Routledge
Taylor & Francis Group

Assessment & Evaluation in Higher Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/caeh20

Online peer feedback on video presentations: type of feedback and improvement of presentation skills

Indira N. Z. Day, Nadira Saab & Wilfried Admiraal

To cite this article: Indira N. Z. Day, Nadira Saab & Wilfried Admiraal (2022) Online peer feedback on video presentations: type of feedback and improvement of presentation skills, Assessment & Evaluation in Higher Education, 47:2, 183-197, DOI: 10.1080/02602938.2021.1904826

To link to this article: https://doi.org/10.1080/02602938.2021.1904826

Published online: 09 Apr 2021.

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
2022, VOL. 47, NO. 2, 183–197
https://doi.org/10.1080/02602938.2021.1904826

Online peer feedback on video presentations: type of feedback and improvement of presentation skills

Indira N. Z. Day, Nadira Saab and Wilfried Admiraal

Leiden University, Leiden, The Netherlands

ABSTRACT

Peer feedback can be a tool to help higher education students develop presentation skills. This process can be supported by using online tools that enable peers to provide feedback annotated to specific moments in the presentation. The current paper investigated whether providing and receiving peer feedback using an online tool was related to improvement in students' presentation skills, whether students provided different peer feedback in comments and annotations, and whether student characteristics played a role in students' presentation skill improvement. Results from 56 second-year undergraduate Education and Child Studies students indicate that providing and receiving feedback did not relate to students' improvement in presentation skill, and that high ability students showed less improvement in presentation skills than medium and low ability students did.

KEYWORDS

Peer feedback; presentations; online feedback

Introduction

During their educational career, students are expected to acquire academic and professional knowledge and skills. Skills that need to be developed during higher education include writing and presentation skills. Whereas students can acquire knowledge by studying textbooks, attending lectures and taking examinations, skill development requires another approach. A main component of skills development is practice. It is important to investigate how students' skill development can be supported.

To further students' skills it is important to provide them with feedback during practice. According to Hattie and Timperley (2007) feedback should help students by letting them know where they are going (with regard to course goals; also called feed-up), how they are doing now (called feed-back), and what the next steps should be to improve and reach their goals (called feed-forward). Several authors have studied the conditions for effective feedback (e.g. Hattie and Timperley 2007; Shute 2008), and focused on the importance of feedback being task or process oriented, instead of evaluating the ability of the feedback receiver.

Feedback can be provided by the teachers, but also by fellow students. Topping (1998) suggests that Vygotsky's concept of scaffolded learning (Vygotsky 1978) may be suitable for a setup of peer feedback where peers identify weaknesses and strong points. This process of peer feedback often provides students with feedback at their specific level of need, which can help them in processing the feedback. Gikandi and Morrow (2016) suggested that peer feedback

CONTACT Indira N. Z. Day i.n.z.day@iclon.leidenuniv.nl

may also help improve students' self-regulated learning and reflection. Welsh (2012) noted that students valued peer feedback as much as tutor feedback, however Admiraal (2014) suggested that students prefer teacher feedback over peer feedback and are more willing to apply teachers' feedback.

Topping (1998) wrote one of the first large reviews on peer feedback and found that peer feedback is usually as beneficial to students learning as instructor feedback is. A meta-analysis conducted by Li et al. (2020) even suggested that peer assessment is more effective than teacher assessment. The majority of the research on peer feedback has focused on peer feedback on writing tasks. Writing competence is an important skill to be developed in higher education, but students are expected to develop a wider set of skills than just writing. Next to writing, presenting is another skill that is of great importance to learn. De Grez (2009, p. 5) defines presentation competence as 'the combination of knowledge, skills, and attitudes needed to speak in public in order to inform, self-express, to relate and to persuade'. Topping noted that there was a limited amount of research into peer feedback for presentation skills. Over two decades later, there have been several studies investigating peer feedback and presentations (e.g. De Grez, Valcke, and Roozen 2012; van Ginkel et al. 2017a, 2017b, all comparing peer feedback to teacher feedback), but 'peer feedback on writing' research still seems to be more ubiquitous (cf. Li et al. 2020). The current study tries to expand the knowledge about peer feedback on presentations by investigating what kind of feedback peers provide, incorporating the use of video and an online feedback platform.

Peer feedback on writing and presenting in higher education

Research on peer feedback on writing tasks is more prevalent, but it can be assumed that several mechanisms of peer feedback transcend specific task characteristics. The following section first discusses literature from writing research, focused on how feedback can improve performance and the role of student characteristics, before discussing results from research regarding peer feedback on presentations.

Several researchers have indicated that peer feedback is not only beneficial for student learning of the students who *receive* the feedback, but that the peer feedback *provider* may also benefit from the peer feedback process. Lundstrom and Baker (2009), for example, compared peer feedback providers with peer feedback receivers and found that the former showed bigger improvement on writing tasks. Additionally, Mulder, Pearce, and Baik (2014) investigated student perceptions of peer feedback and concluded that, after providing peer feedback, some students indicated that they had learned the most from providing reviews. In his review, Topping (1998) already suggested that providing peer feedback is a reflective process, which could be interpreted as learning by assessing.

It is important to note that the benefits of peer feedback for students' learning and performance may be influenced by student characteristics, like their motivation and self-efficacy for the task, their reactions to the peer feedback they received, and their beliefs about peer feedback in general. Shute (2008) suggested that feedback should not threaten students' self-esteem. Therefore, it seems intuitive that students' self-efficacy for the task at hand may play a role in the relationship between peer feedback and performance. Hsia, Huang, and Hwang (2016) indeed found that students with higher self-efficacy for the task and higher levels of intrinsic motivation achieved higher scores on their final performance, irrespective of the type of peer feedback they received. Furthermore, they also suggested that students' self-efficacy and motivation were not affected by the peer feedback process. In a similar vein, Strijbos, Narciss, and Dünnebier (2010) investigated whether perceptions of the received peer feedback influenced students' performance. They found no relation between the reported peer feedback perceptions and students' performance on the final assignment. Huisman et al. (2019) suggest that students'

beliefs about peer feedback may influence their perceptions of peer feedback, and their behaviour during the peer feedback process.

