following subsection sets the premises for the methodology of the study.

1. Topping's (2009:25) first tip is to work with colleagues rather than developing the initiative of peer assessment alone. Peer assessment is quite complex and takes into account several cognitive and social aspects which are quite tedious to keep track of. Therefore, to involve colleagues and receive feedback in the planning phase is of great importance.

2. The second point of interest is that the projects aim has to be clarified before the planning process. Peer assessment can have cognitive, attitudinal, social or emotional gains (Topping 2009: 25). The group of learners should be involved in this stage of planning, as it should be at all times clear what the purpose of the project is. It should also be specified what parts of the learning process are to be assessed.

3. Thirdly the participants should be involved in both developing and distinguishing the assessment criteria (Topping 2009:25). The involvement of criteria development is important as it triggers the learner's ownership towards the criteria, which could make it easier for the learners to make use the criteria when assessing peers. In many instances the criteria suggested by the learners do not differ much from what the teacher would have given them.

4. The next phase would be to match the participants together in pairs or smaller groups. The general idea is that same-ability participants should be paired together (Topping 2009: 25). The low achievers of the subject of assessment should have the teacher's focus during the assessment process.

5. In the fifth stage of implementation of peer assessment, it is important to provide examples, training and practise (Topping 2009: 25). To provide quality training could make the response the peers give each other more reliable. In this stage it should be raised awareness towards the role and behaviours of both assessor and assessee. Exemplifications of constructive feedback could help the learners identifying the desired goal. In this stage the teacher should continuously evaluate the learner's response abilities. The assessors' ability to provide constructive feedback is crucial, thus giving the teacher the important task of assessing and guiding the students to be able to provide constructive response.

6. Topping additionally highlights the importance of guidelines, checklists and other tangible scaffolding tools. Some kind clues or reminders would be beneficial to the assessors, but could additionally be relevant for the assessee. A simple sheet should provide a few reminders on what to do, and how to do it.

7. Additionally, the timescale of the different activities needs to be addressed (Topping 2009: 25-26). Peer assessment requires that all the participants within one group or pair is at the same stage of the project, which sets some requirements to when certain activities should be conducted. This need to be clarified for the participants. In some cases, some of the learners are finished with one stage of the project, while others are not. This needs to be addressed so that every group could at all times move further into the project.

8. During the stage in which the participants provide feedback and response to each other, the teacher needs to provide feedback on how the participants manage their roles as assessor or assessee (Topping 2009:26). The teacher should keep a low profile and only intervene if feedback or coaching is needed.

9. The next stage involves examination of the peer feedback (Topping 2009: 26). This should be done for the whole scale of achievers. If the peer feedback is not adequate this need to be addressed with the assessor. In oral communication this would be to listen to the recorded interactions and investigate if the response correlates to the oral performance.

10. During the feedback stage the teacher needs to keep track of the reliability and validity of the feedback (Topping 2009:26). If more than one peer assesses the same performance, there should be a connection between the two responses. If not, this needs to addressed.

11. The feedback the peers provide needs to be evaluated (Topping 29:26). Without information about their performance as assessors they are not able to provide useful feedback.

2.8 Importance of Group Dynamics.

As previously mentioned, Topping claims that the matching of participants should be based on the participants level of achievement. This would suggest that the groups should be homogenous in relation to level of achievement. There are however, other opinions to what creates the most beneficial group dynamics. Other researcher, such as David W. Johnson, claims that group projects benefit from heterogeneity (Johnson 2006:48). Other, such as Dr. Rebecca Wing-yi Cheng (2008), a researcher at The Hong Kong Institute of Education, denounce that level of achievement is the most important factor when matching participants. Cheng conducted a study which examined the effects of group heterogeneity in project-based learning. The study concluded that “group heterogeneity was not a determined factor. Instead,... both high and low achievers were able to benefit when group processes were of high quality”(Cheng 2008:205), suggesting that the group dynamics are more important than matching participants of equal level of achievement. Chang argues that positive interdependence, individual accountability, equal participation and social skills are the main determiners for successful group projects. Positive interdependence refers to the participants’ need for the others to success in order for themselves to succeed. Individual accountability means that the success of the group depends on the learning of each group member (Cheng 2008:207-208). These factors greatly coincide with peer assessment, as the learners are dependent of each other. Equal participation means that the students to have an equal share of the work, whereas social skills is important to the need of trust in group discussions (Cheng 2008:208). A study by Deutsch, referred to by Johnson, concluded that reciprocity is an important aspect of cooperation, and it can occur both positively and negatively (Johnson 2006: 38). In addition, Johnson claims that a sense of belonging and commitment to the group is crucial to the willingness of helping group members (Johnson 2006:48).

2.9 Possible Problems Related to Peer Assessment.

As in any group collaboration, a group implementing peer assessment can experience negative social processes (Topping 2009:24). Internal social power structures could greatly impact both the reliability and validity of the response provided in any given project. One of the main features of peer assessment is to make the learners responsible for their own learning. If some of the participants in a group project involving peer assessment are diffused to their responsibility as assessor this could be quite severe towards the gains of their peers. In addition, both the assessor and assessee could experience anxiety in the feedback stages of the project (Topping 2009:24). For instance, if an individual with high social capital is being assessed by a person with less social power, it could lower the assessor's chances of providing constructive feedback, as constructive feedback points out features of a product or performance that needs improvement. In addition, if an individual experiences receiving feedback as uncomfortable and problematic, it is likely that the person will not provide feedback to his or her peers. This could severely limit the student's commitment towards the project. If a student who, due to reluctance to be in an evaluative situation, fails to provide concrete response to a peer, it might cause the recipient to experience a lack of commitment towards the project.

A study conducted by Sadler and Good (2006) expressed that self-assessment is more effective than peer-assessment in relation to grading papers. Sadler and Good's results lead them to believe that students at all levels of achievement benefitted from self-assessment. Peer-assessment however, showed only significant gains in the middle-level group (Sadler and Good 2006:25). This further emphasizes that lower level achievers could experience social anxiety when being exposed to peers.

Another problem with peer assessment is that the feedback does not necessarily point the learner in the right direction in terms of improving his performance. Add this to the fact that 80% of the feedback students receive is from their peers (Bjørke 2014:231) it becomes eminently clear that peer response can have damaging effects on students learning.

2.10 But Why Peer Assessment?

Despite several issues related to peer assessment, there are some reported benefits that justifies implementation. A common goal by implementing peer assessment is to give students' more insight into their own learning. Slemmen (2009) claims that peer assessments can contribute to increase the ability to evaluate and identify conditions of a proper performance or product. Furthermore, it could enable students to share and develop ideas together (Slemmen 2009: 191). Paul Black (2004) argues that peer assessment is an important

addition to self-assessment. He states that peer assessment “is uniquely valuable because students may accept criticism of their work from another that they would not take seriously if the remarks were offered by a teacher” (Black 2004:14). Generally, teachers could be reluctant to give the responsibility of assessing students work to themselves. However, peer response could contribute to the assesse’s understanding of the response, as the feedback will likely be provided in language they use themselves (Black 2004:14). Additionally, students are more likely to interrupt assessor if the feedback is not understood. This could create a discourse where both assesse and assessor could benefit from the discussion.

A common concern with peer assessment, and a reason for teachers’ reluctance to give the responsibility of assessment to the students themselves, is the reliability and validity of the feedback provided by peers. Some studies present findings which indicates that the reliability and validity is adequate, such as Sadler and Good (2006), while other studies found them variable (Topping 2009:24). It is however interesting that the number of studies which indicates a correlation between peer assessments and expert assessments are the majority. Despite this, it is noted that the evaluations made by peers tend to be related to the medium-level of achievement (Topping 2009:24). Despite this, as the findings of the majority of the studies indicates that peers are reliably able to assess each other, it becomes important to investigate this further as peer assessment is reported to give cognitive learning gains.

2.11. But What Is Considered Oral Interaction?

The general aim of the project is to investigate whether or not some of the reported benefits from peer assessment in writing can be transferred to oral learning strategies. As such, it is eminently important to define oral interaction. Camilla Bjørke explains that oral skills in language are divided into three subskills; prepared oral production, spontaneous oral interaction and listening (Bjørke 2014:30). This project aims specially to investigate if peer assessment can be used in oral interactive learning strategies, and will therefore be focusing on spontaneous oral interaction and listening. First and foremost, interaction in any language will consist of both production and perception. This can be challenging in a second language, as the speaker needs to pay attention to what is perceived while simultaneously assess how to reply in accordance to the situation and recipient (Bjørke 2014:33). As a result, the production of the speaker is directly influenced by the speaker’s perceptual skills. Without proper perception, the speaker would not cognitively evaluate how to respond to a message, which again could result in a response that is not adapted to the situation and the recipient. This definition of oral skills is further explained by the Norwegian Directorate of Education as the

institution labels oral skills as “*Oral skills* in English means being able to listen, speak and interact using the English language. It means evaluating and adapting ways of expression to the purpose of the conversation, the recipient and the situation.” (Kunnskapsdepartementet. 2013:2) In essence, oral interactive skills are based on the speaker’s ability to perceive information, cognitively evaluate the information, and provide a relevant response that is adjusted to the situation and the recipient.

3. Methodology.

The majority of this section will describe the journey from the initial broaching of the idea to the test group, to the end of the last test round. Keith Topping's list of 11 aspects that need to be accounted for in order to successfully implement peer assessment is the backbone of this journey. It will be commented on the role of the researcher, the data gathering, in addition to a disclosure of factors that could have interfered with the reliability of the data. One of the most important factors that will be explained in this section is the construction of the provided scenarios the participants received, and the general design of the study. Furthermore, as the participants were school students in the age between 17 and 18 at the time of the project there are ethical aspects of the study that will be discussed as well.

