Pure

Syddansk Universitet

Technology enhanced peer learning and peer assessment

Henriksen, Christian Bugge; Bregnhøj, Henrik; Rosthøj, Susanne; Ceballos Garcia, Alejandro Rafael; Kaas, Henrik; Harker-Schuch, Inez Estelle P; Andersen, Ingelise; Larsen, Søren; May, Michael

Technology enhanced peer learning and peer assessment

Christian Bugge Henriksen

PhD, Associate Professor
Department of Plant and Environmental Science,
University of Copenhagen

Henrik Bregnhøj

PhD, E-learning Consultant
Center for Online and Blended Learning,
University of Copenhagen

Susanne Rosthøj

PhD, Associate Professor
Department of Public Health,
University of Copenhagen

Alejandro Ceballos

MSc, Assistant Lecturer
Department of Plant and Environmental Sciences,
University of Copenhagen

Henrik Kaas

MSc, Special consultant
IT Learning Center, University of Copenhagen

Inez Harker-Schuch

MSc, Assistant Lecturer
Department of Plant and Environmental Sciences,
University of Copenhagen
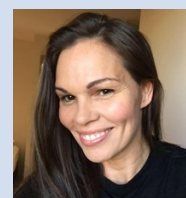
Ingelise Andersen

PhD, Associate Professor
Department of Public Health,
University of Copenhagen

Søren Larsen

MSc, E-learning Project Manager
Digital,
University College Sjælland

Michael May

PhD, Part-time Teacher
Department of Design and Communication,
University of Southern Denmark

The authors have worked together for several years in an online and blended learning project at the University of Copenhagen.

Abstract

This paper explores the application of learning designs featuring formalised and structured technology enhanced peer learning. These include student produced learning elements, peer review discussions and peer assessment in the BSc/MSc level summer course Restoration of European Ecosystems and Freshwaters (REEF), the Master thesis preparation seminars for the Master of Public Health (MPH) and the MOOC course Global Environmental Management (GEM). The application of student produced learning elements and peer review discussions is investigated by analyzing quotes from course evaluations and performing focus group interviews. The application of peer assessment is investigated by analyzing the agreement of peer assessment between students assessing the same assignment. Our analyses confirm previous research on the value of peer learning and peer assessment and we argue that there could also be a huge benefit from developing learning design patterns that facilitate informal peer learning and reinforce knowledge sharing practices.

Introduction

Aligned with the constructivist theory that knowledge development is cultivated through experience, peer-to-peer learning occurs when 'peer-related' students engage in information and knowledge exchange for the mutual benefit of all participating parties; teaching and learning from one another as both part of a planned curriculum and in a broader 'experiential learning' sense. The foundation of peer-to-peer learning originates in the early enculturation process whereby individuals learn socio-cultural 'behaviours' and 'values' from those closest to them - their family, friends and known familiars (Vygotsky 1930-1934/1978). This enculturation forms the springboard for stronger peer-to-peer information and knowledge exchange that takes place during puberty when adolescents transfer respect-authority regimes from their parents to those in a similar age group or 'peers' (Piaget, 1932). This transfer of authority is an important step for peer-to-peer learning as peers replace significant familiars (family) with their peers to obtain new information, ideas and knowledge (or to modify, expand, or reinforce existing knowledge). As the parental-child relationship untethers, the significance of the community of peers increases in the adolescents estimation - so much so that the individual will gladly sacrifice respect and standing with elders rather than risk jeopardising relationships within their peer-community (Burnett & Blakemore, 2009). This pathway, then, forms the basis for self-determined socialisation and the individual's first forays into adulthood; corresponding to the 'second critical stage of intellectual development' that heralds the commencement of abstract reasoning, critical thinking skills and multidimensional problem-solving. As the adolescent moves into adulthood, the frame of 'peers' broadens to include aspects of similar interests (hobbies, careers), social collectives (child-rearing, ideologies) and other familiar concepts - widening the exposure of the adolescent to a broader array of experience, knowledge and information at the very peak of their intellectual development. Individuals derive meaning and trajectory, too, from those around them - lifestyle choices and ideologies are transferred amongst peers; affecting enormous change and upheaval. Adolescents and young adults can be profoundly inspired by

actions and ideals of powerful peers to the point where they will sacrifice their lives to demonstrate their association (e.g. radicalisation). Inspired by the work of Vygotsky and Piaget, Lave and Wenger developed their theory of social learning, including the concept of "situated learning" where peer learning is taking place in a "Community of Practice" that can be both formal and informal (Lave and Wenger, 1990, Wenger, 1998). In a university setting a Community of Practice is formed when students teach and learn from each other either in connection with a course or a project. It follows that peer learning in such a Community of Practice take several different forms:

- Formal peer learning, including student produced learning elements
- Informal peer learning, including and excluding instructor facilitation
- Peer assessment, including peer-reviews

Formal peer learning is when the instructor requires students to take turns in teaching each other. The teaching can happen online or IRL, and it can take the shape of a lecture, a peer review discussion or a student can guide fellow students through an assignment. Formal peer-to-peer teaching also includes student produced learning elements, which can be texts, videos, quizzes/tests, presentations, etc. developed by individuals or groups of students aimed as learning resources for other students to learn from. The benefits of student produced learning elements as exemplified by Hakkarainen et al. (2007) are that students both in the process of designing, producing and engaging in developing case based teaching materials experienced an enhanced positive emotional involvement in the learning process promoting the active and contextual aspects of learning. It is important to note that asking students to produce learning elements for each other, might require that students invest substantial time and effort in gaining e.g. technical competences, video production skills, etc. The instructor need to consider whether these competences and skills are important and matches the learning objectives of the course.

Informal peer learning covers a number of different learning activities: 'Turn to your neighbour' - In a learning environment, students often turn to their neighbour and ask for clarification or explanations when they have not understood something. This can happen during a lecture or when students are working on class work. Turning to your neighbour is, perhaps, the most effortless form of informal peer-to-peer teaching and does not, usually, work to construct highly-complex knowledge.

'Think-pair-share' - This method works in formal classroom environments but embeds informal peer-to-peer interaction within the class. Students are divided into groups, given a topic or problem, told they must 'think' about the problem for some minutes (which promotes higher-quality responses) then they are to pair with a partner and discuss the topic or solution. The instructor then asks a few students to share - thereby increasing the responses from the students as well as the level of response and activity within the classroom.

Vicarious learning - informal interaction amongst peers that serves to improve knowledge or understanding. This occurs when fellow students (with similar status)

share knowledge - but depends on the premise that fellow students have knowledge worth sharing. It is characterised by 'active listening and reflective thinking' (Nehls, 1995) and seeking to understand by being fully absorbed in the activity (Roberts, 2010). The sharing often takes place through 'reflection on practice encounters and relating these through stories to fellow students' (this is also referred to 'culture-making by Bruner, 1986).   This is often coupled with personal and emotional experiences, as well.

Study partnerships or groups - planned or organised studying together can also be considered 'informal' as such activities exist externally to the formal learning environment of their instruction. These allow students to share their understanding and strengthen their knowledge through concentrated study and focus on course material - exchanging their perceptions and reinforcing the context of the course material.

Story-telling/boasting - this is, arguably, the oldest form of teaching in the world and manifests in both formal and informal peer-to-peer group settings. In nursing and medical school environments, practice of procedures is explained to others (often with an element of drama) and the method of the procedure is explained - enabling those listening to 'gain' knowledge or understanding through the story.

Reaching out to another student 'expert' - This is done with note-taking and occurs when a more-able (also more advanced or more senior) student is approached for (or offers to) help a less-proficient student. This is informal as it exists externally from the formal learning environment.

Game-playing - playing quizzes or games with peers often aids in knowledge development. People learn 'by participation' in the game - and, often, through a process of trial and error.

Peer assessment is another form of peer learning that "*can be described generally as a process whereby students evaluate, or are evaluated by, their peers*" (van Zundert et al, 2010), i.e. 'peers' being their fellow students. Peer assessment can take different forms in which students provide qualitative feedback for each other's work (including peer feedback and peer review) and/or give summative grades for written assignments and oral presentations. Dependent on the context and the purpose of using peer assessment, it can be done as a learning activity with open discussions and evaluation or anonymously in a more closed format similar to that of peer review on academic papers. It is crucial that the purpose and clear expectations of using peer assessment is explicitly communicated to students as well as fellow instructors. The primary way of doing this is by preparing clear assessment criteria e.g. as an assessment rubric (Dochy et al., 1999). Airasian and Russell (2008) define assessment rubrics as: "*A rubric is a set of clear expectations or criteria used to help instructors and students focus on what is valued in a subject, topic, or activity*". An assessment rubric is a matrix of explicit criteria that can be used to evaluate student performance.