With regard to presenting, van Ginkel et al. (2015) reviewed the body of research into presentation skills for the context of higher education. In their paper they stress the importance of presentation skills as an outcome of higher education and present a set of design principles for the development of these skills. Some of these principles are to make sure that the presentation task is considered relevant by students, that students should be able to observe models, that students should be able to practice, and that involvement in peer feedback should be encouraged.

Some of these design principles for the development of students' presentation skills are also embedded in tasks where students provide peer feedback on presentations. For example, the process of providing peer feedback also requires students to observe peers. Additionally, formative peer feedback requires students to practice their presentation, before presenting for a grade, since this kind of peer feedback is mostly provided on non-final presentations. Patri (2002) found that students are able to rate presentations of their peers on a similar quality level as instructors do, provided that they are provided with a firm set of criteria.

However, providing peer feedback on a presentation can be challenging for students. They need to focus on the content as well as presentation skills, and when the presentation in class is over, they cannot watch it again. To be able to provide more effective peer feedback on presentations, peers should be able to view a video recording. In this way, they can re-watch the video at their own leisure and rewind specific sections. Palao et al. (2015) demonstrated that video feedback by peers or teachers led to increased physical education performance in high school students compared to verbal teacher feedback. In the context of higher education both Leijen et al. (2009) and Hsia, Huang, and Hwang (2016) investigated peer/self-feedback on video in groups of dance students and found positive results. Leijen et al. (2009) found that students valued the opportunity for self-reflection by being able to re-watch their own performance over the opportunity to watch the performance of peers. Hsia, Huang, and Hwang (2016) expressed that students that provided feedback ratings as well as feedback comments gave similar scores as teachers did. Furthermore, they found that these feedback comments are of higher quality than those of students who only provided comments. Additionally, Admiraal (2014) discovered that students preferred feedback annotated to their presentation videos more than general feedback, because it was linked to their current performance more closely.

The use of video recordings for assignments where students provide peer feedback on presentations also facilitates the use of digital tools for providing peer feedback. According to Li et al. (2020) computer mediated peer assessment appeared to have a larger effect than pen and paper assessment. Furthermore, digital tools have significant added value for presentation assignments. Where an assignment focusing on providing peer feedback on writing will probably take similar shape whether papers are exchanged physically, via email or using a digital environment like Turnitin or Pitch2Peer, the added value of some digital environments for peer feedback on presentations is that students can annotate their comments to specific instances in the video, which would be impossible to do during live presentations. This linking to specific instances also helps with contextualising the feedback, another design principle suggested by van Ginkel et al. (2015).

In the current study, we investigate the role of peer feedback on presentations, using a digital tool, in improving students' presentation skills. The discussion of the literature highlights the benefits of both providing and receiving feedback, and the importance of student characteristics. Additionally, using digital tools to annotate videos may be beneficial. However, to our knowledge, these different facets have not been combined in a study investigating presentation skills, therefore our research answers the following questions:

1. How do provided and received peer feedback relate to the improvement of students' presentation performance?

2. Does students' feedback differ between annotations and general feedback comments?
3. How does students' ability relate to the improvement of students' presentation performance?
4. How do students' peer feedback perceptions, peer feedback beliefs and self-efficacy for presenting relate to peer feedback?

Methods

Participants

Participants in this study were 56 second-year Bachelor of Education and Child Studies students of a research university in the Netherlands. Education and Child Studies is a major predominantly chosen by female students, and 94% of our participants is female. The average age of participants was 19.6 years (*SD* 3.0 years). A majority of participants had experience with providing peer feedback on presentations and feedback on videos (88% and 84%, respectively).

Procedure

This study was approved by the research ethics committee at our institution (no. 2017-07). Data collection took place during the 2017–2018 academic year. Informed consent letters were handed out and students were invited for participation during an earlier tutorial group. Students were required to give a presentation on one of the Master of Education and Child Studies specialisations in the context of a course focusing on academic skills and job orientation. Depending on the size of their tutorial group, students presented alone, in duos or in trios. Most presentations were performed by duos, and therefore presenters will be referred to as presenting duos from now on. The presentation lasted seven to eight minutes and students needed to discuss which courses were a part of the specialisation, whether the specialisation had specific requirements for admission and what the career opportunities for that specialisation are. In preparation for their presentation in class, all students filmed a trial presentation and uploaded it into a Pitch2Peer module on Blackboard.

Three weeks before the final presentation was due, the first author gave an introductory lecture focusing on presentation skills, peer feedback skills, and how to work with the Pitch2Peer platform. After this, students had one week to film and upload their trial presentation, followed by one week to provide online peer feedback. After the peer feedback process, presenting duos had one final week to adapt their presentations based on the feedback before the final presentation in class. A schematic overview of the tasks for students and the researcher each week can be seen in Figure 1.

All participating students completed the questionnaire in class before they held their final presentation. Final presentations were filmed by the first author or a student assistant. Only duos where both students agreed to participate in the research were filmed. Table 1 shows an overview of the total dataset.

Trial presentations, peer feedback comments and annotations were downloaded from the Pitch2Peer platform and in all cases students' names were replaced by student IDs to ensure confidentiality.

