International Journal of Vocational Education Studies 2025 Vol. 2/No. 1
https://doi.org/10.14361/ijves-2025-020104

Supporting Peer Feedback Formulation in a Video-Based Digital Platform in Technical Vocational Education

Sietse Brands, Bas Kollöffel, Elwin Savelsbergh & Maaike Endedijk

Abstract *Vocational education and training (VET) students often have difficulties with formulating peer feedback for various reasons. This is a problem as only well-argued peer feedback positively affects learning. Thus, supporting peer feedback formulation is desirable. This study explores the effectiveness of a prompt consisting of evaluative markers and sentence openers on the quality of peer feedback formulation. 48 VET students in a technical domain in the Netherlands watched the same five videos on practice and then provided peer feedback to the performers. An experimental group (24 students) had access to evaluative markers and sentence openers to support feedback formulation. The control group (24 students) did not have access to these supports. The generated feedback was analysed on feedback quality indicators and compared using independent sample t-tests. The results indicated that prompt-based supports can improve the quality of student-formulated peer feedback. We discuss the importance of carefully designed prompts and how they can contribute to better peer feedback formulation.*

Title *Supporting Peer Feedback Formulation in a Video-Based Digital Platform in Technical Vocational Education*

Keywords *Feedback Formulation; Vocational Education; Video Annotation; Sentence Openers; Evaluative Markers*

Dr. Sietse Brands, Dr. Bas Kollöffel, Dr. Elwin Savelsbergh & Prof. Dr. Maaike Endedijk (Professional Learning and Technology, University of Twente; Professional Learning and Technology, University of Twente; Research Group Curriculum Development in Primary and Secondary Education, HU University of Applied Sciences Utrecht; Professional Learning and Technology, University of Twente; Professional Learning and Technology, University of Twente);
s.brands@utwente.nl; b.j.kolloffel@utwente.nl; Elwin.savelsbergh@hu.nl; m.d.endedijk@utwente.nl;
https://orcid.org/0009-0009-8561-0740;https://orcid.org/0000-0002-1044-3345;https://orcid.org/0000-0001-7140-8634;https://orcid.org/0000-0001-9193-8018;

1 Introduction

Peer feedback is becoming more popular as a form of formative feedback. It requires students to take up the teacher role and (formatively) assess the performance of their peers, and is expected to contribute to learning effectively (Harris & Brown, 2013; Panadero & Lipnevich, 2022; Shute, 2008). Peer feedback has several advantages over tutor feedback, as giving and receiving peer feedback requires students to take a critical view towards the practice and opinions of their peers, which may lead to improved understanding and performance over tutor feedback (Gielen et al., 2010). This is because peer feedback is analysed more critically by the receiver, as peers are novices that could give inaccurate or even incorrect feedback.

But to be effective, feedback has to be formulated in such a way that it is accepted by the receiver (Alqassab et al., 2023; Panadero & Alqassab, 2019; Strijbos et al., 2010). Prior research established that effective peer feedback is specific and concrete and that it contains congruent reasoning (Hattie & Timperley, 2007; Shute, 2008). But how specificity and justifications of peer feedback can be enhanced has not been studied extensively. Furthermore, peer feedback is common in domains such as health care and teacher education, but not in other vocational education and training (VET) domains, such as commercial and technical domains. Because of this, students in these domains have little experience with giving and receiving peer feedback, and are therefore also less likely to engage in critical thinking activities themselves (Jossberger et al., 2010).

Past research on peer feedback has been focused on other aspects of peer feedback. Most notably, feedback accuracy, meaning whether peer feedback is consistent with that of experts (Andrade, 2019). Therefore, most supportive measures for formulating peer feedback have been designed to help provide feedback on the elements that experts pay attention to. As a result, rubrics have become very popular supports (Panadero & Jonsson, 2020). But rubrics do not help students build a clear justification for their feedback.

Furthermore, students have difficulties writing feedback that is specific and descriptive and need to be supported to do so (Boldrini & Cattaneo, 2014). Thus, further support to generate effective feedback is necessary, not only in terms of what to focus on, but also on how to formulate feedback. This is especially important in VET, as students often do not have sufficient language skills to write cohesive argumentation for their feedback. For example, demographic studies show a decrease in literacy skills (OECD, 2013), and teachers report that VET students have difficulties with reading comprehension and writing (Jossberger et al., 2015). As language skills of students may be inadequate, we argue that supporting students with writing activities such as formulating peer feedback is of great importance.