Objectives

The objectives of this paper are to investigate and evaluate learning designs featuring the application of student produced learning elements in peer-to-peer teaching, and the use of formalized peer feedback and peer assessment in online and blended learning courses at the University of Copenhagen by applying a mixed methods approach.

Methodology

Methodological considerations

Evaluation of peer learning can be done in many ways, depending on what should be evaluated. If the effects of peer learning on learning outcome should be evaluated the most direct way to do this would be to either include or exclude peer learning from a course and determine the difference in final grades. However, this approach is problematic in a university setting since all students should receive the same teaching. Alternatively, questionnaires or focus group interviews could be performed to ask students about their own perceptions of peer learning. This will not directly determine the learning outcome from peer learning, but instead give an idea about whether peer learning is able to create a formal or informal community of practice that will facilitate the achievement of learning outcomes. If peer learning includes elements of peer assessment, it will be important that the homogeneity of the peer assessment is determined in order to verify whether the peer assessment is valid. In this study we are applying a mixed methods approach with questionnaires, focus group interviews and statistical data analysis depending on what is most suitable for the three courses we are investigating. (See Appendix A for a description of the course).

Evaluating experiences with learning designs featuring student produced learning elements

In the Restoration of European Ecosystems and Freshwaters (REEF) course the experiences with student produced learning elements were evaluated in 2011 by extracting data from the students' evaluation and subsequent reflections in the classroom. We used an online questionnaire in order to get the feedback from students on the use of student produced videos in the course. The following six questions with a five-level Likert scale and and an open-answer question was included to investigate how students perceived the creation of the videos as a learning process:

1. To what extent have you experienced the use of video as a motivating factor for your work with the course?
2. In this course, the use of video to document and report the fieldwork was preferable compared to handing in a regular written report.
3. The use of video to communicate scientific work on the internet is a powerful tool that students should learn how to master for use in their future career.
4. The process of making a video instead of a written report gave me a better understanding of the studied subject.
5. The videos produced in this course have great potential in helping rising public awareness about restoration of ecosystems in Europe.

6. The production of a video helped me to work independently and stimulated my creativity
7. What do you think about using Youtube and other online tools in university teaching?

In the subsequent classroom reflections the discussion focused on the following questions: (i) how was it to use another format than a traditional written report to document natural sciences field work? (ii) what is the potential for using the videos as instructor assisted peer to peer teaching material and (iii) what is the potential for using these videos as material to raise public awareness on restoration of freshwater ecosystems?

Evaluating experiences with learning designs featuring peer review discussions

In the Master of Public Health (MPH) experiences with peer review discussions during thesis seminars in 2012 were evaluated with the following questionnaire:

1. Do you recommend that we offer seminars to the 2nd year MPH-students next year? Yes/no
   Please feel free to comment on your answer.

If yes:

1. Which themes should we work with in the seminars?
2. How frequent should we have seminars?
3. Should the seminars be autumn and spring or only spring?
4. Would you prefer seminars in the evening or week-ends instead of two week days?
5. Any suggestions for improving the seminars?

Based on the answers to the answers to the questionnaire focus-group interviews were subsequently made with former and current Danish MPH-students.

Evaluating experiences with learning designs featuring peer assessment

In the Global Environmental Management course experiences with peer assessment was evaluated by performing a statistical analysis of the homogeneity of the peer assessment following the methodology developed by Fleiss et al (2003). In the GEM course peer assessment was performed by using a rubric with a set of criteria for assessment, each with a marking scale with a description of the adequacy needed for each particular mark. For example one of the criteria for assessing the Technology essay was: *Give from 1-4 points depending on how well the technology is described.* The scales were as follows:

- 1 point: This part of the question is not answered
- 2 points: The description is unclear or its purpose / function seems not to be understood by the author
- 3 points: It is described, but I'm in doubt whether its purpose or function is

understood by the author
- 4 points: It is well described, clearly understood by the author

The distribution of the criteria included in the analysis of peer assessment in the GEM course is shown in Table 1

Table 1. Distribution of criteria included in the analysis of peer assessment in the GEM course.