Variables

Presentation

Trial and final presentations were rated for quality by the first author and a student assistant, using a presentation rubric consisting of the same four categories (content, speech, presence

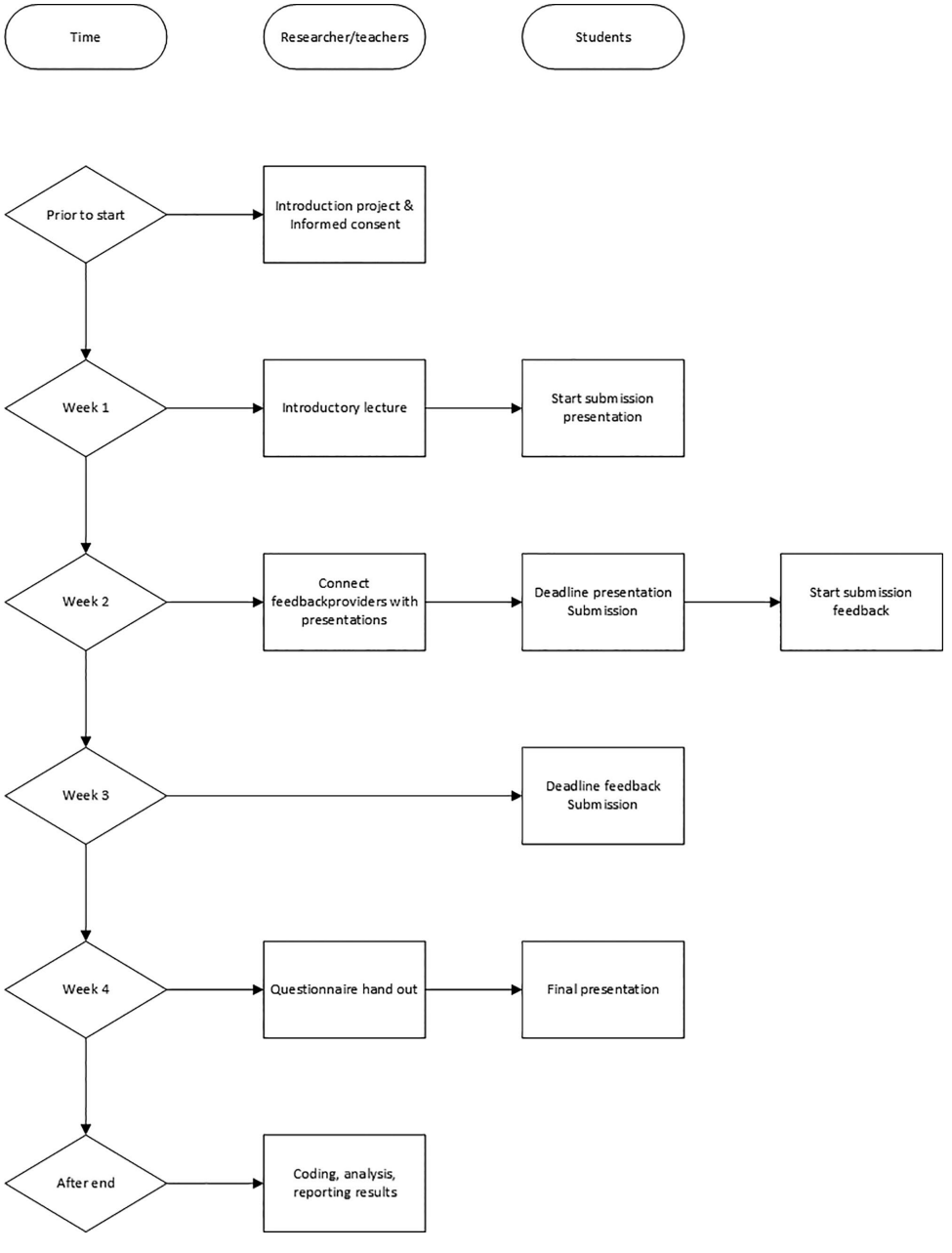

Figure 1. Schematic overview of student and researcher tasks.

Table 1. Overview of the dataset.

Data	*N*
Participating students	56
Trial presentations	33
Final presentations	33
Feedback providers	54
Providers per presentation	1-3
Peer feedback perceptions and beliefs questionnaire	52
Presentation Self-efficacy questionnaire	53

and audio-visual tools) that students used to rate the videos. Each category consisted of 4–6 subcategories. An overview of the categories with a short description can be found in the supplemental material. For the content category, subcategory three was discarded for analysis. This category focused on whether students had all the mandatory components in the presentation, and since all students did it was not informative. For each subcategory, raters gave a score of 1–5, which was then transformed to a mean score per main category. Videos were analysed using a cross design, where raters did not rate trial and final presentation of the same students. Students' trial presentation scores were used as a presentation ability measure.

Peer feedback

Using Pitch2Peer, each presenting duo rated one presentation (1–5 stars) on four criteria, focusing on (1) content of the presentation, (2) speech of the presenters, (3) presence of the presenters, and (4) audio-visual tools used, and gave an overall score. Furthermore, students were required to write a short paragraph with general feedback comments and annotate at least one moment in the video with specific feedback for that moment.

For coding peer feedback, event sampling has been used. This means that each code refers to a meaningful unit, which can be a sentence fragment or a sentence. A new coding unit was started when a different coding category of the rubric applied, or when a new sentence started in the same coding category. For example, the sentence 'You made a nice introduction, it's immediately clear what kind of master it is, and that it's relatively small' was divided in three fragments, where the clauses separated by commas were separately coded.

Following earlier research (van den Berg, Admiraal, and Pilot, 2006; Huisman et al. 2017) on peer feedback comments, the comments and annotations were coded using feedback *functions* (analysis, evaluation, revisions, and possible elaborations on the final two) and feedback *aspects* (in this case the four main categories of the rubric: content, speech, presence and audio-visual tools). Each fragment was coded using a matrix of the functions and aspects, and the number of comments in each category was counted and used for analysis.