3.1 Qualitative method

With the limited test group, consisting of only 14 individuals, there are no grounds for making any conclusions based on results from this project alone. However, as Tove Thagaard (2004) writes that "An important aim with qualitative research is to gain knowledge of social processes. Interpretation is therefore of great importance to qualitative research" (Thagaard, 2004: 11). As mentioned in the theory section, social processes and power structures is an important part of group activities such as peer assessment. As such, a qualitative approach to investigate aspects of peer assessment seems reasonable. Thagaard further explains that "Qualitative studies may also focus on behaviour and interaction between humans"(2004:11). As mentioned in the theory section, peer assessment and social constructivist approaches to teaching is influenced by the learner's interaction with his or hers surroundings, which clearly suggests that a qualitative approach would be reasonable when investigating the benefits of peer assessment in oral learning strategies.

3.2 Selection of Test Subjects.

As Thagaard (2004:57) mentions, accessibility to the field is important in any research proposal. As being the teacher in English to a group of students, a sample group of 14 participants was available. These were all in the age of 17 to 18, and were all students at high school level in Norway. As such, they emerged as an ideal test group. Furthermore, being the teacher of the group for the past six month prior to the project gave greater insight into some of the power structures in the test group, as well knowledge of their proficiency in English. This would be beneficial when matching the participants into groups, a process which Keith Topping claims should be based on level of competence. In addition,

The headmaster of the school agreed to the usage of the participants in the study, under the condition that the names of the participants were omitted from the study. In addition, the participants signed a letter of consent to agree to the participation. As Thagaard (2004:23) highlights, it is important that the test subjects are informed about the aim of the project and the main features of the project. The aim of the project was explicitly conveyed to the test group, as the project was broached to them several months ahead of the start-up.

3.3 The Role of the Researcher.

It is important to highlight the researcher's role in the project. The researcher is also the EFL teacher of the test participants. This aspect creates possibilities and issues related to the project. The most prominent benefit of this could be that the researcher has deeper insight into the level of competence of the different participants, in addition to insight into certain social patterns of the test group. However, the teacher's subjective impression of the social patterns of the test group could also be wrong, and crucial power structures might be hidden from the teacher. In an attempt to remedy for this, a discussion with colleagues concerning the social power structures of the participants was implemented before the start of project. It is important to note that the colleague's opinions were subjective, but it is plausible to claim that this would account for revealing hidden social power structures in the test group.

As Thagaard (2004:79) mentions, it is important to reflect on the relation between the researcher and the test subjects. The dual role of both researcher and teacher could have caused anxiety in some of the participants. As mentioned in the theory section, many students could be anxious in performing in front of the teacher, who is continuously assessing the performances of the students. In order to prevent this from happening, it was explicitly conveyed to the test group that their performances in this project would not be graded.

3.4 Staging Oral Interaction with Formal Criteria.

The following section aims to explain the construction of the scenarios which would create the basis of task the participants were going to perform. One of the most problematic dilemmas when merging peer assessment and oral interaction is the inclusion of criteria of assessment needed in peer assessment and the complex yet informal nature of spoken language. Of course, spoken language follow the same linguistic rules as written language, but albeit to a different extent. To provide constructive and applicable feedback to peers when addressing grammatical and linguistic errors would require considerate linguistic insight from both the assessor and the assessee. As such, it would not be reasonable to make the participants talk to each other, while their peer assessed the quality of their interaction. It could occur that

the assessors would be able to identify who of the group that were high and low performers, but it is unlikely that they would be able to provide constructive feedback that the assessee could use to better his performance later on. In addition, evaluating if an oral interaction is adapted to the situation and the recipient would be based on subjectivity, if not addressing language errors. Because of this, the format of the interaction would need to include formalities and criteria that could be taught in a short period of time, while simultaneously include some of characteristics of spoken oral interaction.

3.4.1 Selecting ways of communication.

During the initial stages of the project, different formats of communication were evaluated in terms of relevance to the test group. Different recipients would require different information and different ways of communication. As a result, the criteria would need to address general ideas of contextualized interaction. What is considered the proper way of communication in different situations could be highly subjective, and thus not suitable to use as basis for setting formal criteria. As mentioned, earlier the criteria would need to be attainable for the participants, both as assessors and assessee. With a wider range of possible recipients, the criteria would need to be general. To provide concise and constructive feedback on more general criteria would set high requirements concerning the assessors' competence in English. Because of this, it was decided to use three different possible formats; pan-pan- and mayday-calls, and addressing the passengers of a vessel in a certain situation. These three formats were later limited to pan-pan and mayday-calls.

At this point it is of eminent importance that the reader gains some insight into the nature of two possible radio communication formats that were used in the project. Without knowledge of the difference between a pan-pan- or a mayday call the reader has no basis from which he or she could understand some of the observations and analysis that are founded on the performance of the participants. Mayday calls are also referred to as distress calls (Lees 2004: 34), whereas pan-pan calls are a form of urgency communication (Lees 2004:46). They differ in the way that distress calls are prioritized over urgency calls in radio frequencies (Lees 2004:47). As such, it signifies that the nature of mayday calls is more severe and possibly life threatening, whereas pan-pan calls are used when there is an urgent situation, but one that could be handled by the crew. In short, the provided scenario is a form of risk assessment. One example of the sort of risk assessment is in the case of smoke or fire in the engine room. Smoke would indicate that the engine needs attention and maintenance, whereas flames and fire in the engine would be a more severe situation.

Despite the fact that pan-pan and mayday calls are different in their level of emergency, the information required to be included in these two formats is quite alike. The formal requirements of a pan-pan- or a mayday call are the following: mayday/pan-pan, identification of the vessel, location, the nature of the distress, number of persons aboard and description of the vessel (Special Notice to Mariners 2001). In terms of the structure the formats of pan-pan and mayday-call is quite similar. Because of this, the written scenarios that created the basis for the assesses' evaluation need to include all the relevant information that is required in an emergency call. Furthermore, information regarding the nature of the distress needed to be implemented as well.

However, simply including relevant information in a written scenario would not require the participants to pay attention to anything else that the structure of the transmitted distress call. As such, irrelevant information was added as well. These bits of information needed to not be out of context. Sentences as "The captain did the dishes yesterday" would be simple to evaluate as irrelevant to the participants, and would therefore require little of their cognitive abilities. Other information, as inclusion of a forecast showing great weather for the coming weeks, would be more relevant to the context. This information would however not be relevant to include if the emergency situation occurred because of fire in the engine room.

3.4.2 Connection Between Marine Radio Communication and Oral Interaction Skills.

How can the formal requirements of distress- and urgency calls be used to investigate the benefits of peer assessment for EFL learners of English in oral communication? At this point, it is important to emphasise the earlier mentioned definition of what oral interaction skills actually are. As previously mentioned "*Oral skills* in English means being able to listen, speak and interact using the English language. It means evaluating and adapting ways of expression to the purpose of the conversation, the recipient and the situation."

(Kunnskapsdepartementet 2013:2). By using this definition, with emphasis on the importance of the situation, the project connects the formalities of marine radio communication with the ideas of perception, cognition and production in oral interaction. This required the project to be structuralised differently than a genuine emergent situation at sea, as the perception skill would be used differently. In order to fully include the aspect of perception related to oral interaction, the scenario had to be read by the assessor and not by the assessee. This would involve the listening skills of the participants, as failure to perceive aspects of the provided scenario could lead to usage of the wrong format related to the emergency. In addition, the

cognitive skills of the participants would additionally be tested, as merely perceiving the scenario does not guarantee a successful conveyed distress call. The participants would need to evaluate the input received through the scenario to produce a relevant response. As such, the project accounts for the perceptual skills when listening to the scenario, the cognitive skill when evaluating the situation, and the production skill when conveying the message.

However, the cognitive aspect was implemented as there are differences in when these emergency calls are made. Using the wrong format would indicate that the participants had not cognitively assessed the provided scenario correct. It then became obvious that the scenarios needed to include information that could indicate either of the formats being the correct one, but that with further investigation were possible to assess and evaluate to end up with the correct format. Failure to perceive all the relevant information could additionally result in a wrongly conveyed emergency call.

3.5 Spontaneous Interaction or Prepared Production?

It can be argued that by preparing criteria of assessment, the project is designed to investigate prepared production rather than spontaneous speech. This is even further backed by the fact that the production is based on a set of formal requirements both in terms of content and structure. Bjørke (2014) provides definitions of both variations of communication. In prepared production, the speaker has the opportunity to gather missing information and generally practice in advance (Bjørke 2014:33). The project certainly includes aspects of prepared production as the students were allowed to take notes, and ask the assessor to repeat aspects of the scenario. However, spontaneous interaction is described as a situation where the recipient has limited amount of time to understand the language input, while simultaneously think of relevant a response (Bjørke 2014: 33). The participants did not have the opportunity to look up information of how to convey a distress call or practice in advance, and had to respond quickly. This project falls into neither the category of spontaneous interaction or prepared production, but rather a combination of the two. To provide a formal set of requirements was necessary in order to provide a basis from which the peers could assess each other. If the formal requirements were not used as the basis of the project it would require the participants to be qualified assessors of all kinds of spoken English interaction. Although the goal of EFL is to teach students to communicate in English in many given contexts, it would be impossible based on the time limitation. Instead, different fragmentized types of English communication could be labelled as responsive communication and interaction rather than a rehearsed presentation.

3.6 The timeline of the project.

During the first stage of the project, the test group watched videos of licenced certifiers of radio operators, which showcased examples of both pan-pan and mayday calls. The participants were asked to highlight important aspects that needed to be included in a distress call. The criteria that was most notable by the participants was the need to speak slowly and concisely. In addition, the participants highlighted what fragments of the distress call that needed to be included. Together, this created the basis of the criteria of assessment, in accordance to Topping's third listed feature to implement peer assessment successfully (Topping 2009:25).