A possible way of support is to provide students with sentence openers. Sentence openers are predefined phrases to start a sentence (van Joolingen et al., 2005) and have been used in online learning environments to support students in formulating hypotheses (van Joolingen et al., 2005) and arguments (Yiong-Hwee & Churchill, 2007). This type of support has not received a lot of research attention as of yet. Thus, it remains unclear under which conditions it provides favourable results.

Therefore, exploring the effects of sentence openers on formulating peer feedback is the focus of the current study. Specifically, we focus on exploring the effects of support-

ing VET students with sentence openers to write well specified and justified feedback. In general, we expect that better facilitated peer feedback may lead to better learning results and a more critical outlook on the practice VET students reside in in their daily or prospective jobs. The current study is conducted in a technical VET domain.

2 Theoretical Framework

2.1 Defining and Exploring Peer Feedback in VET

Feedback is a broad concept that can be defined and executed in various ways. In general, feedback is aimed at identifying mistakes and good practice, questions on decisions and performance, and providing suggestions for future performance to learn and improve practice (Alqassab et al., 2023; Gielen et al., 2010; Panadero & Lipnevich, 2022). Feedback can be directed at the task, process, self-regulation or personal level (Hattie & Timperley, 2007), be given through various modes (e.g. in person, written or video-recorded) and by various actors (e.g. self, peer, expert, computer) (Shute, 2008). Different types of feedback have their place in education and research and will depend on the learning task at hand.

In VET, various forms of feedback can be found. For example, e-learning environments allow for predefined feedback on mistakes given by a computer. In apprenticeships, in-person feedback by a tutor is more common (Mikkonen et al., 2017; Schaap et al., 2023). Further, workplace simulations or video recordings allow peers to also evaluate performance and provide additional feedback, either in-person or in writing (Gavota et al., 2010; Ortoleva & Bétrancourt, 2016). There is a difference between peer and tutor feedback. Gielen et al. (2010) describe that the main difference is the authority of the feedback provider. Tutors are experts in their field, while peers are not. Because of this, the accuracy and correctness of peer feedback may differ from teacher feedback. Despite this, peer feedback can be even more effective than tutor feedback, due to the uncertainty of feedback correctness. This uncertainty triggers the process of reflection in the feedback receiver (Gielen et al., 2010). Furthermore, also the feedback giver can benefit from peer feedback, training their noticing skills, feedback formulation and developing a critical stance on practice (Panadero & Lipnevich, 2022) and personal interdependence (Gielen et al. 2010). We conclude that peer feedback is an important learning activity in VET.

While peer feedback was found to be helpful for the learning processes in VET in general, VET students generally have little experience with evaluating the performance of others (Cattaneo & Motta, 2020; Schaap et al., 2012), as they tend to lean on the evaluations and feedback of their tutors and colleagues more than their own and peer student evaluations (Josserger et al., 2010; Mikkonen et al., 2017). This trend is underlined by very little scientific publications on the topic of peer feedback in VET. In the current study, we focus on feedback given by a student that is directed at a peer. This poses specific challenges that we will discuss in the next paragraphs.

2.2 Peer Feedback Quality

The process of generating and receiving effective peer feedback is delicate. Kollar and Fischer (2010) describe the process of peer feedback as one that moves from task performance to evaluation, to feedback formulation, to feedback reception to task revision. This means that generating feedback requires specific attention to how the feedback is received, taking into account the feelings and understanding of the receiving end. Then, also the receiving end needs to make an effort to process the feedback, understanding the intentions, reasoning and value of the received feedback (Panadero & Lipnevich, 2022).

An idea of what high quality peer feedback entails also needs to be explicated. We adopt the peer feedback quality indicators of Gielen et al. (2010), who proposed several criteria indicators for written peer feedback. The most relevant criteria indicators are specificity of the feedback, justification of feedback and providing suggestions for alternatives. Furthermore, the researchers mention appropriateness of feedback compared to assessment criteria and clear formulation.