Questionnaire

The questionnaire was used to measure several student characteristics that the literature suggests may be important in the peer feedback process, which are students' peer feedback perceptions (Strijbos, Narciss, and Dünnebier 2010), peer feedback beliefs (Huisman et al., 2019) and presentation self-efficacy (Hsia, Huang, and Hwang 2016).

Peer feedback perceptions. Students' peer feedback perceptions were measured using an adapted version of the peer feedback perception questionnaire developed by Strijbos, Narciss, and Dünnebier (2010; Dutch translation by Agricola et al. 2016). This questionnaire consists of six scales with three items each, which participants rate on a five-point Likert scale where 1 = not at all applicable to me, and 5 = totally applicable to me. The scales are (1) fairness, (2) usefulness, (3) acceptance, (4) willingness, (5) positive affect, and (6) negative affect. Items were slightly adapted to reflect that students answered the questionnaire after they had used the peer feedback to improve their presentations. An example item from the fairness scale was 'I consider the feedback I received justified' Reliability and descriptive statistics are presented in Table 2.

Peer feedback beliefs. The Beliefs about Peer Feedback Questionnaire (BPFQ) developed by Huisman et al. (2019) was used to measure peer feedback beliefs. Four scales measured (1) students' valuation of peer feedback as an instructional method (4 questions), (2) students' valuation of peer feedback as an important skill (3 questions),

(3), students' confidence in their own peer feedback (two questions), and (4) students' confidence in peer feedback from their peers (two questions). All questions are rated on a five-point Likert scale ranging from 1 = completely disagree to 5 = completely agree. An example item from the valuation of peer feedback as an important skill is 'Being capable of dealing with critical peer feedback is an important skill'. Reliability and descriptive statistics are presented in Table 2.

Self-efficacy for presentations. A nineteen-item questionnaire based on the questionnaire used by Adams (2004) was used to measure presentation self-efficacy. This questionnaire consists of four scales measuring students' self-efficacy for speech (seven questions), audio-visual (renamed from *display* in the original for consistency with the rubric categories; four questions), content (five questions) and presence (three questions). For each question participants rate how well they believe they can perform specific tasks for the final presentation on a five-point scale ranging from 1 = not well at all to 5 = very well. This questionnaire was translated to Dutch by the first author and the translation was checked, using a back-translation procedure, by a native speaker of English at our institution. All items asked students 'How well do you believe you can perform the following tasks during your final presentation'; a task from the speech scale was 'Speak with the appropriate level of formality'. Reliability and descriptive statistics are presented in Table 2.

Reliability

Data of one tutorial group (consisting of six trial videos, 13 feedback comment files, and seven final videos) was used to calculate inter-rater reliability for videos and comments. For the inter-rater reliability of annotations eight feedback comment files from the same tutorial group were used. Inter-rater reliability for the videos, the comments and annotations was calculated using intraclass correlations (ICCs). For the videos mean scores per rubric category were calculated per rater, whereas for the comments and annotations mean scores were calculated per feedback function per rater. Subsequently ICCs and 95% confidence intervals were calculated using a two-way mixed-effects model, looking at absolute agreement of average measures (Landers, 2015). According to Koo and Li (2016), ICCs lower than .50 are considered poor, ICCs between .50 and .75, between .75 and .90, and above .90 represent moderate, good and excellent reliability respectively. Table 3 displays ICCs for the videos, and Table 4 shows ICCs for feedback comments and annotations.

After reliability was calculated, the first author and a student assistant each coded the videos and comments of separate tutorial groups. Since the agreement on comment annotations was so high, all annotations were coded by a student assistant.

Table 2. Descriptive statistics for peer feedback beliefs and perceptions, and presentation self-efficacy.

Peer FB beliefs					Peer FB perceptions					Presentation self-efficacy				
	α	*N*	*M*	*SD*		*α*	*N*	*M*	*SD*		*α*	*N*	*M*	*SD*
Instructional method	.697	47	3.66	0.57	Fairness	.716	52	4.02	0.55	Content	.685	53	3.54	0.49
Professional Skill	.614	53	4.23	0.46	Usefulness	.808	52	3.47	0.69	Speech	.552	53	3.67	0.39
Confidence Own	.588	53	3.80	0.51	Acceptance	.663	47	4.68	0.35	Presence	.667	53	3.70	0.60
Confidence Peers	.788	53	3.53	0.68	Willingness	.844	52	3.57	0.89	Audio-visual	.638	53	3.91	0.43
					Positive Affect	.810	52	3.58	0.66					
					Negative Affect	.919	52	1.27	0.54					

Table 3. Intraclass correlations for presentations.

		95% Confidence interval	
	ICC	Lower	Upper
Content	.623	.135	.835
Speech	.651	-.059	.869
Presence	.602	.137	.821
Audio-visual	.791	.520	.908

Table 4. Intraclass correlations for feedback functions in comments and annotations.

	Comments			Annotations		
		95% Confidence interval			95% Confidence interval	
	ICC	Lower	Upper	ICC	Lower	Upper
Analysis	.980	.932	.994	n.a.	n.a.	n.a.
Evaluation	.935	.779	.981	.990	.956	.998
Explanation of evaluation	.763	.132	.933	.954	.790	.991
Revisions	.793	.119	.944	1.00	n.a.	n.a.
Explanation of revisions	.744	.187	.924	1.00	n.a.	n.a.

Note. Both raters did not use the code analysis across the eight files for annotations; therefore ICC could not be calculated. Codes for revisions and explanation of revisions were 100% corresponding between raters.

Statistical analyses

To investigate how providing and receiving peer feedback influences students' presentation performance, we conducted analyses of variance (ANOVAs) and regression analyses.

For all students a mean score for their trial and final presentation was calculated by calculating a mean score across the four main category scores. To investigate whether the type of feedback comments students receive influences their improvement from trial to final presentation, a presentation improvement score was calculated by subtracting the trial presentation score from the final presentation score.