The next phase of the project had significant focus on the fifth point for successfully implementing peer assessment, which was to provide examples and training in giving feedback. The participants were asked to investigate a written scenario, in order to evaluate what sort of message that specific scenario would indicate. This was implemented as the participants needed to have an understanding of how that scenario would turn out in a transmitted distress call, before they could actually provide formative assessment to a peer. In this instance however, they were not going to evaluate each other. Based on the provided written scenario, three different conveyed messages were recorded and conveyed by the researcher, with intentional mistakes and errors in relation to the criteria of assessment. These recordings became the first assessment material the students were to evaluate, and were different both in terms of format and quality of language. This would enable them to practice providing feedback without the social anxiety that could occur while being assessed and assessing classmates. The feedback related to the recorded messages were written down in order to enable the teacher to assess the quality of the feedback, and possibly provide guiding on how to provide relevant feedback, all in accordance to Topping's emphasis on the importance of evaluating the quality of the feedback.. There were deviations in the students' ability to give constructive feedback. Some of the students gave concise response to both the way the message was conveyed, and the content of it. Others were not able to identify aspects of the message that needed improvement. Some participants assessed all three examples as flawless, although they were different interpretations of the same provided scenario. In essence, only one of these recordings used the correct format. It is also of importance to mention that none of the participants gave feedback on the language itself, which would require advanced linguistic insight.

Several issues related to the criteria surfaced during this phase of the project, as the criteria confused some of the pupils. They expressed that some of the criteria was ambiguous and irrelevant to the project. Also, it became eminently clear that the criteria was somewhat designed as a checklist of information, rather than an instrument aiming to give constructive feedback. As a result of this, it became necessary to revise and simplify the criteria in addition to provide more recordings which the participants could assess. The new criteria of assessment (appendix 1) was made using some of the earlier suggestions the students proposed. The design of the criteria sheet was decided after a brief discussion with the test group. The learners received coaching on how to give feedback, as well as it was emphasized that they needed to point out the most positive aspects of the performance. This was done in order to implement some of Hattie and Yate's claim that positive as well as constructive feedback is needed (2014:64). In addition, it was emphasized that they needed to identify the most crucial aspect that needed attention from the assesse in order to further improve their performance. After the second round of practice in assessing recordings, the assessments were more concrete and it was decided that the participants could start to assess each other's performances. This decision was based on the researcher's assessment of the feedback. It was required that each participant was able to provide concrete feedback to all three of the recorded videos before moving on to the next phase of the project. This was done to make sure that the participants could identify different errors and mistakes their peers could do.

Before they were going to assess one another, the participants were provided with written scenarios. None of the members within a group received identical scenarios. There were a total of four groups. Group 1 and 3 consisted of four members while group 2 and 4 consisted of three. Scenario 1, 2, and 3 were given to four participants, while scenario 4 was given to the fourth members of group 1 and 3. They were then asked to evaluate and make their own interpretation to what the best answer to their specific scenario was. This was deemed necessary as they were going to assess the content their peers conveyed in their message, and would therefore need insight into what format and what information that were relevant to the situation. To atone for possible misinterpretations, the participants were asked to discuss the scenario with the participants who received the same scenario. The teacher monitored these discussion groups to see that they came to an agreement. The teacher spent more time discussing with the participants who received scenario 4, as this group consisted of only two participants. They were also asked to practice their own pronunciation of the scenario, as mispronunciations in their presentation of the scenario could cause problems for the assesse.

Before the first rounds of testing, pupil 1 (assessor 1) was asked to discuss the scenario with pupil 3 (assessor 2), as they were ones acting as assessors for pupil 2 (assesse). They had to have a common idea of what parts of the scenario that would be relevant for pupil 2 to implement into his or hers conveyed message, in order to assess pupil 2's performance. The researcher observed this process, and the testing did not take place before a common understanding was achieved between the assessors. Pupil 2 was allowed to take notes during the testing, in addition to ask the assessor to repeat parts of the scenario. This was implemented because of the availability of asking questions is prominent in oral interaction. In addition, in an actual distress situation at sea, the crewmembers always have the information of the ships' position and identification at their disposal. This is additionally information that needs to be precise, yet can be hard to perceive. Next, the message was conveyed by the assesse to the recipient. This was to investigate if there were any important aspects of the message that were not accounted for in the criteria. The roles rotated so that each of the participants acted as assessor 1, assessor 2, assesse and recipient. The two groups who consisted of only three members had a fourth member who took the role of the recipient in the task, who was one of the highest achievers from the other groups. Immediately after each test round, the assessors filled out a response sheet, in which they were required to at least pinpoint something that was good about the performance, and something that could have been improved. After all the participants of a group had gone through with the first test round, they received the written response from their assessors and spent five-ten minutes reading it. They were then asked to hand in the written response for it to be copied, before the feedback was returned to them. This was done to gain access to the written feedback, which was crucial data to the project.

At the start of the next session the participants yet again were asked to read their received response. The students did get 45 minutes to read though the response, and prepare for the last tests rounds. During this stage they were additionally asked to rehearse the scenario they were going to read to the assesse. Only two of the groups were able to go through with the recordings this session, whereas the other groups were scheduled to do the testing the following day. However, the two remaining groups were not able to do the second tests until three weeks after the initial testing. Because of the lack of momentum in the project, and the extension of the schedule, the data from these two groups were omitted from the study.

3.7 Data gathering.

The data, which creates the basis for the discussion chapter, is primarily based on the recorded performances of the test subjects, in addition to the written feedback they received from their peers. It is important however, to highlight that the peers could have given response to each other that was conveyed orally, and not in the presence of the researcher. To optimize the insight into the students' response and feedback to each other, observation of second order was used to remedy for possible spontaneous feedback and discussions that could occur after the tests. With observation of the discourse that could occur after a test round, as well as a copies of the written feedback, it is likely to enable sufficient insight into the constructive feedback each test subject would get. Bjørndal (2011) defines observation of second order as a situation where the observer is an active agent in the process that is to be observed. Bjørndal further explains that first order observations are preferable as it enables the observer to solely focus on the observation (Bjørndal 2011:33), but this was not possible due to the teacher's responsibility of guiding the participants through the project. In addition, observation is a method that is highly influenced by the observer's motivation for during the observation, as well as the condition of the observer at the time of the observation. This however, is not crucial to the project as the main focus will be on the written response that the students used while preparing for the second test round. However, it is of note that the teacher took notes related to the discourse of the different groups, which could give some insight into the social structures of the groups. In addition, by merging the role of researcher and teacher, earlier experiences with the test group could be used to explain certain aspects of the project, although being based on a subjective opinion.

There are aspects of the data-gathering that could have enabled more insight into the social structures of the test group, which is particularly relevant when discussing the group compositions. By using a survey which enquired about preferred peers could have complemented the teachers' assessment of the social structures of the test group. However, by discussing with other teachers, who had experience with the test group for the last six months, it was deemed unnecessary to devote more time to the creation of the groups.

3.8 Group Compositions

The group compositions themselves required careful planning. As Topping (2009) suggests, the participants should be matched to peers that are of the same level of competence. However, as mentioned in the theory section, there are other arguments for how to properly organised group compositions. Cheng (2008) focused on the quality of groups processes when

investigating group compositions. In essence, Cheng's study concluded that interdependence, individual accountability, equal participation and social skills were most important factors in successful group compositions. Peer assessment, by its very essence, accounts for interdependence and equal participation. As assessors, the participants were mutually dependent on each other in terms of receiving feedback. In addition, they were all required to act as both assessors and assessee's, resulting in equal participation. Social skills is also of importance, and crucial when establishing trust between the participants (Cheng 2008:208), which is considered crucial by Johnson (2006). In an attempt to account for both Toppings' and Johnson's view on matching participants, the students were placed in a group with at least one person they were regarded as comfortable with. The other two participants were regarded as have the approximately same level of competence in of English. As a result of these group compositions, several issues related to group work could occur. Lower achievers could feel anxious when being assessed by a higher achievers, and higher achievers could disregard response provided by lower achievers. However, the students were generally placed in a group where they had at least one person they were comfortable with, and very often they matched in terms of language competence. However, as the test group consisted of 14 participants it was seemingly impossible to perfectly match the participants, while taking into account both the level of competence, and social patterns. This could be the cause of several issues related to providing feedback. It is also of note that the attempt to match up same-ability peers were only based on the teacher's assessment of their earlier work in written and spoken English.

3.7 Limitations and Shortcomings.

The following section aims to highlight some of the flaws and shortcomings of the project. These shortcomings renders any conclusions made by this thesis invalid. This however, does not indicate that the project does not give some insight into the significance of the several important aspects related to peer assessment.

3.7.1 Lack of Control Group.

It is important to mention that a comparative analysis of two groups where one received feedback and the other did not would be beneficial to identifying the significance of the feedback provided. This was not implemented as the test groups were already limited. Additionally, it would require both the control group and the test group to be adequately similar in level of competence, which would be further difficult to account for with the limited number of participants. As a result, the cases will be discussed in relation to their

possible improvements in their second test, in addition to the performances of their peers and the received feedback. Other factors as social power structures will additionally be mentioned, and will be based on the subjective interpretations of the researcher. Signs of social anxiety could also emerge from the performances of the participants when combined with the provided response. It is important to be aware of the lack of a control group as any tendencies shown in the results will have no benchmark to be compared to in terms of scrutinising the significance of the results. Despite this, these tendencies can give some indications to important aspects of implementing peer assessment in oral learning activities in EFL.

3.7.2 Possible Exposure to the Scenario Prior to Testing.

During the phase where the assessor evaluated and examined the scenario they had been given, there was little to no control on how these scenarios were treated in terms of exposure to the assesses. It could have been the case that some of the participants would have been exposed to the scenario before their test, which would limit the effect of evaluating the scenario during the recording. In an effort to remedy for this possibility, the participants were informed that the transmitted message was not to be graded.