First, feedback should be sufficiently concrete and specific (Alqassab et al., 2023; Gielen et al., 2010; Shute, 2008). Concrete feedback means that the feedback refers to specific behaviours or aspects of the performance. In other words, the feedback does not just state: "Well done", but includes a clear topic to indicate *what* was done well. Second, justification of the feedback is of importance. This requirement is related to concreteness of feedback, indicating the reasoning behind both positive and negative feedback. As peers are not authorities in the domain, their feedback might be inaccurate, incomplete, or even plain incorrect (Strijbos et al., 2010). Therefore, justification is necessary, as it allows the receiver to better understand and evaluate the quality of the feedback, prior to revision of the performance (Gielen et al., 2010). Thirdly, and in line with justification, Hattie and Timperley (2007) conclude that both positive and negative feedback can be beneficial, but this depends on the focus of the feedback. In general, negative feedback yields positive learning outcomes, especially if aimed at the performance and provides a form of justification. And better yet, suggestions for future practice should be included (Gielen et al., 2010). Positive feedback can also yield beneficial results on motivation, but generally has a lower effect on cognitive learning outcomes compared to negative feedback. Further, appropriateness of feedback towards assessment criteria is seen as an important aspect of peer feedback. As is the way feedback is formulated, meaning addressing the receiver personally (Gielen et al., 2010).

2.3 Supporting Peer Feedback Quality

There are several effective supports to help students formulate quality peer feedback. The general consensus is that students need to be directed in their attention, to focus on relevant performance aspects. For this, the use of rubrics and prompts are common supports. A rubric is a list of performance aspects, guided by indications for good and bad quality performance of these aspects (Panadero, 2017). Rubrics have been studied extensively in formative assessment contexts, and were found to positively influence the way students provide feedback, as they help students focus on relevant performance quality criteria and to more accurately evaluate the quality of their performance (Panadero &

Jonsson, 2013, 2020). The effectiveness of rubrics has been studied often and proven to a sufficient degree. Because of this saturation, the current study will not focus on rubrics, but rather on prompting students in various ways.

The use of prompts was found to be an important way of supporting students in feedback activities. A prompt can be defined as a written question or statement directed at the performance, with the aim to direct student attention on specific elements or to stimulate them to write feedback in a specific way (Gielen et al., 2010). In general, the effectiveness of prompts for feedback purposes is acknowledged, allowing for more and higher quality feedback and reflection (Cattaneo & Motta, 2020; Gielen et al., 2010). The implementation of prompts is rather diverse, as they can take many different forms and may address various topics.

Mostly, prompts are predefined questions that either pop-up at a certain time during the activity, or are visible to students during the whole activity (Kori et al., 2014). But besides prompting students with questions, we argue that prompts can also manifest as sentence openers. With sentence openers, students do not receive directive questions, but are prompted to finish predefined sentences (Gielen et al., 2010; van Joolingen et al., 2005). While sometimes mentioned in the literature, sentence openers have not been studied often. The available studies do indicate a positive attitude towards their use, especially for students new to formative assessment (Aamri, 2018; Barkhuizen & Wette, 2008; Farkas, 2019). Sentence openers may reduce cognitive load for students, as they do not have to think about formulating their feedback as much. On the other hand, forcing students to use sentence openers might have adverse effects on creativity and motivation, which may reduce students willingness to provide feedback (Farkas, 2019; Gielen et al., 2010). We argue that this downturn is not to be expected with VET students with lower language proficiency and little experience with peer feedback formulation, but that they may in fact help these students formulate higher quality feedback than they would without sentence openers.

Using symbols, such as traffic lights to prompt evaluation is another way of prompting, and has not been researched much. Such evaluative markers come in the colours green, yellow and red, respectively representing good, questionable and negative performance. In this sense, traffic light colours prompt students to evaluate. Similar to written prompts, evaluative prompts allow learners to quickly pick an evaluation and apply it to the performance, again possibly lowering cognitive load. While such evaluative markers have been mentioned in a couple of studies (Colasante & Douglas, 2016; Harris & Brown, 2013; Hulsman & van der Vloodt, 2015; Lai et al., 2020; Lavoué et al., 2015), instances of their use are scarce and the effects of these supports have not been thoroughly studied yet. Thus, exploring the effects of evaluative markers on peer feedback seems to be a worthwhile endeavour, especially in a domain where peer feedback is not commonly used yet.