Results

Table 5 shows the mean scores for all four separate rubric categories and the mean score across categories for both the trial and the final presentation. Repeated measures ANOVAs indicate that scores of the presentation increased from the trial to the final presentation, for each category of the rubric (results not pictured) and for the mean total score across categories, $F(1,49) = 56.30$, $p < .001$.

Relationship between received and provided feedback and improvement of presentation performance

To investigate the role of received and provided feedback, sum scores for each feedback function were calculated across feedback aspects, resulting in a total of ten sum scores, five for each function in received comments and five for each function in received annotations. With these new scores, two linear regressions were conducted. For the first analysis, the five independent variables were the amount of received comments for each function, and the dependent variable was the total presentation improvement score (see Table 5). Results of the regression indicate that the received general feedback comments do not influence students' improvement (all 5

β's < .05, n.s. R^2 = .02). A similar result came from the analysis of the feedback annotations, (all 5 β's < .10, n.s R^2 = .11).

Since providing feedback may be more beneficial for students' improvement than receiving feedback, similar feedback function sum scores were calculated only for the amount of general feedback comments each student provided. A regression analysis with the five provided feedback function variables as independent variables and the presentation improvement score as the dependent variable showed that this was not the case for our sample. Students' improvement is independent of the amount of feedback they provided, (all 5 β's < .20, n.s, R^2 = .07).

Differences in provided feedback between feedback comments and annotations

Tables 6 and 7 show the total amount of feedback comments and annotations that students provided. Across aspects it is apparent that the most frequent occurring feedback function is evaluation, where students gave implicit or explicit quality statements about different aspects of the presentation.

To look at the differences in provided feedback between the general comment and the annotations, raw scores were transformed to proportions, displayed in Table 8. We see similarity in the amounts of provided feedback functions and aspects across feedback comments and

Table 5. Mean scores for trial and final presentations.

	Trial presentation		Final Presentation		Improvement	
Category	*M*	*SD*	*M*	*SD*	*M*	*SD*
Content	3.83	0.41	4.12	0.48	0.27	0.55
Speech	3.62	0.57	3.95	0.56	0.32	0.72
Presence	3.44	0.55	3.94	0.49	0.50	0.63
Audio-visual	4.11	0.50	4.40	0.42	0.29	0.55
Total	3.71	0.36	4.10	0.33	0.39	0.36

Table 6. Descriptive statistics of provided feedback comments.

	Aspects							
Functions	Content				Speech			
	Tot N	Mean	*SD*	Mode	Tot N	Mean	*SD*	Mode
Analysis	4	0.07	0.43	0	0	0.00	0.00	0
Evaluation	200	3.70	2.09	5	148	2.74	1.56	2
Explanation of Evaluation	43	0.80	1.56	0	21	0.39	0.56	0
Revisions	45	0.83	1.19	0	56	1.04	0.95	1
Explanation of Revisions	20	0.37	0.65	0	22	0.44	0.72	0
	N	Mean	*SD*	Mode				
Other	76	1.41	1.51	0				
Off Task	9	0.17	0.38	0				

Table 6. Continued.

	Aspects							
	Presence				Audio-visual			
Functions	Tot N	Mean	*SD*	Mode	Tot N	Mean	*SD*	Mode
Analysis	2	0.04	0.27	0	0	0.00	0.00	0
Evaluation	86	1.67	1.74	1	76	1.41	1.09	1
Explanation of Evaluation	23	0.43	0.79	0	16	0.30	0.57	0
Revisions	33	0.61	0.90	0	20	0.37	0.68	0
Explanation of Revisions	11	0.20	0.45	0	7	0.13	0.44	0

Table 7. Descriptive statistics of provided feedback annotations.

	Aspects							
	Content				Speech			
Functions	Tot N	Mean	*SD*	Mode	Tot N	Mean	*SD*	Mode
Analysis	10	0.21	1.18	0	0	0.00	0.00	0
Evaluation	81	1.72	1.91	0	69	1.47	2.12	0
Explanation of Evaluation	13	0.28	0.65	0	8	0.17	0.48	0
Revisions	26	0.55	0.80	0	23	0.49	1.04	0
Explanation of Revisions	17	0.36	0.82	0	3	0.06	0.25	0
	N	Mean	*SD*	Mode				
Other	8	0.17	0.48	0				

Table 7. continued.

	Aspects							
	Presence				Audio-visual			
Functions	Tot N	Mean	*SD*	Mode	Tot N	Mean	*SD*	Mode
Analysis	0	0.00	0.00	0	0	0.00	0.00	0
Evaluation	57	1.21	1.30	0	13	0.27	1.06	0
Explanation of Evaluation	15	0.32	0.59	0	3	0.06	0.25	0
Revisions	28	0.60	1.33	0	8	0.17	0.38	0
Explanation of Revisions	10	0.21	0.55	0	2	0.04	0.20	0

Table 8. Proportions of feedback provided per function and aspect in comments and annotations.

	Aspects							
Functions	Content		Speech		Presence		Audio-visual	
	Comment	Annotation	Comment	Annotation	Comment	Annotation	Comment	Annotation
Analysis	0.004	0.025**	0.000	0.000	0.002	0.000	0.000	0.000
Evaluation	0.218	0.206	0.161	0.175	0.94	0.145**	0.083	0.033**
Explanation of Evaluation	0.047	0.033	0.023	0.020	0.025	0.038	0.017	0.007
Revisions	0.049	0.066	0.061	0.058	0.036	0.017**	0.022	0.020
Explanation of Revisions	0.022	0.043**	0.024	0.008*	0.012	0.025	0.008	0.005
	Comment	Annotation						
Other	0.083	0.020**						
Off Task	0.010	0.000*						

Note. Significant differences in proportions of provided feedback are denoted in the annotation column.
*$p < .05$, **$p<.01$.

annotations. *Z*-tests were conducted to test whether there were significant differences between the proportions, and results are presented in Table 8.