3.7.3 Possibility of Feedback Received Orally

The basis for evaluating the significance of the feedback was done solely on the written scenario in combination with the improvements of the participants' second tests. However, it is important to note that it occurred that the participants discussed the performance of an assesse in-between test rounds. As these interaction between the participants were neither recorded nor actively observed by the researcher it is possible that some of the participants received feedback that is not accounted for in this study. Feedback not accounted for could render observations and analytics concerning peer modelling unfounded, as some of these assumptions are based on improvements that occurred where constructive feedback were scarce.

3.7.4 Different Evaluation Setting Compared to Training.

One of the aspects that could improve the quality of the feedback would be to train learners in providing feedback in the same setting as the testing. In this project, the participants evaluated recorded videos which they could reinvestigate and watch as many times as they needed in order to assess them. This situation is quite different to the actual testing. With the recordings they had two sessions with a total of three hours at their disposal to write constructive feedback. In the test rounds they spent just a couple of minutes in providing feedback, to a performance they could not listed to. In further projects, the performances of the assesses

should be recorded and made available to the assessors. This would still require the participants' perceptual skills to be implemented, while simultaneously give the assessors better conditions to provide formative feedback. However, this could even further increase the level of anxiety that the participant could be experiencing in a situation of assessment.

3.7.5 Limited Number of Test Subjects.

It is important to note that any results or conclusions based on a study with limited number of participants would be highly debateable. This certainly is the case for this study, as the original group of 14 participants, which is a low number of participants, was later reduced to 7 test subjects.

3.7.6 Insufficient Prepared Assessors.

Although being explicitly tasked with analysing the scenario before reading it to the assessee, some participants expressed that they had not identified to most relevant information in their scenario before assessment. This would be a problem as they would not have any basis for assessing their peers, and could severely limit the participants' ability to provide constructive feedback to their peers. In these cases, the researcher intervened and discussed the scenario with both assessor 1 and 2, with the assessee not being present. A common understanding was achieved before the assessee was allowed into the room. This could have limited the assessors ability to give concise feedback, as they had not spent time to evaluate the scenario that they were going to read.

4. The Study.

In the following section a selection of the test rounds will be presented. Group 1 and 4 were selected for further examination. At first glance, there were differences in relation to the improvements in group 1 and 4. This section will investigate each of performances of the selected participants, in addition to the provided feedback as a result of the first test round. There will be commented on the performances of each individual participant, in addition to the feedback they received after the first recordings. The feedback the participants gave to their peers will also be commented on, as there could possibly be correlations between the assessments they make and their own performance. All these comments are included in order to emphasise aspects of the recordings that is of importance, and to express these aspects to the reader of this thesis. The transcribed recordings have been scrutinized to look for errors both linguistic and contextual. Underlined sentences show information that the context did not demand to be conveyed further. These sentences indicate a lack in the participants' ability to extract the most significant information of the scenarios. Sentences that are highlighted in red signifies misplacement in terms of the formal criteria related to distress calls.

Ungrammatical features and words are highlighted in blue. Grammatical errors that occur because of pauses in the message is not regarded as linguistic errors in this study, although the number of pauses will be commented on. It is also of note that the metalanguage of the study was Norwegian. Most of the provided feedback was written in Norwegian, with some exceptions. Feedback translated by the teacher is marked with a “*”. It is also of note that the assessors' reading of the provided scenario was recorded. This was done to ensure that formal deviations in terms of position, call sign etc. could be traced back to the assessor's reading of the scenario. If the assessor would provide the assessee with information that differed to that of the written scenario, it would not be regarded as an error.

The structure of the presentation will be the following; transcribed performance in the first test round, received response, conveyed response, and the performance in the second test round. Each of these segments will also be commented on by the researcher. This is also an assessment on the quality of the response each participant gives and receives.

4.1 Results.

Many participants failed to take part in the second test rounds in accordance to the schedule. As a result, two of the groups were not able to do the second round of testing until three weeks after the initial schedule. The length of this period is substantial, and could severely influence the results received in the second test round. As such, the data from these two

groups is omitted from the results section, limiting the numbers of participants to seven individuals. The remaining two groups were able to go through with the project in accordance to the schedule, and concluded the project one week after the initial test round.

4.2 Conveyed Messages and Response of Group 1.

4.2.1 Test subject 1.

Test 1.

There is fire in the engine room

There is also a oil leakage

A helicopter is on the way

And the.

Comments from the researcher:

The participant did specify the format after the recording, and did select a mayday call as the proper format. In the transmitted message the test subject did not include crucial information such as position of the vessel, call sign, or the number of people aboard. The message contained the most relevant information in terms of the nature of the emergency, but added two pieces of information that was not deemed relevant to the situation.

Response:

Subject 4:

Positive feedback: **“Everything was good”*

Constructive feedback: **Be more certain on what you are going to say*

Subject 2:

P: **He talked loud and clear”*

C: **Practise and reformulate who the message should be conveyed to.*

Comments from the researcher:

The feedback provided by subject 4 was slightly ambiguous as subject 4 regarded the transmitted message to be good, but added that the assesse needed to be more certain in terms

of conveying the information. Subject number 2 was able to point to certain aspects of the message that were good, namely the voice usage during the message. The constructive feedback provided by test subject 2 aimed to increase the assesse's awareness to the relation between the message and the recipient.

Response given as assessor:

To Subject 3:

P: **Include the position*

C: **Use the correct format*

To subject 4:

P: **He understood that it was a mayday- call. He notified the correct recipient.*

C: **Talk a bit louder.*

Comments from the researcher:

The participant seems to be aware of the criteria to what format the situations would require, as he highlighted that his peer remembered to include the position in his conveyed message. It is interesting that the participant notified his assesse included the position of the vessel, which is something subject 1 failed to include in his own conveyed message.

Test 2.

Pan-pan, pan-pan, pan-pan

This is striker

Call sign EE 67

Our position is 35 degrees north 14 minutes and 12 seconds, 1 degree east 55 minutes and 36 seconds

We have 15 crewmembers on board

And **the** we have spotted sparks of fire in the engine room

The passengers are (pause) anxious

And we need backup right away, because we don't have firefighters on board.

Comments from the researcher:

In the second test round subject 1 made one mistake in terms of the structure of the message. The participant was able to include all the necessary information in terms of formal requirements. It is of note that subject 1 made one mistake related to the position of the vessel, as the position read by the assessor was 34 degrees north.

4.2.2 Test subject 2.

Test 1.

Passengers on board Alara, this is the captain.

We have a problem in the engine room

And we are trying to sort it out

Just keep calm and await **for** further instructions.

Comments from the researcher:

The key information of this scenario was the sighting of fire in the engine room. The participant decided to address the passengers rather than transmitting a mayday or a pan-pan call.

Response:

Subject 4:

P: **Everything.*

C: **Nothing.*

Subject 3:

P: **Talked loud and clear*

C: **Find out who the message should be conveyed to. Could have been two additional recipients.*

Comments from the researcher:

Subject 4 wrote that “*nothing*” needed to be improved. Despite this, the participant regarded the assesse to have used the wrong format based on the feedback in the assessment sheet.

Subject 3 did provide more substantial feedback, despite that this feedback was based on a subjective opinion. As every scenario included a key sentence that would strongly indicate the proper format, it appears that subject 3 did not identify this, as subject 3 thought there could be two additional recipients. Both assessors regarded the format to be wrong in the provided sheet which addressed level of performance. This assessment was supported by the researcher.

Response given to peers:

To subject 3:

P: **Talked loud and clear*

C: **Did send the message to wrong recipient.*

To subject 1;

P: **He talked loud and clear”*

C: **Practise and reformulate who the message should be conveyed to.*

Comments from the researcher:

The response subject 2 gave to his peers focused largely on using the correct format. This is interesting to behold as subject 2 did in fact not use the correct format himself.

Test 2.

Pan-pan, pan-pan, pan-pan

This is Narvik

Call sign ME 45

We have smoke in the engine room

And all our firefighters are sick

So we need back-up firefighters to assist us in the engine room

Our position is 13 degrees North, and (pause) 12 minutes and 45 seconds

And east 45 degrees, 53 minutes and 35 seconds.

Comments from the researcher:

It is to be noted that the number of crewmembers was not mentioned by the assessor. The omission of this information is not regarded as a mistake made by the assesse, as it was not included in the scenario. With this in mind, test subject 2 made one mistake in relation to the required formalities to the conveyed message, and one mistake in relation to the content of the message. It is additionally important to note that the segment that should have been omitted indicates that assistance was needed, signifying that a mayday-call should have been transmitted instead of a pan-pan call. Subject 2 was able to perceive the most relevant information, and replied using the correct format.

4.2.3 Test Subject 3.

Test 1.

Mayday, mayday, mayday

This is NS creator

We are taking in water

Our position is 10 degree north, 4 minutes and 001 degrees west

We are 15 crewmembers on board

And the water pumps are working

And the water level is not rising anymore

And the call sign is NS 556

Comments from the researcher:

Subject 3 made several errors in this conveyed message. Firstly, a number of required formalities was misplaced. Secondly, the situation did not indicate that a mayday call was necessary, as the water level had stabilized. It is of note that the most essential parts of the transmitted message were included, in terms of the required formalities.

Response:

Subject 1:

P: **Include the position*

C: **Use the correct format*

Subject 2:

P: **Talked loud and clear*

C: **Did send the message to wrong recipient.*

Comments from the researcher:

The response subject 3 received was not formative, as solely stating that subject 3 used the wrong format would not help the assessee in the second testing. Neither subject 1 or 2 gave any feedback that could have enabled participant 3 to improve his performance. Both assessors regarded that subject 3 used the wrong format, a conclusion which is supported by the researcher.

Response given to peers:

To subject 2:

P: **Talked loud and clear*

C: **Find out who the message should be conveyed to. Could have been two additional recipients.*

To subject 4:

P: **Does include most of the information*

C: **Should practise pronunciation*

Comments from the researcher:

Subject 3 included every required segment the distress call in relation to the required criteria. As such, he was able to express that subject 4 included mostly all of the required information. Subject 3 also highlighted pronunciation, as one of the only ones during the entire study.