2.4 Current Study

In the current study, we study the effectiveness of two types of prompts, evaluative markers, and sentence openers, on peer feedback quality produced by VET students in a digital video-annotation environment. Our research question is as follows: Does a prompt-

based support consisting of evaluative markers and sentence openers promote the quality of peer feedback formulated by VET students in a video-based digital platform? Based on the literature, we expect that this prompt-based support may increase peer feedback quality.

3 Method

3.1 Participants and Procedure

Participants in this study were dual VET students in the technical domain of Electrician Education in the Netherlands. In dual VET, students work four days a week as an apprentice and visit school the other day. Participants for this study were enrolled at various vocational schools and participated on their school day. The study procedure was approved by an ethics committee of the University of [redacted]. After agreeing to participate, each student was invited into a face-to-face session with the researcher. The researcher explained the ethical research affordances, the study procedure and finally the video-annotation software. There were no incentives for participation.

Participants were asked to give feedback on the videos as if they would write the feedback to a peer. Participants were told to focus on general aspects of performance in the domain, these being safety, procedure, and workplace organisation. The subject of feedback was a set of five short (one minute) videos of performance of common tasks by anonymous apprentices in the domain. Thus, each participant gave feedback on the same five videos. The same researcher was present during the process, primarily to answer technical questions about the video-annotation software. From pilot studies, we knew students had little experience with feedback and that it was important to help them get started by asking some generic questions without value judgement, like: "What did you think of this performance?" and "what would you say to this person if he was your peer?"

Feedback will be given through means of video-annotation. Video-annotation is a relatively new tool that allows users to create time-based comments in a video (Evi-Colombo et al., 2020). With video-annotation, instances of feedback are directly related to specific moments in a video. Participants were randomly assigned to either a control or experimental group. The experimental group had access to supports embedded in the digital video viewing environment. The control group used the same environment, but without the embedded supports. The supports embedded into the video-annotation software were evaluative markers and sentence openers. The evaluative markers represented traffic light colours: red, yellow, and green. When they selected any of these markers, it would be linked to a specific time in the video. Additionally, selecting the marker would initiate a comment, with a sentence opener. The sentence openers were respectively "Not good, because", "I have some doubts or a question, namely" and finally "Good, because". The control group did not have access to these supports and had to write their feedback from scratch.

3.2 Video Vignettes

Participants were asked to watch five videos on workplace practice. The videos all captured a part of a common domain specific procedure, performed by a student, and were recorded using a head-mounted camera. The idea behind these videos was to give a first person view of the workplace, showing the work-in-progress, tool usage and workplace organization. We asked experts to review the videos to check whether there are some typical mistakes and good behaviours in them for students to comment on. The content of the videos was as follows:

- Clip 1: a student conducting a last-minute risk analysis before working on a large consumer unit indoors.
- Clip 2: a student working on a transformer outside. He is in the process of creating an earth screen by twisting copper wires and then cutting them to attach them to the installation.
- Clip 3: a student working in a pit outside. He is in the process of branching a cable under voltage to connect a house to the power grid.
- Clip 4: a student working on a large indoor consumer unit similar to the first video. This student has started work and is close to finishing.
- Clip 5: a student measuring voltage on an electricity meter that is being replaced.

3.3 Data Analysis

In this study, students created video-annotations that contained written feedback. They would write one or more annotations per video. The video-annotations that participants created were analysed on feedback quality. In principle, each video-annotation was seen as a unit of analysis. But students would sometimes only use one annotation to comment on multiple aspects of performance. In these cases, we took the feedback on each topic separately. For the remainer of this paper, we will refer to the unit of analysis as a 'comment'.