These results suggest that more evaluation and revision feedback on students' presence is provided in annotations, compared to general comments. Furthermore, students ask more content related questions and further explain their content revisions in the annotations, indicated by the higher proportion of content analysis and explanation of revision comments in the annotations.

Differences in improvement for students with different presentation ability

To examine whether students' presentation skills were related to their improvement and the feedback they provide, students were divided in three ability groups of equal size. The 18 students in the low ability group scored an overall score of 3.55 or lower on the trial

presentation, the 17 students with a score between 3.56 and 3.84 were classified as medium ability, and the 18 students scoring 3.85 or higher were classified as high ability.

Analysis of variance suggests students in the different ability groups showed different improvement, $F(2,47) = 12.01$, $p < .001$. Bonferroni Post Hoc tests indicated that low and medium-ability students showed more improvement than high-ability students, $p < .001$ for low-ability and $p = .035$ for the medium-ability students, but that low and medium-ability students did not show differences in improvement. A second analysis of variance indicated that the three ability groups did not differ on their final presentation scores, $F(2,47) = 2.03$, $p = 0.143$.

Relationship of feedback beliefs and perceptions and presentation self-efficacy with presentation performance improvement

Mean scores for students' peer feedback beliefs, peer feedback perceptions and presentation self-efficacy can be found in Table 2. As can be seen, students generally have positive perceptions and beliefs with regard to peer feedback. Interestingly, students rate their own peer feedback skills as significantly higher than the skills of their peers, $t(52) = 2.70$, $p = .009$. Overall, students' presentation self-efficacy is good, indicated by all scales having means higher than 3.

Neither students' peer feedback beliefs nor their peer feedback perceptions were related to their improvement from trial to final presentation (β's for all 3 scales < .10, n.s. R^2 .02 for beliefs, and β's for all 4 scales < .30, n.s. R^2 .01 for perceptions). Students' self-efficacy for presenting, similarly, does not relate to the improvement: β's for all four scales < .20, n.s. $R^2 = .05$.

To see whether students' peer feedback beliefs and presentation self-efficacy related to the feedback students provided, separate regression analyses were conducted using the four content related categories of provided feedback comments as outcome variables and for each analysis all the peer feedback beliefs scales and the relevant presentation self-efficacy scale (e.g. self-efficacy about the content of the presentation for comments related to the content) as predictors. None of these analyses returned significant results.

Discussion

The current study investigated students' use of a digital tool to provide peer feedback. Our first question focused on whether students' provided and received peer feedback related to the performance on students' final presentations. Students' presentation scores did improve from the trial presentation to the final presentation, but this improvement was not related to the amount or type of peer feedback students provided and received.

For the second research question we looked deeper into the functionality of the tool to try and see whether students provide different feedback in a general comment than when they give feedback annotations. Results indicated that, in providing feedback, the majority of students focused on evaluating their peers' presentations, and mostly provided similar feedback in general comments and annotations.

The final two research questions investigated the relationship of student characteristics by looking at students' presentation ability and the role of students' peer feedback perceptions and beliefs and their presentation self-efficacy. Results indicated a statistically significant improvement in presentation score between the trial and final presentation, but high-ability students showed less improvement than low and medium-ability students. This may be caused by a ceiling effect, since the mean score for the trial presentation already was 3.7 out of 5. The ceiling effect is further evidenced by the three ability groups no longer differing on the final presentation scores. With regard to the other student characteristics, students' peer feedback beliefs and perceptions and presentations self-efficacy did not relate to their improvement, or the feedback they provided.

The most striking finding might be that feedback (either provided or received) seemed to not be related to students' improvement, even though previous research (e.g. Topping 1998; Lundstrom and Baker 2009; Li et al. 2020) has indicated that providing or receiving peer feedback is related to improved performance. A possible explanation for this finding could be that the majority of feedback comments could be coded as evaluations. This means that peers mostly gave information about the current state of the presentation, often without providing recommendations for improvement. With regard to the feedback framework of Hattie and Timperley (2007) this means that students often did not provide feed-forward. According to Hattie and Timperley (2007, p. 90) '[t]his feedforward question can have some of the most powerful impacts on learning'.

That students did improve their presentation skills, regardless of the feedback they provided or received, could be explained by several mechanisms. First of all, Boud and Molloy (2013) suggest that curricula should include ample opportunities for students to produce work central to learning outcomes, and in this case the peer feedback process helped students gain more insight in features of good presentations and presentation skills. Furthermore, students were able to reflect on their own performance, because they could re-watch the video of their trial performance. Self-reflection can improve learning through similar mechanisms as peer feedback, like the increased time on task and reflecting on assessment criteria (Topping 1998).

Our results expand on results found by Huisman et al. (2017, 2018) who found that the nature of received peer feedback was not related to students' writing improvement and that students' peer feedback perceptions were not related to an increase in writing performance. The current study applied the same coding scheme to identify the nature of peer feedback and used the same questionnaire for feedback perceptions, but in the context of a presentation assignment.

Previous research by Van der Pol, Admiraal, and Simons (2006) suggested that anchored discussions (where comments, and following replies, could be annotated to text) led to more constructed knowledge than traditional forum discussions. The current study did not employ a between-subjects design where the influences of annotations and general comments can be directly compared. Yet, analysis did not portray a large difference in types of feedback provided in the two different feedback modes, and neither of them influenced students' performance.

Students' beliefs and perceptions about peer feedback were not related to their results. Students answered questions about their peer feedback beliefs and perceptions right before they had their final presentation, and thus *after* they had already incorporated the feedback, which may have influenced their answers on the questionnaire. However, previous research (Strijbos, Narciss, and Dünnebier 2010; Huisman et al. 2018) also found no relation between peer feedback beliefs or perceptions and results.