Essay:	Trend	Management	Technology
Number of criteria:	2	4	3
Marking points per criteria:	1-4	1-3	1-4
Maximum mark (points)	8	12	12

A statistical analysis of peer assessment scores were performed for each individual criteria as well as per essay (sum score). Analyses were performed using a random coefficient model subdividing the variance in the observed peer assessment scores into the variance due to the performance of students, the variance due to the peers and the residual error variance. The model was used to assess the inter-rater agreement of the peer assessments, agreement being defined as the degree to which assessments on the same student from different raters are consistent. Formally the agreement is defined as the correlation between two peers rating the same student and is derived from the model as the ratio between the student variance and the total variance (a single measures Intra Class Correlation (ICC), range from 0 to 1). Values close to 0 indicate no agreement, values above 0.75 excellent agreement and values between 0.4 and 0.75 has been suggested to represent fair to good agreement (Fleiss et al.,2003). The model can also be used to derive the agreement of average rather than individual peer scores (the residual error variance is scaled down by k in the ICC formula when an average score based on k peers is used (ICC average measures). A scale can be used to discriminate between the student performances if the residual error variance is small compared to the variability between the students. To assess discrimination, an ICC type measure was defined as the ratio between the student variance and the sum of the student and residual error variance. To evaluate whether the self-assessment scores were systematically different from the peer assessment scores in the GEM course, the above model was expanded with a fixed coefficient distinguishing between self- and peer assessment, thereby allowing for estimation of the mean difference in scores between the two kinds of assessments.

Results

Student produced learning elements in the REEF course

The results of the questionnaire on the students' evaluation of the student produced videos are shown in Table 2. Overall students answered very positively to the

questions. On average 75% of the students strongly agreed or agreed to all six questions, 12.5% of the students answered neutral and 12.5% disagreed or strongly disagreed. More than 75% of the students agreed or strongly agreed that the making of videos were motivating, preferable compared with handing in a report, helped them to work independently, stimulated creativity and should be learned by future students. 56% of the students strongly agreed or agreed that making a video gave a better understanding of the studied subject whereas 29% disagreed or strongly disagreed. This shows that whereas a significant majority of students were generally satisfied with the application of student produced videos in the course, their perceptions of their effects on learning outcome is not unequivocally positive indicating that the learning outcome from applying student produced videos depends very much on the learning preferences of the individual student.

Table 2. Student's evaluation of student produced videos

Question	Very much/ strongly agree	To some extent/ agree	Neutral	It did not motivate me/ disagree	It de-motivated me/ strongly disagree
1.To what extent have you experienced the use of video as a motivating factor for your work with the course?	37.5%	37.5%	12.5%	8.3%	4.2%
2. In this course, the use of video to document and report the fieldwork was preferable compared to handing in a regular written report.	41.7%	41.7%	8.3%	8.3%	0%
3. The use of video to communicate scientific work on the internet is a powerful tool that students should learn how to master for use in their future career.	45.8%	29.2%	25%	0%	0%
4. The process of making a video instead of a written report gave me a better understanding of the	16.7%	41.7%	12.5%	25%	4.2%

studied subject.					
5. The videos produced in this course have great potential in helping rising public awareness about restoration of ecosystems in Europe.	20.8%	50%	8.3%	20.8%	0%
6. The production of a video helped me to work independently and stimulated my creativity	41.7%	45.8%	8.3%	4.2%	0%
7. What do you think about using Youtube and other online tools in university teaching?	Mostly the students answered very positively to this question. Students stated that this is very interesting and useful way of learning and teaching that provided them with new tool adapted to the new development in information sharing. These are just few of the statements we got: *"I think, that soon it will be one of the fundamental ways of learning". "It was very unusually for me, but I think the online tools offer so many new possibilities". "It is an additional tool and a great opportunity to involve and inform people which are not connected to University. Written reports contain more detailed information but videos/pictures address more the emotions of people and are therefore pretty powerful" "Making the film was an interesting experience. I could learn a lot and had a lot of fun :) It was a nice change from the constant sitting at a desk and reading books, notes and articles. Besides images help you better communicate and perpetuate them very well in my head".*				