Each comment was analysed on feedback quality indicators following a coding scheme. Coding was done dichotomously. For each comment we analysed which quality indicators were present. Those present were coded as a '1' and those that were not, were coded as '0'. Specifically, we coded feedback on the valence of feedback, meaning whether it was positive, critical or a question, specificity, and presence of justifications and suggestions. Definitions of these codes are presented in Table 1. This coding scheme is an adaptation of the work of Gielen et al. (2010). The scheme of Gielen et al. (2010) was applied to a context in which students wrote longer sections of feedback, not related to a video. Because of this difference in context and feedback mode, we could not apply their framework directly. We did not include the balance between positive and negative feedback, as we only found five of these instances in our data sample. The elements appropriateness of the feedback compared to assessment criteria and clear formulation (Gielen et al., 2010) were deemed less important for the current study and were thus omitted. This is because in our study, a wide range of performances was presented, which could also be interpreted in multiple ways. Further, we deemed it inappropriate

to include feedback formulation in this simulated context, as students did not know the persons, they would give feedback to.

We found there are two ways students constructed their sentences, and these greatly influenced coding of their feedback. In the first way, students would formulate feedback as follows: [specific topic] is [evaluation], because [reasoning]. The other way is [evaluation], because [specific topic], which means [reasoning]. In this second formulation, a pitfall in coding was that we were inclined to code the specific topic as the reasoning. This can be illustrated in the following sentence: "Not good, because you did not wear gloves." In this case, one could easily identify not wearing gloves as a justification for the evaluation. However, it does not provide any reasoning why this is wrong. A solid justification would include an argument such as "you risk getting a shock so close to live wires". Because of these findings, we decided the term 'reasoning' was better fitting than 'justification' for our data.

A second rater coded 50 comments, which were compared and discussed with the first rater, finally resulting in a kappa of 0.82 for the nature of comments, $k = 0.87$ for specificity, $k = 0.83$ for reasoning and $k = 0.84$ for suggestions.

Table 1: Overview of the Feedback Aspects and Definitions

Feedback quality		**Definition**	**Example**
Nature of comment	Positive	Positive feedback on the performance was given.	'Nicely organised tools.'
	Critical	Negative feedback on the performance was given.	'Do not place your tools on the ground.'
	Question	Questions raised to clarify aspects of the performance or the reasoning behind the way of working	'Why did you measure voltage on so many places?' 'What are you going to measure?'
Quality indicators	Specific	The comment explains *what* is right or wrong, clearly relating to a specific aspect of the performance.	'He's wearing his safety gloves.'
	Reasoning	The comment explains *why* something is right or wrong and/or *how* something might affect outcomes.	'You should press this three times, not two.' 'Tighten it right away, or you might forget.' 'You might trip over your tools.'
	Suggestion	The comment includes a suggestion for an alternative.	'I would punch holes instead of jabbing my instrument into the cable.'

After categorization and evaluating the quality indicators, we performed independent sample t-tests to compare the experimental and control group to analyse whether the experimental group created higher quality feedback.

4 Results

In total, 48 students participated in this study, of which 24 were assigned to the experimental group and 24 to the control group. Overall, the data sample consisted of 293 feedback comments. The experimental group created significantly more comments ($M = 6.67$, $SD = 1.58$) than the control group ($M = 5.54$, $SD = 2.04$) at ($t(46) = -2.135$, $p = .038$). The effect size for this was medium with a Cohen's d of 0.62. Most feedback was critical, as 168 comments consisted of critical feedback and 99 were positive. Not many questions were posed (26 comments). In some cases, feedback was clearly critical but was presented by students as a question. In those instances, we chose to code the comment as 'critical', regardless of the question form. Thus, the comments that are defined as questions in this data, do not directly contain a value judgement on the performance.

Table 2: Overview of the Feedback Nature per Group

	No support			**With support**			**t-test**
	%	*M*	*SD*	*%*	*M*	*SD*	*p*
Positive	29	1.58	1.98	38	2.54	1.69	.078
Critical	67	3.71	2.23	49	3.29	1.81	.484
Question	5	0.25	0.68	13	0.83	1.09	**.032**
Total	100	5.54	2.04	100	6.67	1.58	**.038**

Table 2 presents an overview of the feedback nature per group. In this table, ratios and means are used to compare the nature of comments between groups. For example, the percentage ratio shows that in the control group, 67 percent of comments contained critical (negative) feedback, while in the support group this was 49 percent. The t-tests indicate that the support group created less critical comments than the control group, but significantly more questions (medium effect size at Cohen's $d = 0.64$). Although not significant, the support group also appeared to create more positive comments than the control group.