Limitations and suggestions for further research

The current study focuses on a relatively small sample of students in one bachelor programme of one university. Since presentation skills are important for professions across the spectrum, it is important that the effects of peer feedback on these skills are studies beyond Child and Education Studies.

Research by van Zundert, Sluijsmans, and van Merriënboer (2010) suggests that peer feedback training greatly improves students' peer feedback skills and their attitudes towards peer feedback. In the current study, students did not follow a peer feedback training, which may have hindered their ability to provide high quality peer feedback. However, all students did follow a short introduction on how to provide peer feedback. In addition, almost 90% of students had prior experience with providing peer feedback on presentations. We have no information, though, whether these students applied adequate peer feedback or what they have learned from these experiences.

With regard to the students' presentations not improving as a function of received peer feedback it should be noted that all students gave relatively good trial presentations, with an average score of 3.7 for the trial. Research with a more mixed ability group may show different results.

Concluding remarks

Irrespective of feedback, students have improved their performance from trial to final presentation. This may be because of looking at another presentation (modelling) or because students re-watched their own presentation and had time to reflect on facets that they could improve. Another factor that may have been influential is that students were provided with an extra opportunity for practice, which helped them hone their skills. Since studies focusing on peer feedback on presentations are still relatively scarce, and our results remained inconclusive, we suggest that peer feedback on presentations should be further examined in the future.

Acknowledgements

The authors would like to thank Lise van der Plas and Elja Vellekoop for their assistance in coding the videos, comments and annotations.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Indira N. Z. Day completed her doctoral dissertation about the characteristics, perceptions, and results of continuous assessment in higher education at the Leiden University Graduate School of Teaching in 2018, where she also worked as a researcher studying the use of peer feedback on video presentations. She currently works as a post-doctoral researcher at Utrecht University where she studies the professional development of teachers in higher education and studies student well-being at University of Applied Sciences Leiden.

Nadira Saab is a professor at ICLON (Graduate School of Teaching), Leiden University, the Netherlands. In addition, she is a strategic advisor at Kennisnet (a public organisation for education and ICT). Her research covers several topics, including technology enhanced education, diversity and inclusiveness, motivation, formative assessment, professional development of teachers, and collaborative learning. She supervises several PhD students and participates in various national and international research projects. She is project leader of the 'Kenniswerkplaats Diversiteit' (Knowledge Lab Diversity) of which the aim is to construct knowledge about (cultural) diversity and inclusiveness. Saab is chair of the board of the special interest group Learning and Instruction of the VOR (Vereniging Onderwijs Research), the national organisation of Educational Research in the Netherlands, and she is member of the board of the Catharine van Tussenbroek Fonds. Saab is a Fellow at the Leiden Teachers Academy, which is aimed at developing and sharing good practices of (higher) education with other teachers at Leiden University.

Wilfried Admiraal (Leiden University) is full professor Teaching and Teacher Education and director of the Leiden University Graduate School of Teaching. He graduated in Social Psychology and in 1994 he defended his PhDthesis on teachers' coping strategies with stressful classroom events. His research interest combines the areas of teaching, technology and social psychology in secondary and higher education. He is and has been involved in various national and international educational projects on the use of technology in secondary education, teacher education and higher education. He supervises some 30 PhDs and published over 200 peer-reviewed journal articles. He is editorial board member of various scientific journal such Educational Psychology and Teaching and Teacher Education. He is president of the Netherlands Association of Educational Research (VOR). More information on projects and publications: https://sites.google.com/site/wilfriedadmiraal/.

ORCID

Indira N. Z. Day http://orcid.org/0000-0002-5511-1061
Nadira Saab http://orcid.org/0000-0003-0751-4277
Wilfried Admiraal http://orcid.org/0000-0002-1627-3420

References

Adams, K. 2004. "Modelling Success: Enhancing International Postgraduate Research Students' Self-Efficacy for Research Seminar Presentations." *Higher Education Research & Development* 23: 115–130. doi:10.1080/0729436042000206618.

Admiraal, W. F. 2014. "Meaningful Learning from Practice: Web-Based Video in Professional Preparation Programmes in University." *Technology, Pedagogy and Education* 23 (4): 491–506. doi:10.1080/1475939X.2013.813403.

Agricola, B., F. J. Prins, M. F. van der Schaaf, and J. van Tartwijk. 2016. "Changes in Students' Self-Regulated Learning, Feedback Perception and Motivation during Undergraduate Research Projects." *Paper Presented at the European Association for Research on Learning and Instruction (EARLI)*, Munich, Germany.

Boud, D., and E. Molloy. 2013. "Rethinking Models of Feedback for Learning: The Challenge of Design." *Assessment & Evaluation in Higher Education* 38 (6): 698–712. doi:10.1080/02602938.2012.691462.

De Grez, L. 2009. "Optimizing the Instructional Environment to Learn Presentation Skills." Doctoral diss., Ghent University.

De Grez, L., M. Valcke, and I. Roozen. 2012. "How Effective Are Self-and Peer Assessment of Oral Presentation Skills Compared with Teachers' Assessments?" *Active Learning in Higher Education* 13 (2): 129–142. doi:10.1177/1469787412441284.

Gikandi, J. W., and D. Morrow. 2016. "Designing and Implementing Peer Formative Feedback within Online Learning Environments." *Technology, Pedagogy and Education* 25 (2): 153–170. doi:10.1080/1475939X.2015.1058853.

Hattie, J., and H. Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112. doi:10.3102/003465430298487.