Test 2.

Pan-pan, pan-pan, pan-pan

This is Bergen

BN 35

The Engine has stopped working

And we have 16 crewmembers on board

And one have food poison

My position is 65 degrees north, 47 minutes 11 degreed west 26 minutes and 22 second

Comments from the researcher:

In the second test the participant misplaced one segment, and included one that was irrelevant to the situation. However, subject 3 where able to use the correct format.

4.2.4 Test Subject 4.

Test 1.

Mayday, mayday, mayday

(pause) This is Roald Amundsen

We need some assistance

We have one crewmember with a heart attack

The position is 13,5 North, and 1 west.

Call sign 45 EE

Comments from the researcher:

Participant 4 did not include irrelevant information. There were two segments that should have been placed differently in the message. The only information omitted was the number of crewmembers on board. Subject 4 were able to perceive the most relevant information, and did use the correct format.

Response:

Subject 3:

P: **Does include most of the information*

C: **Should practise pronunciation*

Subject 1.

P: **He understood that it was a mayday- call. He notified the correct recipient.*

C: **Talk a bit louder.*

Comments from the researcher:

The positive feedback of both assessors pointed to aspects of the message that were good. However, subject 3 did not supplement the response by specifying what information that was missing from the message. The constructive feedback from subject 1 could have been applicable, as increased awareness to speaking clearly could better the communicative aspect of transmitting a distress- or urgency call.

Response given to peers:**To subject 1:**

P: **“Everything was good”*

C: **Be more certain on what you are going to say*

To subject 2:

P: **Everything.*

C: **Nothing.*

Comments from the researcher:

The feedback subject 4 gave to his peers is quite interesting to behold. Not in a sense that the feedback were particularly constructive, but rather the opposite. Subject 4 was the only participant of group 1 to use the proper format of the distress call. As such, it would be likely that subject 4 would specify the cases in which his peers used wrong format. This was however not the case, and it will be discussed later in this thesis.

Test 2.

This is Eidsvoll

Call sign (pause) EIS 45

Position 65 degrees north 13 minutes and 5 seconds, 13 degrees east 45 minutes and 55 seconds.

It's (pause) one of the crewmembers [overboard](#)

And the (pause) it's [did](#) some oil leakage

And it's 54 passengers and 15 crewmembers.

Oh, I think I forgot the pan-pan

Comments from the researcher:

The number of pauses in the second test were substantial when compared to the first test. The first test of participant 3 lasted 34 seconds, while the second one ended at 54 seconds. In terms of the structure of the conveyed message only one segment was misplaced. However, the participant added some irrelevant information.

4.3 Examination of group 1.

In order to investigate the results in relation to the research question of the project it is important to investigate possible improvements in the second test round as compared to the first test round. Additionally, it is relevant to see if the improvements coincide with the provided feedback. A correlation between the feedback provided and the improvements made

could indicate that the students were able to give constructive feedback that is approachable to the peers.

When examining subject 1's first test round it became obvious that the task at hand was unclear to the student. Subject 1 was able to specify the format of the scenario, but did not include any of the formal requirements of a distress call, except for mentioning the nature of the emergency. Subject 1 was regarded as one of the highest achievers in terms of using and implementing spoken English, in addition to having social capital in test the group. This could have influenced subject 4's response to the performance. Subject 4 responded that "everything was good" despite subject 1 omitted several important features of the mayday call. At first glance one could argue that subject 4 did not know what information needed to be included in the conveyed message for it be complete, but subject 4 was able to include the most essential information of a distress call in his own performance. However, one explanation of this could be that subject 4 observed the performance of subject 3, who included every segment of the distress call. Another explanation, which coincides with reported pitfalls related to peer assessment (Topping 2009:24), is that subject 4 experienced anxiety whilst assessing subject 1. Either way, participant 4 were unable to concretise the feedback to subject 1. This would coincide with some of the reported pitfalls of peer assessment. Despite this, subject 1 second performance was far superior compared to the first one. One explanation to this this increase in performance could be the effect of peer modelling. Topping and Ehly stated that "Modeling on and by peers might thus lead to greater metacognitive awareness, and thereby more self-regulation"(Topping and Ehly 2001:117). Subject 1 could have been unsure of the requirements of the task, but that this was clarified by watching the performance of his peers. It is possible to assume that this improvement would not have been that significant if subject 1 was unable to observe and assess the performances of his peers. As such, it seems reasonable to assume that peer modelling, rather than the received feedback, impacted subject 1's performance in the second test round. However, it is possible that subject 2's response; "*Practise and reformulate who the message should be conveyed to*" could have been applicable to subject 1. As mentioned in the theory section, intermediate learners or advanced learners need feedback which addresses professional strategies for improving their performance (Hattie and Yates 2014:66). It could be that the simple word "practise" could have caused subject 1 to spend more time preparing before the second test round. At this point it also important to highlight the social processes between subject 1 and 2. Both were equally proficient in both spoken and written English. In addition, when asked who they would prefer

to be matched with in group projects, they both named each other. As a result of this, it could be reasonable to highlight Cheng's focus on the importance of social skills and trust in group dynamics (Cheng 2008:208), and that trust-building is an important factor of peer assessment.

4.3.1 Perceptual Mistakes

One of the mistakes subject 1 made in the second test round was related to the position of the vessel, as the coordinates conveyed by the participant were incorrect. Initially, this would not be relevant in an actual emergency situation as this information is provided by the vessels instruments, but for this project it gives some insight into the role of reception. Although being aware of the necessity of providing the position of the vessel, subject 1 was unable to correctly take note of the coordinates. All of the assesses were explicitly asked to write down the coordinates. Misperceptions of the provided information indicates that awareness of relevant information does not warrant a successful piece of communication.

4.3.2 Inability to Concretise Feedback.

The case of subject 4 can be used to showcase that there are pitfalls related to peer assessment. As mentioned, subject 4 was unable to give constructive feedback to subject 1. When assessing subject 2, who used the wrong format, subject 4 replied that "Everything" about the message was relevant to the provided scenario. If this was the only provided feedback it could have had damaging effects to participant 2's performance in the second test. Earlier, in the theory section, the idea of judgments of learning was broached. If a learner experience high JOL it renders him or her unmotivated to further improvements (Metcalfe and Kornell 2004:464-465). As such, a feedback stating that everything about the performance is adequate could make the learner unlikely to put effort into enhancing his or hers performance.

4.3.3 Summary of Group 1.

In the first test round, only one of the four participants responded in accordance to the situation and the recipient, whereas all four used to correct format in the second tests. As such, it is not controversial to claim that group 1 benefitted from peer assessment. It is difficult to trace the improvements to the feedback itself, but through the result of subject 1's improvements it is plausible to claim that peer modelling had a significant impact on the English oral skills of the participants. This with the basis or oral skills defined as "being able to listen, speak and interact using the English language. It means evaluating and adapting ways of expression to the purpose of the conversation, the recipient and the situation." (Kunnskapsdepartementet. 2013:2).

4.4 Conveyed Messages and Response of Group 4

4.4.1 Test Subject 12.

Test 1.

Pan-pan, pan-pan, pan-pan

This is NS creator.

(Lack of position)

We are currently taking in water

But the (pause) all the water pumps are (pause) are in order.

And the water level doesn't seem drastic.

There are fourteen crewmembers on board

And some of the crewmembers have not experienced this before

The call sign is NS 556

Comments from the researcher:

The only information missing was the position of the vessel, which is considered crucial. Otherwise, some pieces of irrelevant information were included, and one information segment was misplaced.

Response:

Subject 13:

Positive feedback: **Talked loud and clear. Included every bit of information. Correct format.*

Constructive feedback: **some mistakes in relation to the structure of the information*

Subject 14:

P:**Said what happened, and that pumps are functioning and that it stopped after 30 minutes.*

C:**Concrete information. Order of the information*

Comments from the researcher:

Both assessors evaluated the assessee as intermediate in terms of pronunciation and the clearness of the message. It is of note that both assessors gave feedback to participant 12 that the structure of the conveyed message needed to be improved. Both assessors focused on the structure of the urgency call. This feedback is considered constructive, as the structure of the urgency call was formalized through the criteria of assessment.

Response given to peers:

To subject 14:

P: **He included most of the relevant information*

C: **Talk a bit more clear.*

To subject 13:

P: ** Good that he addresses the passengers*

C: ** Say that a mayday message has been transmitted.*

Comments from the researcher:

Through the response subject 12 gave to his peers, one could assume that he had a clear focus on what information that is relevant in certain situations. This can coincide with the fact that he were able to include all the different components of an urgency call except for the location of the vessel.

Test 2.

Striker Striker Striker (format of the distress call is absent, but this fragment of the message is regarded as the name of the vessel)

Call sign EE67

Position 34 Degrees North, 15 minutes and 12 seconds, 1 Degree east.

There **have** been a fire in the engine room.

There are 15 crewmembers

And the passengers are starting to catching on the (pause) what's going on.

Comments from the researcher:

The participant made fewer mistakes in terms of the structure of the distress call when compared to the first test. The only missing information was the specification of the format. After the recording the participant immediately noted that the proper format would be a mayday call, which would be the correct format to use for this scenario. The number of sentences containing irrelevant information were lessened from three in the first test round, to merely one in the second test.

4.4.2 Test Subject 14.

Test 1.

Mayday, mayday, mayday

This is Roald Amundsen, Roald Amundsen, Roald Amundsen.

Call sign 45 Echo Echo

We have thirteen people on this vessel

And our position is 13 degrees north and 1 degree west

One of our fellow crewmember **has** a heart attack

And we need immediate assistance

And we have fuel enough for staying at sea for three weeks

And the weather is nice

But the Forecast says that **is** gonna be bad weather in the next few days.