In the subsequent classroom discussion students stated that they appreciated the use of video as learning tool and they found it highly useful learning how to use video in the documentation of scientific work. The students expressed that they were really happy with the "movie nights", that gave students the opportunity to collectively watch the work of the other students and give direct and instant feedback to each other. This was one of the most valuable experiences in the course both from the instructor's and student's perspective. In connection with the embedded field work where sampling techniques are deployed, students liked that the video gave them the opportunity to perform the actual technique in order to show them in the video assignments. Students also commented on that the ability to use all available resources (voice over, graphs, pictures, written text, music, animations, etc) provided great flexibility, but also made the editing and uploading process more complex and time consuming. During the process of making the videos about their field work (video 2), students spent a lot of time editing and uploading the videos. This was exacerbated by the lack of previous knowledge with editing videos, lack of computers with enough power to handle large

video files, weak internet connection and tiredness since students dedicate time to the video editing and uploading process in the afternoon after a long day of field work. This situation was avoided with the last video (video 3), where students were asked to follow the Dogme concept and they did not use any other resource but just the footage taken on the field and spend much less time in editing and uploading. With respect to the use of the videos as tools for raising public awareness about restoration projects, students expressed that they did not feel comfortable with having their videos available on Youtube, however most of them gave us permission to use the video for marketing the REEF course at academic and public institutions. An important element here is that most of us have an intrinsic fear to be judged by other. The act of making a video with your voice or image in Youtube or any other public platform is a huge challenge for many people.

Peer review discussions in the MPH seminars

The response rate from questionnaires was around 50% = 10 students. All the students recommended the seminars to continue, but the Danish students preferred autumn, whereas the European Master of Public Health students preferred spring:

"I felt like the first set of seminars, before Christmas break, was not the most productive. Joining the EUROPUBHEALTH and the Danish students at that point, and having them "compare" their projects was, in my opinion, counter-productive. The Danish students were a lot further ahead in their projects, and personally comparing my progress with theirs just caused me anxiety. That is why I stopped attending. On the other hand, the presentation of the projects in the spring terms was, I thought, very useful, and very productive, given that we are all so much closer to completing our projects. However, I think that they should all be in the same period, so that the projects would be at comparable stages."

Whereas the students felt it was a benefit to participate in the seminars, there were two huge barriers. One of them was language. A Danish student said:

"Talking English was a great barrier to get 100% of the seminar, and it was difficult to relate to the other international students work, especially because our deadlines were different,(there is one week apart) and they quickly got behind, and were more focused on the methods instead of the discussion, which were more relevant for the MPH-students to discuss."

Another Dane said:

"In my opinion Danish and EUROPUBHEALTH students should not have seminars together ..... many good details and questions get lost, when the seminars are in English."

However. Whereas most of the Danish students were impeded by being taught together with non-Danish students some of them found it *very fruitful from a professional as well as a cultural point of view.*
The other barrier related to presenting in Adobe Connect. The EUROPUBHEALTH easily

felt familiar with making a presentation at Adobe, whereas some of the Danish students were afraid of using a new tool.

"*I have never seen anybody present research or health promotion by means of the internet, so why should we spend our time on preparing a presentation, especially when we are so busy by preparing our master thesis?*"

The instructor´s evaluation of the seminars is that European master of Public Health students really benefitted from presenting in the seminars, maybe because that group of students was closely related and felt a kind of responsibility for each other and therefore commented on the projects. The Danish students who presented were very satisfied afterwards. It took some time to prepare this presentation, but *"it was funny"*. It was especially the younger Danish students who accepted this challenge. This way of presenting topics has been used among global health students who during field trip once a week should present the progress of their proposals to classmates in another country.

This group of students accepted Adobe immediately without questions and got fruitful feed-back. Unfortunately, it was not possible to continue the seminars. The support from the students was weak and since it was a 0 ECTS course it is difficult for instructors to find time to lead the seminars. The seminars closed down in 2015.