Hsia, L.-H., I. Huang, and G.-J. Hwang. 2016. "Effects of Different Online Peer-Feedback Approaches on Students' Performance Skills, Motivation and Self-Efficacy in a Dance Course." *Computers & Education* 96: 55–71. doi:10.1016/j.compedu.2016.02.004.

Huisman, B., N. Saab, J. van Driel, and P. van den Broek. 2017. "Peer Feedback on College Students' Writing: Exploring the Relation between Students' Ability Match, Feedback Quality and Essay Performance." *Higher Education Research & Development* 36 (7): 1433–1447. doi:10.1080/07294360.2017.1325854.

Huisman, B., N. Saab, J. van Driel, and P. van den Broek. 2018. "Peer Feedback on Academic Writing: Undergraduate Students' Peer Feedback Role, Peer Feedback Perceptions and Essay Performance." *Assessment & Evaluation in Higher Education* 43 (6): 955–968. doi:10.1080/02602938.2018.1424318.

Huisman, B., N. Saab, J. Van Driel, and P. Van Den Broek. 2020. "A Questionnaire to Assess Students' Beliefs about Peer-Feedback." *Innovations in Education and Teaching International* 57 (3): 328–338. doi:10.1080/14703297.2019.1630294.

Koo, T. K., and M. Y. Li. 2016. "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research." *Journal of Chiropractic Medicine* 15 (2): 155–163. doi:10.1016/j.jcm.2016.02.012.

Landers, R. N. 2015. "Computing Intraclass Correlations (ICC) as Estimates of Interrater Reliability in SPSS." *The Winnower* 2: e143518. 81744. doi:10.15200/winn.143518.81744.

Leijen, Ä., I. Lam, L. Wildschut, P. R.-J. Simons, and W. Admiraal. 2009. "Streaming Video to Enhance Students' Reflection in Dance Education." *Computers & Education* 52 (1): 169–176. doi:10.1016/j.compedu.2008.07.010.

Li, H., Y. Xiong, C. V. Hunter, X. Guo, and R. Tywoniw. 2020. "Does Peer Assessment Promote Learning? A Meta-Analysis." *Assessment & Evaluation in Higher Education* 45 (2): 193–211. doi:10.1080/02602938.2019.1620679.

Lundstrom, K., and W. Baker. 2009. "To Give is Better That to Receive: The Benefis of Peer Review to the Reviewer's Own Writing." *Journal of Second Language Writing* 18 (1): 30–43. doi:10.1016/j.jslw.2008.06.002.

Mulder, R. A., J. M. Pearce, and C. Baik. 2014. "Peer Review in Higher Education: Student Perceptions before and after Participation." *Active Learning in Higher Education* 15 (2): 157–171. doi:10.1177/1469787414527391.

Palao, J. M., P. A. Hastie, P. G. Cruz, and E. Ortega. 2015. "The Impact of Video Technology on Student Performance in Physical Education." *Technology, Pedagogy and Education* 24 (1): 51–63. doi:10.1080/1475939X.2013.813404.

Patri, M. 2002. "The Influence of Peer Feedback on Self and Peer-Assessment of Oral Skills." *Language Testing* 19 (2): 109–131. doi:10.1191/0265532202lt224oa.

Shute, V. J. 2008. "Focus on Formative Feedback." *Review of Educational Research* 78 (1): 153–189. doi:10.3102/0034654307313795.

Strijbos, J. W., S. Narciss, and K. Dünnebier. 2010. "Peer Feedback Content and Sender's Competence Level in Academic Writing Revision Tasks: Are They Critical for Feedback Perceptions and Efficiency?" *Learning and Instruction* 20 (4): 291–303. doi:10.1016/j.learninstruc.2009.08.008.

Topping, K. 1998. "Peer Assessment between Students in Colleges and Universities." *Review of Educational Research* 68 (3): 249–276. doi:10.3102/00346543068003249.

van den Berg, I., W. Admiraal, and A. Pilot. 2006. "Peer Assessment in University Teaching: Evaluating Seven Course Designs." *Assessment & Evaluation in Higher Education* 31 (1): 19–36. doi:10.1080/02602930500262346.

van der Pol, J., W. Admiraal, and P. R. J. Simons. 2006. "The Affordance of Anchored Discussion for the Collaborative Processing of Academic Texts." *International Journal of Computer-Supported Collaborative Learning* 1 (3): 339–357. doi:10.1007/s11412-006-9657-6.

van Ginkel, S., J. Gulikers, H. Biemans, and M. Mulder. 2015. "Towards a Set of Design Principles for Developing Oral Presentation Competence: A Synthesis of Research in Higher Education." *Educational Research Review* 14: 62–80. doi:10.1016/j.edurev.2015.02.002.

van Ginkel, S., J. Gulikers, H. Biemans, and M. Mulder. 2017a. "Fostering Oral Presentation Performance: Does the Quality of Feedback Differ When Provided by the Teacher, Peers or Peers Guided by Tutor?" *Assessment & Evaluation in Higher Education* 42 (6): 953–966. doi:10.1080/02602938.2016.1212984.

van Ginkel, S., J. Gulikers, H. Biemans, and M. Mulder. 2017b. "The Impact of the Feedback Source on Developing Oral Presentation Competence." *Studies in Higher Education* 42 (9): 1671–1685. doi:10.1080/03075079.2015.1117064.

van Zundert, M., D. Sluijsmans, and J. van Merriënboer. 2010. "Effective Peer Assessment Processes: Research Findings and Future Directions." *Learning and Instruction* 20 (4): 270–279. doi:10.1016/j.learninstruc.2009.08.004.

Vygotsky, L. S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge MA: Harvard University Press.

Welsh, M. 2012. "Student Perceptions of Using the PebblePad e-Portfolio System to Support Self- and Peer-Based Formative Assessment." *Technology, Pedagogy and Education* 21 (1): 57–83. doi:10.1080/1475939X.2012.659884.