Comments from the researcher:

Subject 14 included all the required information in the distress call. Additionally, he also used the correct format. Despite this, subject 14 also included information that was not relevant to distress call.

Response recieved:

Subject 12:

P: **He included most of the relevant information*

C: **Talk a bit more clear.*

Subject 13:

P: **Loud/clear. Included everything*

C: **Did not need to include that the weather was going to be worse the next few days.*

Comments from the researcher:

Both test participants number 12 and 13 regarded participant 14's pronunciation to be quite good. They also assessed 14 to perform from intermediate to high achiever in terms of speaking clearly. In addition, the assessors both assessed that the content of subject 14 could need some attention. Student 13 emphasized that there were parts of the information in the message that were not needed in the message.

Response given to peers:

To subject 12:

P: **Said what happened, and that pumps are functioning and that it stopped after 30 minutes.*

C: **Concrete information. Order of the information*

To subject 13:

P: *He chose the right scenario. Told what was going on to happen and what is happening.*

C: *Concrete information. Not being shaky in the voice.*

Comments from the researcher:

Subject 14 highlighted subject 13's need of structuralizing his conveyed message. This feedback is considered formative, as it addressed an aspect of subject 13's performance that needed improvement. In addition, this feedback is applicable to subject 13. However, subject 14 also noted that both his peers needed more "concrete information", while not addressing what specific parts of the message that needed to be concretized. As such, some of the feedback was not considered applicable and formative.

Test 2.

Pan-pan, pan-pan, pan-pan

This is Eidsvoll, Eidsvoll, Eidsvoll

Call sign is Echo Echo 45

There is one man over board

At this vessel

And the (pause) and they are staying close to the ship

And the passengers are monitoring the situation

We have an oil leakage on the vessel

Our position is 65 degrees south 13 minutes and 15 seconds

13 degrees east 45 minutes and 57 seconds

We're 15 total crewmembers and 54 passengers on board.

Comments from the researcher:

Subject 14 included all the required information in a pan-pan call, but made one mistake in relation to the structure of the message. He additionally included information that is needed in the conveyed message. There were additionally some grammatical mistakes.

4.4.3 Test Subject 13.

Test 1.

Dear Passengers

There has been sighted fire in the engine room

And the (pause) and a oil leakage

And a helicopter is coming.

Comment from the researcher:

This performance displayed a flaw in the design of the project. The format of addressing the passengers was not ideal to the project as it was not based on the strict formal criteria of assessment. As a result, this format was omitted from the other tests. However, this is an important finding as it gives some insight into some of the benefits that occur as a result of being an assessor (Topping 2009: 22), which will be further discussed.

Response:

Subject 14:

P: He chose the right scenario. Told what was going on to happen and what is happening.

C: Concrete information. Not being shaky in the voice.

Subject 12:

P: Good that he addresses the passengers*

*C: * Say that a mayday message has been transmitted.*

Comments from the researcher:

Both of the assessors evaluated 13's performance to be intermediate. In the response sheet they both stated that the assessee made minor mistakes in terms of language, clarity of voice and content. The reasoning behind this was that this specific format would be based on subjectivity on behalf of the assessors, as the students were not thoroughly taught how to address passengers. In addition, this way of communication does not follow the same strict formal requirements that distress calls do, which would make it harder for the students to give attainable feedback as well as improving their performance based on that feedback.

Response given to peers:

To subject 12:

*Positive feedback: *Talked loud and clear. Included every bit of information. Correct format.*

*Constructive feedback: *some mistakes in relation to the structure of the information*

To subject 14:

P: **Loud/clear. Included everything*

C: **Did not need to include that the weather was going to be worse the next few days.*

Comments from the researcher:

Subject 13 was able to identify that both peers included every bit of relevant information, and that they used the proper format. The participant also highlighted the need of subject 12 to focus on the structure of the urgency call, a response which is considered formative. The constructive feedback provided to subject 14 were less constructive, as it would not be relevant in any different scenario that then one subject 13 were tasked with.

Test 2.

Mayday, Mayday, Mayday

This is Bergen

Call sign BN 35

Our position is 65 degrees north, 47 minutes and 18 seconds

11 degrees west, 26 minutes and 22 seconds.

The engine has stopped

and one of our fellow crewmember [have had been foodpoison](#) and is ill.

We have 0 passengers on board.

Comments from the researcher:

Subject 13 included all the required information of a distress call, and used the correct format. One sentences of irrelevant information was included as well.

4.5 Examination of Group 4.

The results from group 4 give some interesting insight into possible gains and problems related to peer assessment. Most notably is the improvement of participant number 12's performance in relation to the desired formal requirements linked to a pan-pan or mayday call.

However, it is additionally important to analyse student 14 performances, as there was not any significant improvement of the second test when compared to the first test round. In addition, subject 13's performance in the second test was superior to that of his peers, despite not having used a distress call in the first test round.

Among all the participants in this project, none received more concrete feedback than subject 12. In the first test he included information that was not relevant to the urgency call. By including a lot of contextual information in the message it indicated that the participant was not fully aware of the criteria of a successful distress call. In addition, the participant included almost all the provided information of the scenario, which suggests lack of ability to extract the necessary information that the situation demanded. Participant 12 received feedback on the structure of the distress call, yet was unable to structure the distress call according to the formal requirements. However, in the first test the participant misplaced 1 segment, and omitted one. In the second test all the required segments were present, although one was misplaced. Add this to the fact that the participant did not include any irrelevant information and it becomes eminently clear that the test subject performed better in the second test round. By omitting irrelevant information, the participant showcased an increase in cognitively assessment of the scenario, as he was able to identify to most prominent aspect of the provided scenario. Another possible reason for subject 12's increase in performance could be the value of peer modelling. Both subject 13 and subject 14 were generally more proficient in both spoken and written English. It is also of note that participant number 13 performed a near perfect distress call before subject 12 was going to go through with his second test. As Topping and Ehly (2001:117) peer modelling could be effective as it causes the observer to imitate the performance or product of a peer that is regarded as competent by the observer. This could have been the case in this instance, as subject 12 knew that subject 13 was one of the highest achievers in the test group.

Subject 12 did perform better in the second test round in relation to the formalities of the conveyed message, yet there was little to no improvement in terms of the language spoken. As highlighted in blue writing in the results section, the number of linguistic mistakes were not significantly less in the second taking. However, the number of pauses in the message were lessened. In essence, the performance of subject 12 indicates that the used method can be implemented in EFL when the aim is to improve the structure and situational context of the spoken language. This does however demand a basis which highlights that oral skills in EFL is when the speaker communicates in accordance to the situation.

Subject 13 was the first participant to go through with the test. This was not a coincidence as subject 13 was regarded as one of the highest achievers in terms of spoken English in the entire test group. The reason for this decision was based on the assumption that the peers would feel less anxiety after a confident individual made a few errors, or that they would learn through peer modelling. However, the first test round gave answers to the usage of addressing the passengers as a way of communication. Subject 13 received feedback that he needed to *“Say that a mayday message has been transmitted.”* and that he needed to provide *“Concrete information. Not being shaky in the voice.”* The first response is contextual, and not necessarily applicable in other conveyed messages. The second response was more relevant, but not concise. To further investigate subject 13 performances becomes important as he hardly had any mistakes in his second test round. In terms of providing his peers with constructive feedback, participant number 13 was able to identify mistakes participant 12 made in his performance in relation to the formal criteria. This indicates that subject 13 had insight into the structure of a conveyed distress call at the time of the first test round, which could additionally indicate that the participant had the insight into criteria of assessment and the formal requirements prior to observing the performances of the peers. However, subject 13 stated that participant 12 included all the relevant information. This was not the case, as participant 12 omitted the position of the vessel, which is regarded crucial in an urgency situation.

The omission of a format, the one that addressed passengers, creates an interesting standpoint from which certain aspects of project can be discussed. There were no traceable improvements in subject 13 second performance as there were no benchmark to compare it to. However, the successful performance of subject 13 becomes of importance considering that no received response could have bettered his performance. One explanation to why subject 13 second performance was one of the superior ones can be the role of peer modelling. Ehly and Topping (2001:117) claimed that peer modelling could lead to greater metacognitive awareness and self-regulation. They also claim that peer models are competent, while not flawless. There are indications that the performances of subject 12 and 14 influenced the learning of subject 13. As their performances were not perfect, and subject 13 were tasked with assessing them, it could have lead subject 13 to greater awareness to the criteria of assessment. This assumption is strongly indicated in the case of subject 13, as the participant did not receive any response regarding the formalities of urgency- and distress calls. However, it is additionally plausible that subject 13, who was regarded as one of the best

participants in both spoken and written English, had learnt how to properly convey a successful distress call before the first test rounds. There are no evidence to either confirm or deny this claim. However, none of the other participants, either high or low achievers, performed near the level of subject 13's second performance in their first performance. As such, it is reasonable to claim that there are aspects of participant number 13's performance that indicate that participants gain metacognitive advantages when assessing peers. This coincides with the claims made by Topping (2009:23).

The performances of subject 14 is additionally important to investigate. The number of grammatical errors increased in the second test, and subject 14 was additionally not able to limit the number of irrelevant bits of information. The received response was situational, and not necessarily applicable in other scenarios. It is of importance to note that subject 14 addressed the structure of the message in the assessment of subject 12, but made structural errors himself. As a result, this indicates that there is not necessarily a correlation between the assessor's feedback and the assessor's self-regulative abilities. The performances of subject 13 and 14 are contradicting when addressing the benefits of being the assessor. Both participants assessed the structure of subject 12, but only subject 13 was able to implement the same feedback in his own performance. Subject 14 noted that the order of the information provided in subject 12 performance needed to be improved, although there were misplaced segments and irrelevant information in both of his own performances. It is also of note that participant 13, who performed first in each test round, performed superbly in the second test round in terms of the formal requirements, but that this had little or no effect on the performance of subject 14. In essence, the importance of peer modelling seems vague when scrutinizing subject 14's performance. Furthermore, it is also plausible to claim that subject 14 gained little to no increased metacognitive awareness through assessing his peers.