Peer assessment in the GEM course

The number of participating students in a MOOC is not well defined, but the sign-ups in the GEM course increased from 1519 to 2736 from the beginning to the end. A total of 199, 170 and 138 students delivered the three essays called 1. Trend, 2, Management and 3. Technology Essays and 138 students delivered all three essays. The distribution of deliverables and self and peer assessment scores in the GEM course are shown in the table:

Table 3. Distribution of deliverables and self and peer assessment scores in the GEM course

Essay:	Trend	Management	Technology
Deliverables	199	170	138
Self assessments	175	149	125
Total assessments incl self-assessments	754	647	522
Total assessments excl self-assessments	579	498	397

One student made exactly 50 peer assessments in each of the essays, which is far higher than anyone else. To avoid placing too much emphasis on one person in the analyses, the assessments from this person have been excluded in all analyses (however similar results are obtained when keeping these assessments in the analyses). Most students did the 3 required peer assessments, and therefore most of the deliverables have also been assessed by 3 peers. The distribution of peer assessments (excluding self assessments) is shown in the below table 4:

Table 4. Distribution of peer assessments (excluding self assessments)

Essay:	Trend (n=199)	Management (n=170)	Technology (n=138)
Peer assessments per delivery:			
1-2	51	26	22
3	115	131	111
4-5	33	13	5
Assessments performed per assessor:			
1-2	1	1	0
3	148	130	108
4-10	28	23	16

The non-regular distribution of number of assessments per delivery and per student (both should be 3) are mainly results of a) Exclusion of the 3*50 assessments by one student (about 10% of the data material), meaningless assessments per delivery, b) The personal wish from some students to do a few assessments more, and b) Possibly limitations in the automatic distribution when some students choose not to rate a delivery they receive. The results from the statistical analysis is presented in Table 5. There are three assignments/essays with a total of nine criteria. The results are analysed both per criteria and per essay. Explanations of the headings are given in the text below. **Table 5. Results from the statistical analysis of peer assessment in the GEM course**

	Tr1	Tr2	M1	M2	M3	M4	Te1	Te2	Te3
Score distribution									
1	6	84	10	23	33	51	6	8	54
2	82	163	118	215	230	172	28	20	50
3	323	286	519	409	384	424	124	112	164
4	343	221					364	382	254
Mean	3.33	2.85	2.79	2.6	2.54	2.58	3.63	3.66	3.18
Mean, % of max	83	71	93	87	85	86	91	92	80
Mean of sum score	6.18		10.51				10.47		
Mean of sum score, % of max	77%		88%				87%		
ICC discrimination	0.19	0.32	0.32	0.23	0.17	0.2	0.32	0.39	0.5
ICC discrimination sum score	0.28		0.37				0.55		
Pairwise differences in peer assessments, %									
0	47%	40%	71%	54%	53%	51%	57%	60%	47%
1	45%	42%	28%	43%	42%	43%	36%	33%	37%
2	8%	16%	1%	3%	5%	7%	7%	7%	12%
3	0%	3%					0%	0%	4%
ICC single measure	0.16	0.27	0.27	0.18	0.15	0.16	0.26	0.29	0.41
ICC average measures, 3 assessors	0.36	0.52	0.53	0.40	0.35	0.37	0.51	0.56	0.68
ICC sum score average measures, single measure	0.23		0.27				0.42		
ICC sum score average measures, 3 assessors	0.48		0.55				0.68		
No. assessors needed for ICC average measures >0.4	4	2	2	3	4	4	2	2	1
No. assessors needed for ICC sum score average measures > 0.4	3		2				1		

No. assessors needed for ICC average measures >0.75	16	9	8	14	17	16	9	7	5
No. assessors needed for ICC sum score average measures > 0.75	10		8				5		
Self assessment higher mean score	0.14	0.28	0.13	0.16	0.22	0.29	0.32	0.34	0.52
95% CI	0.02 - 0.25	0.13 - 0.43	0.05 - 0.20	0.06 - 0.25	0.11 - 0.32	0.18 - 0.40	0.21 - 0.44	0.23 - 0.46	0.35 - 0.69
Self assessment higher score, % of max.	5%	9%	7%	8%	11%	15%	11%	11%	17%
Self assessment higher mean score, sum score	0.41		0.79				1.19		
95% CI	0.19 - 0.64		0.51 - 1.06				0.89 - 1.49		
Self assessment higher score % of max., sums	7%		10%				13%		

The means of the scores are divided by the maximum possible score, resulting in the "% of max" to make them comparable. The average scores for the criteria varies between 71-93% and between 77-88% for the summary of the essays. No student ended below the 60% average of all three essays, that was set as the cut-off mark. The marks are relatively high, but that was also the intention in this MOOC where the students don't obtain ECTS points, that it is possible to pass if you do a fair effort. Those who actually did bother to write an essay rather than just doing the quizzes could be expected to also do the necessary effort. The ICC discrimination is not impressive as such, ranging from 0.28 to 0.55, supporting the observation that most marks are in the higher end. The essential results are the other half of the table summarizing the differences between assessors rating the same essays / students. The "Pairwise differences in peer assessments, %" shows the percentages of the differences between two assessors' assessment of a single criteria that are 0 (the same), 1, 2 or 3. More than half (between 40-71%) of the assessments made are precisely the same when two assessors look at particular criteria in an essay. 28-45% of the differs by 1 score mark and mostly less than 10% (1-19%) differ a score mark of 2 or 3. The ICC average measures show that with three assessors we will have a fair to good reliability on assessments from the summarised (essay-based) results (ICC = 0.48-0.68), whereas on the single criteria level (ICC = 0.36-0.68) the assessment sometimes require four assessors. The results also show that on an assignment/sum score level we would

need 5-10 assessors for the assessment to have an "excellent" strength. Most of the students also did self assessment (89% of the deliverables) and when comparing the scores given as self-assessment with the scores given by the peers the students tend to score themselves higher than their peers, on average between 0.13 and 0.52 higher scores corresponding to 5 to 17% higher than their peers, based on single criteria, or generally 7 to 13% higher based on the collective scoring of each essay.

Discussion and conclusion

Based on a mixed methods approach featuring questionnaires, focus group interviews and statistical analysis depending on the course being investigated, our study has given some valuable insights into the use of student produced learning elements, peer review discussions and peer assessment in a university setting.

By evaluating the use of a learning design featuring student produced learning elements in the REEF course we have demonstrated that student produced videos is a viable alternative to written group assignments be a good complement to the written text as peer to peer learning elements and in the creation and communication of scientific information. Among other things students reported that working with videos were motivating and stimulated both independence and creativity. This corresponds with a study by Hafner and Miller (2011) who found that the use of student produced videos were highly motivating and increased independent learning. Whereas most students generally perceived the use of videos as a good thing, they did not all agree that they were superior to a written report with respect to learning outcome. This contrasts with the studies by Sildis (2006) who found that video projects may increase language learning and Hoogerheide et al (2016) who found that explaining on video was more efficient for learning than explaining in text. The production of videos by students that have been educated in a text based education system imposed a challenge which required students to invest a lot of time and effort in mastering video editing and uploading skills, which may have taken time away from achieving the core scientific learning outcome. However, in a study comparing the quality of an assignment submitted as a video podcast or on a CD Özdener and Güngor (2010) found that there was no difference in the learning outcome and academic quality, the quality of the design and presentation was higher in the video podcast than on the CD. The use of videos published on the internet has the potential to raise public awareness beyond the classroom, but our study shows that it is important to provide proper time and guidance for it.

Even though several universities and international education programmes offer online thesis seminars with peer feedback to students either as a voluntary or mandatory part of their studies, (e.g Aalto University and Cross Border University in Finland, and University of Amsterdam and Leiden University in the Netherlands) we have not been able to find any other studies describing the evaluation of such online thesis seminars. In the MPH seminars it was evident that students who took time to present and comment on presentations by other students really benefitted from the time they spend on this learning activity. Speaking English as well as technical issues were the main barriers

among the Danish students. Since the students learn so much it is recommended that learning designs that integrate thesis seminars and peer feedback are reintroduced in the MPH as a mandatory requirement before the thesis can be submitted. With respect to the use of peer assessment our study of the GEM course has shown that there is a reasonable good agreement between assessors in determining the average score for multiple essays, requiring only three assessors to give a reliable average score, whereas 5 to 10 assessors are required to give reliable scores for individual essays. It follows that courses with a limited number of peer assessed deliverables must secure that each deliverable is assessed by a sufficiently large number of peers to validate the assessment. However, as argued by Suen (2014) there is a risk that unmoderated peer assessment may suffer from lack of credibility. In a MOOC course on human centered interaction design Kulkarni et al (2013) compared student assessments with teacher assessments and found that that even though the majority of students assessed their peers within 10% of the teacher assessments, students generally assessed their peers higher than the teachers (7%). Furthermore, they also assessed peers from their own country higher than peers from other countries. It has been suggested that the reliability of peer assessment in MOOCs may be improved by guidance and training (Ashton and Davies, 2015). Nevertheless, peer assessment has it's limitations and (Admiraal et al, 2015) suggests that in the context of MOOCs self-assessment and peer assessment should be used as assessment for learning and not assessment of learning.