It thus becomes crucial to further investigate the connection between subject 12 and 13, with special emphasis on the participant who enhanced his performance the most. These two participants generally cooperated in every group project the sample group had conducted during the six months prior to this study. It is reasonable to assume that continuous cooperation would make the participants more comfortable in performing in front of each other. However, it is important that there was a difference in the students' level of competence in spoken and written English. It was common consent among the teachers of the test group that subject 12 benefitted in a range of different subjects from cooperating with participant 13. The general impression was that subject 12 scores tended to drop when he attempted

individual performances. This background information could explain his improvement in the second test round, as subject 12 would possibly acknowledge the response provided by subject 13. The case of subject 12 and 13 could be an example which highlight the importance of trust and peer modelling. The case of subject 12 and 13 can be further explained when applying Cheng's focus on the quality of group processes (2008:216). As mentioned in the theory section, Cheng focused on four elements that are important to group processes: positive interdependence, individual accountability, equal participation and social skill (Cheng 2008:207). As peer assessment is a cooperative learning strategy, social skills becomes eminently important. It is possible that subject 12 and 13 had developed a bond of trust and positive interdependence, as they used to work together. This could have caused them to prepare sufficiently before the first test round, as they knew that their knowledge of the task could be beneficial to their peers. This could further explain the positive results of subject 13 in the second test, despite him not receiving any constructive feedback. Furthermore, as mentioned in theory section, a sense of commitment to the group is important to a participant's willingness to help others (Johnson 2006:48). It is reasonable to claim that the bond between subject 12 and 13 caused them to prepare sufficiently. In addition, the participants would be individually accountable, as they were going to provide constructive feedback to their peers. In general, the case of subject 12 and 13 supports the claims of Cheng: that group processes are more important than matching participants in either homogenous or heterogeneous groups.

4.5.1 Summary of Group 4.

The results of group 4 is slightly ambiguous when addressing the importance of peer assessment. The case of subject 13 stands out as a great example on how a learner can benefit from assessing peers. Additionally, the benefits related to peer modelling though peer assessment also becomes eminent when investigating the results of subject 12 and 13's performances. There were also a correlation between the aspects subject 13 focused in his feedback to his peers and the performance in his second test. Despite all this, subject 14's results did not show a correlation between the feedback he provided and his own performance. In addition, subject 14 were unable to perceive and evaluate what specific information that needed to be included in his performance. In essence, the results of group 4 expressed the importance of group dynamics and processes.

4.6 Summary of the Results.

In essence, the performances of subject 1, 2, 3 and 12 in the second test round were superior when compared to their first recordings. Subject 1, 2 and 3 were the ones who increased their oral skills the most, as all three was unable to use the correct format in relation to what the situation demanded in the first tests, but were able to both use the correct format, and include most of the formal requirements of the transmission. This signifies an increase in the active listening skill, and an increase in the cognitive skills related to this specific way of spoken communication in English.

These cases all highlight different aspects of peer assessment that could be beneficial to the participants. The increase of subject 13's performances were difficult to distinguish as the format of addressing passengers were not properly implemented in the project in terms of the criteria of assessment. It is however possible to assume that the performance of participant 13 was influenced by him taking the role of the assessor in the first test round. On the other hand, subject 14 assessed the structure of his peers' performance, but was unable implement this into his second performance. The general impression was that the improvements of the performances was linked to peer modelling and the benefits of taking the role of the assessor, rather than the response itself. In the second test round all seven participants were able to cognitively evaluate the perceived scenario, and respond with the format the situation would indicate. This differs to the first tests, were four out of seven used to correct format. This suggests that the participants increased their ability to express themselves orally adapted to the situation, which is the established definition of oral skills in this project.

The participants were asked to write a response to their peers, but were additionally tasked with assessing them based on a low-to-high achievement scale (appendix 1). For the selected test group, the general assessments were based around the intermediate level of achievement. This coincides with earlier research (Topping 2009:24). Out of the seven participants further scrutinized, only one was assessed by any assessor as a higher achiever in terms of including the most relevant information and in terms of language. There were cases where the level of achievement-sheet did not match the feedback provided. One participant expressed that his peer did everything perfectly, but noted the performance as average in the assessment sheet. This could indicate that the assessor did not acknowledge the performance as perfect, but experienced anxiety in providing concise feedback, or that he felt unqualified to provide this kind of response. This would coincide with some of the reported pitfalls of peer assessment (Topping 2008:24), as this could additionally been a sign of free loafing.

4.8 Perception.

As mentioned in the theory section, perception is an essential part of successfully communicating in any language. In this project the role of reception was crucial as it determined whether or not the assesse would get insight into every segment of information that was provided in the different scenarios. If a participant would not able to identify the most relevant information it would be hard for him to determine the proper format of the distress call. The scenarios were designed to be complex in terms of the gravity of the situation, with the exception of one or two sentences that would strongly indicate the proper format. If the participants failed to identify these sentences, or misinterpret them, it could lead them to using the wrong format. If the participant chose the correct format it would indicate that the participant's reception skills were adequate to identify the most relevant information in the given scenario. Thus, it was through the production of the participants that their perceptual skills were identified. However, there are issues related to this as there could be instances where the participants did perceive the most relevant information, but are unable to successfully convey a distress call. Despite this, it is likely to assume that students able to identify crucial information in a given scenario, with some insight into how to give a distress call, would be able to at least present the nature of the emergency. There were other pitfalls related to perception. The provided scenario was read by a peer, which could have caused misinterpretations. None of the participants were regarded as a native or near-native speakers of English. Despite this, as all seven participants used to correct format in the second test, it is reasonable to claim that there was an increase in the participants' active listening skills during the course of this project, related to maritime emergencies.

4.9 Lack of Motivation and Miscalculated High Judgment of Learning.

As mentioned in the theory section, a high judgment of learning could severely limit the student's motivation for further working with the project. This would occur if the task at hand was not challenging to the student. Even before the first test rounds, one of the participants expressed that he was competent to successfully convey a distress call. However, this participant made several errors in both tests. The student did additionally express that response was not needed. As the test subject made several errors in both tests, it indicates that a miscalculated high JOL could even further limit the learning gains than simply high JOL. However, this is circumstantial as this participant did the second test three weeks after the first one, but an interesting note nevertheless.

4.10 Inflating Social Power Structures.

Level of competence and social structures were both taken into account in the group composition. By not merely focusing on level of competence in English or social aspects it could have resulted in the groups not being properly constructed according to either. Even though a student considered one of the group members as a friend, he would possibly need to cooperate with two participants who the participant was not comfortable with. In addition, merely considering a peer as a friend does not warrant high quality group processes. The same could have occurred in relation to the idea of same-ability match making, as some might have experienced that the feedback they received from the peers were not valid and reliable, based on their own impression of their assessors' ability to provide constructive feedback. If an assessee would regard himself superior to his assessor, in terms of level of performance, it is unlikely that the feedback and response provided would be deemed necessary by the assessee. These problems could have occurred in the case of subject 4, as he was unable to provide any feedback to his peers, who were both regarded as superior in spoken and written language. This was the most prominent case in terms of social anxiety, as subject 4 outperformed both subject 1 and 2 in the first tests. However, this could additionally indicate that subject 4 were not able to phrase any feedback that would be deemed formative. Either way, this would coincides with some of the common concerns related to implementation of peer assessment (Topping 2009:24)

5. Conclusion.

In general, the study showcased multiple issues and benefits related to peer assessment. Failure to participate, inability to provide constructive feedback, free-rider effects, and social anxiety were all present in the project. By contrast, it could be argued that there are benefits from taking the role of both assessee and assessor. There were also indications of peer modelling having an effect on the performance of learners. These findings coincide with the claims and findings of researchers as Topping and Ehly (2009) and Sadler and Good (2006) and Black (2004).

Sadler and Good state that “Students should be trained to grade accurately and be rewarded for doing so” (2006: 28), which highlight the importance of students’ ability to reliably assess a product or a performance. Although Sadler and Good’s study focused on grading rather than formative assessment this is irrelevant as both would require the assessor to gain insight in the criteria of assessment, which should increase metacognition in relation to a task. Although many of the participants were unable to provide constructive feedback to their peers, most were able to identify aspects that were not properly accounted for in the performance. To identify features of a performance that is not up to standards is certainly the first step of providing constructive feedback, which indicates that implementation of peer assessment over a longer time-span would increase the quality of the formative assessment. There was additionally indications that while the feedback was not always formative, there were benefits related to assessing peers as pinpointing features others missed in their performance often was included in the assessors own performance as assessee. The general idea that peer assessment can promote self-assessment and self-regulation (Topping 2009:23) is supported through some of the performances in this study.

Peer assessment can be successfully implemented in EFL learning, when addressing ways of communication that uses a basis of formal criteria that are applicable to the learners. However, during the project several additional questions emerged. The study was not able to pinpoint what features of peer assessment that were most crucial. In peer assessment, there are multiple complex aspects that needs to be accounted for in order to identify the most important features. Improvements in the second test could be the result of the received feedback, peer modelling, criteria insight as gained by being the assessor, or simply by the fact that the participants did the test a second time. In order to account of the latter, as well as identifying the magnitude of the response, it would be a requirement to make use of a control group. This would also require the number of participants to be quite substantial in order for

the control group to be representative of the test group. Using a control group as a benchmark for the expected improvements of the performance would be essential to exclude to improvements related to repeating the test a second time. It would be optimal to do a comparative analysis of a group who simply repeats the task, a group who only receives response, a group who receives response in addition to watching the performance of peers, and a last group who would receive feedback and assess peers.