By investigating different approaches to technology enhanced peer learning our study has shown that student produced videos may be used as an alternative to written assignments, but that proper guidance and instruction is needed to secure that students have enough time to achieve the desired learning outcome. Whereas the advantages of peer reviewed online thesis seminars seems straightforward, proper guidance and training is also needed to improve the quality and learning outcome of online peer assessment. An overall conclusion from our study is that the application of learning designs featuring student produced videos, peer reviewed online thesis seminars and online peer assessment are valuable additions to the more traditional instructor driven learning and assessment, but that teacher guidance to facilitate formal peer learning is very important. It should be noted that this conclusion is based on the mixed methods approach applied in the study, and that a more quantitative evaluation of student produced videos and peer reviewed online thesis seminars and a more qualitative evaluation of peer assessments may give other results. Furthermore, it is suggested that future studies should also explore how informal peer learning could be facilitated to reinforce knowledge sharing practices in online and blended learning courses.

References

Admiraal, W., Huisman, B., Pilli, O. (2015). Assessment in Massive Open Online Courses. Electronic Journal of E-learning 13, pp 207-216

Airasian, P. W., & Russell, M. K. (2001). *Classroom assessment: Concepts and applications*. Boston: McGraw-Hill.

Ashton, S., Davies, R.S. (2015). Using scaffolded rubrics to improve peer assessment in a MOOC writing course. Distance Education 36, pp 312-334

Bruner, J.S. (1986). Actual Minds, Possible Worlds. Harvard University Press, Massachusetts, USA

Burnett, S., Blakemore, S-J. (2009). The development of adolescent social cognition. Annals of the New York Academy of Science 1167, pp 51-56

Dochy, F. J. R. C., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. Studies in Higher education, 24(3), 331-350.

Fleiss, J. L., Bruce, L., & Myunghee C. P. (2003). Statistical methods for rates and proportions. John Wiley & Sons.

Hafner, C.A., Miller, L. (2011). Fostering learner autonomy in English for science: A collaborative digital video project in a technological learning environment. Language Learning & Technology 15, pp 68-86

Hakkarainen, P., Saarelainen, T., & Ruokamo, H. (2007). Towards meaningful learning through digital video supported, case based teaching. Australasian Journal of Educational Technology, 23(1), 87.

Hoogerheide, V., Deijkers, L., Loyens, S.M.M., Heijltjes, A., van Gog, T. (2016). Gaining from explaining: Learning improves from explaining to fictitious others on video, not from writing to them. Contemporary Educational Psychology 44-45, pp 95-106

Kulkarni, C., Wei, K.P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., Koller, D., Klemmer, S.R. (2013). Peer and Self Assessment in Massive Online Classes. ACM Transactions on Computer-Human Interaction 20, 33

Lave, J. & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge University Press.

Nehls, N. (1995). Narrative pedagogy: rethinking nursing education. Journal of Nursing Education 34, 204–210

Piaget, J. (1932). The Moral Judgement of the Child. New York: The Free Press.

Roberts, D. (2010). Vicarious learning: a review of the literature. Nurse Education in Practice 10, pp 13-16

Sildis, T.I. (2006). The Effect of a Student Video Project on Vocabulary Retention of First-Year Secondary School German Students. Foreign Language Annals 39, pp 54-70

van Zundert, M., Sluijsmans, D., van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. Learning and Instruction, 20, 270-279.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes (A. R. Luria, M. Lopez-Morillas & M. Cole [with J. V. Wertsch], Trans.) Cambridge, Mass.: Harvard University Press. (Original work [ca. 1930-1934])

Wenger, E. (1998). Communities of Practice: Learning, Meaning, Identity. Cambridge University Press

Özdener, G. and Güngor, Y. (2010). Effects of video podcast technology on peer learning and project quality. Procedia - Social and Behavioral Sciences 2, 2217-2221