5.1 Aspects Worthy of Implementation in Future Studies.

In the future it would be beneficial to conduct a study which investigated the benefits of peer assessment linked to implementation over a longer period of time. This could result in increased social gains, with social patterns being forged over time, and make the students more capable of assessing each other. Many of the problems related to peer assessment would be more accounted for if the participants were comfortable in the role of both assessee and assessor. Future studies should additionally investigate the correlation between the quality of feedback and the improvements of the assessed product, and make use of control groups to specify this correlation.

5.2 Closing Remarks.

The problem with implementing peer assessment in the short term is that there were no apparent changes or improvements in terms of the spoken language. However, as expressing oneself relevant to the situation and the recipient is considered an important part of oral skills, it is reasonable to claim that peer assessment can be successfully implemented to EFL learning in oral learning strategies. This study were not able to pinpoint the feature of peer assessment that were most crucial to noted improvements in the performances, but sheds light on some of the aspects of peer assessment and their relevancy. Further studies should examine the long-term effects of peer assessment in spoken language. It is likely that if the groups used in peer assessment are comfortable in the groups, it could create a group dynamic where the individuals are more comfortable in using English. By doing this, it is possible that spoken language-improvements could be identified. This study indicates that peer assessment can be implemented successfully in the short term for EFL learners, as the majority of the performances in the second test rounds were superior when compared to the first ones.

Bibliography

Bjørke, C. and M. Dypedahl and G. Myklevold. 2014. *Fremmedspråksdidaktikk*. Oslo: Cappelen Damm AS

Bjørndal, C. R.P. 2011. *Det Vurderende Øyet*. Oslo: Gyldendal Akademisk.

Black, P. and M. Wilson and S. Yao. 2011. Road Maps for Learning: A Guide to the Navigation of Learning Progressions. *Measurement: Interdisciplinary Research and Perspectives*, 9:2-3. 71-123. DOI : 10.1080/15366367.2011.591654

Black, P. and C. Harrison and C. Lee and B. Marshall and D. William. 2004. Assessment for learning in the Classroom. *Working Inside the Black Box*. Phi Delta Kappan. 86:1. 8-21.

Cheng, R. W and S. Lam and J. Chan. 2008. When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project-based learning. *British Journal of Educational Psychology*, 78. 205-221. DOI: 10.1348/000709907X218160

Fjørtoft, H. 2014. *Norskdidaktikk*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS

Hattie, J. and G. Yates. 2014. *Visible Learning and the Science of How We Learn*. New York: Routledge.

Hattie, J. and H. Timplerley. 2007. The Power of Feedback. *Review of Educational Research*. 77:1. 81-112. DOI: 10.3102/003465430298487

Imsen, G. 2005. *Elevenes verden*. Oslo: Universitetsforlaget.

Johnson, D. and R.T. Johnson and O.K. Haugaløkken, and A.O Aakervik. 2006. *Samarbeid i skolen*. Namsos : Pedagogisk Psykologisk Forlag AS

Krumsvik, R. J. and R. Säljö. 2013. *Praktisk-Pedagogisk Utdanning*. Fagbokforlaget Vigmostad & Bjørke AS.

Kunnskapsdepartementet. 2013. *English subject curriculum* <http://data.udir.no/k106/ENG1-03.pdf?lang=eng> (18.10.17)

Lees G. D. and W. G. Williamson. 2004. *Handbook For Marine Radio Communication*. UK: Informa Professional

Metcalfe, J and N, Kornell, 2004. A region of proximal learning model of study time allocation. *Journal of Memory and Language* 52:4. 463-477. DOI: 10.1016/j.jml.2004.12.001

Sadler, P and , Good. 2006. The Impact of Self- and Peer-Grading on Student Learning, *Educational Assessment*, 11:1, 1-31

Slemmen, T. 2009. Vurdering for læring. In *La stå*. Oslo: Gyldendal Akademisk.

Thagaard, T. 2004. *Systematikk og innlevelse*. Bergen: Fagbokforlaget.

Topping, K. 2009. Peer Assessment. *Theory Into Practise*, 48:1. 20-27. DOI: 10.1080/00405840802577569

Topping, K and S. W. Ehly. 2009. Peer Assisted Learning: A Framework for Consultation. *Journal of Educational and Psychological Consultation*, 12:2. 113-132. DOI: 10.1207/S1532768XJEPC1202_03

1st Coast Guard District. 2001. *Special Notice to Mariners*.

http://www.uscg.mil/d1/prevention/NavInfo/navinfo/documents/A-Emergency_Procedures.PDF (13.05.17)

6.1 Appendix 1. Evaluation Sheet.

Vurderingsskjema.

Ditt nummer: _____

Partners nr. : _____

Din partners uttale er forståelig, men med en del feil.	Din partners uttale er generelt sett god, men med noen feil.	Din partners utale er veldig god.
Partneren din snakker ikke tydelig.	Partneren din snakker tydelig for det meste.	Partneren din snakker høyt og tydelig gjennom hele gjennomføringen.
Partneren din inkluderer enkelte deler av nøkkelinformasjonen som er relevant for scenarioet.	Informasjonen partneren din kommer med er stort sett relevant for scenarioet.	Partneren din har med all informasjon som er relevant for scenarioet.
Partneren din har valgt feil måte/mottaker for scenarioet.	Partneren din har valgt rett måte og rett mottaker for scenarioet.	-----

Hva var det beste med prestasjonen til partneren din? Her Må du skrive noe.	
Hva er det viktigste partneren din må øve på før neste forsøk? Her MÅ du skrive noe.	

What to look for:

Passengers:	Language: Clear. Professional. Concrete, simplistic language. Slowly and calmly. Act like you are in control. Content: No technical information. Ask them to remain calm.
Coast Guard/Distress call:	Language: Concrete, detailed. Clear and concise. Slowly and calmly. Content: Mayday/Pan-Pan. Identification. Call sign. Position. The nature of the emergency. Number of persons on board.

6.2 Appendix 2. Scenarios Used During the First Test Round.

There are 13 crewmembers on board the ship. The weather is quite still, and the sea is calm. However, for the next few coming days there will be an increase in wind, so the sea is not going to remain calm. The position of your vessel is 13,5 degrees north, and 1 degree west. One of your fellow crewmembers has had a heart attack. Your ship's name is Roald Amundsen. All of the crewmembers have substantial training in giving first aid. Your call sign is 45EE. You have enough fuel and resources to stay out at sea for the next two-three weeks.

You are on board the ship NS CREATOR. You are currently taking in water. The position of your vessel is 10 degrees North, 4 minutes, and 1 degree east, 45 minutes. All the water pumps are fully functioning, and are working to pump out the water. In total, you are 14 crewmembers on board the vessel. Some of the crewmembers have never before experienced taking in water and are starting to get anxious. Your call sign is NS556. After 30 minutes it seems that the water level is not rising.

There has been spilled some oil on the deck. You have 13 crewmembers on board, in addition to 68 passengers. You are on board the ship vessel West Elara. There have been sighted sparks of fire in the engine room earlier, but you are unsure of the current status of the engine. The position of your vessel is 1-degree north, 15 minutes, and 22 degrees west, 15 minutes. A few of the passengers has asked questions about the numbers of crewmembers that seem somewhat stressed out. The call sign of your vessel is WE 771.

Your call sign is HE 455. You are on board the passenger boat called Namsos. There has been sighted fire in the engine room, and passengers is catching on that something is going on. At this moment the captain has already transmitted a mayday call. Your position is 12 degrees north, and 01 degrees east. There has been a slight oil leakage, and it is difficult to say if it is of significance. A helicopter is coming to your aid.

6.3 Appendix 3. Scenarios Used During the Second Test Round.

You have been out at sea for the last three to four weeks. You do not have any number of passengers on board, but are a total of 16 crewmembers on board. The weather is quite still, and the sea is calm. It seems that the engine has stopped working after several weeks of malfunctioning. Your position is 65 degrees north, 47 minutes and 18 seconds. 11 degrees west, 26 minutes and 22 seconds. One of your fellow crewmembers has food poisoning, and has fallen ill the last few days. As a result of this the workload is quite substantial for the rest of the crew, and you are starting to feel exhausted. The vessels name is Bergen. You have a

shipment of coffee beans that has expired. The call sign of your vessel is BN35. Your mechanics have managed to operate the engine up to now.

There are 15 crewmembers in total. You are on board the vessel Eidsvoll. You have a lifeboat available. One of your crewmembers has fallen over board, and is currently staying quite close to the boat. You have a number of passengers on board, and they are all carefully monitoring the situation. Several of them has volunteered to help out, but they seem to more of a hindrance than a resource. Your call sign is ES45. There is additionally an oil leakage on board, but you have integrated good procedures for this. Your position is 65 degrees South, 13 minutes and 5 seconds, and 13 degrees east, 45 minutes and 57 seconds. The number of passengers on board is 54.

You have a sufficient number of licenced firefighters on board. The call sign for your vessel is EE67. You have two civilians on board. Your position is 34 degrees north, 14 minutes and 12 second, and 1 degree east, 55 minutes and 36 seconds. There are a total of 15 crewmembers on board. There has been appearing sparks of fire in the engine room. A pan-pan message was transmitted from the vessel just yesterday. The passengers are catching on that something is going on. The safety instructions can be found beneath the passenger's seats. It is of utmost importance that the passengers remain calm. Your ship's name is Striker.

Your position is 13 degrees north, 13 minutes and 45 second, and 45 degrees east, 53 minutes and 25 seconds. You have a numbers of crewmembers that have fallen ill due to food poisoning. Despite the requirement of having multiple licenced firefighters on board you do not have any firefighters on board. Your call sign is ME45. It seems that the provisions have been contaminated, and some of it has been spilled into the ocean. There has been sighted smoke in the engine room, but you are not quite sure. Your ship's name is Narvik. Two weeks ago one of your crewmembers fell overboard. Your position is 13 degrees south, 34 minutes and 13 seconds, and 67 degrees east, 27 minutes and 12 seconds.