Table 3: Overview of the Feedback Quality Indicators per Group

	No support			**With support**			**t-test**
Quality indicator (n)	%	*M*	*SD*	%	*M*	*SD*	*p*-value
Unspecific (44)	34	0.11	0.32	66	0.18	0.39	.103
Specific (249)	47	4.92	2.02	53	5.46	2.27	.386
Reasoning (112)	44	2.04	2.18	56	2.63	2.28	.369
Suggestion (102)	46	1.96	1.40	44	2.29	1.57	.442

The results of the analysis of quality indicators of feedback are presented in Table 3. Characteristics of feedback quality were the presence of specificness, reasoning and suggestions for improvement. In this table, we reported the percentage ratios and means per quality indicator per group. We included both specific and unspecific comments as mutually exclusive quality indicators to obtain deeper insight in the data. The quality indicators specific, reasoning and suggestion are not exclusive, which means that multiple quality indicators could be present in a single comment. From this data, it becomes apparent that for the total dataset, most comments were specific, but that other indicators were found less often in both groups. Independent t-tests indicated no significant differences between groups.

As the number of comments between groups differs, further analysis of the quality indicators is performed based on the nature of the comment. Similar to Table 3, Table 4 presents the quality indicators per group, but now split out between positive and critical comments. This allows for a more comparable view of the comments. The distribution of positive and critical comments is quite skewed, possibly because of the relative ease with which students can create a positive remark in the support group, using the evaluative markers. Stating something was good requires less effort than saying it was wrong, as this requires reasoning.

Table 4: Overview of the Feedback Quality Indicators Found Within Positive and Critical Comments per Group

		No support		**With support**		**t-test**
		M	*SD*	*M*	*SD*	*p*-value
Positive comments	Unspecific	0.46	0.72	1.21	1.25	**.015**
	Specific	0.96	1.55	1.33	1.79	.441
	Reasoning	0.17	0.38	0.92	1.21	**.006**
	Suggestion	0	0	0.08	0.28	.162

		No support		With support		t-test
Critical comments	Unspecific	0.04	0.20	0	0	.328
	Specific	3.67	2.28	3.29	1.81	.531
	Reasoning	1.83	2.14	1.50	1.53	.538
	Suggestion	1.92	1.44	2.00	1.69	.855

Table 4 shows that the support group more often created positive comments that contained reasoning. Furthermore, the support group created more unspecific positive comments than the control group did. For the other quality indicators, no notable differences were found.

5 Discussion

The current study was conducted to explore the effects of embedded feedback support in a video annotation environment on peer feedback quality. Our research question was aimed at finding out whether prompt-based supports improved the quality of peer feedback. Based on our results, we can conclude that the support helped students formulate higher quality feedback, mostly in the form of a higher quantity of comments and more reasoning for positive comments. No increase was found for other feedback aspects such as the specificity or suggestions. Further, students with access to support asked more questions to the performer than the control group did. Our overall conclusion is that both the marker and sentence opener prompt affected student feedback.

5.1 Interpreting the Results

We found that the support led to students creating more comments than without the supports. We believe this was caused by the relative ease of creating comments in the support group. Because of the supports, students only needed to click on a marker to make a positive comment, whereas the experimental group would need to think about a formulation, requiring a lot more effort from the students. Similarly, students in the experimental group also asked more questions than the control group. It seems that students are more likely to use the markers to create different types of feedback than just critical feedback. We believe this can be interpreted as a positive outcome, as balanced feedback valence is believed to be most effective for learning (Prilop et al., 2021). Based on the results of this study, it appears that evaluative markers can influence the nature of feedback that students give, reminding students that feedback entails more than giving critique.

A significant difference between groups was found for reasoning, with the support group providing more reasoning for positive feedback. Part of the explanation for this probably lies in the varying number of positive comments between groups, as the experimental group did provide more comments in general, and though insignificant, also more positive comments than the control group. As stated, we postulate that the evalua-

tive markers may have been the cause of the higher number of positive comments. Further, it seems logical that the sentence openers have influenced students to also provide reasoning for why something was evaluated positively.

We found no differences for the quality indicators specificity and suggestion. A possible explanation for the comparable specificity of comments is that both groups used means of video-annotations to create comments. Studies have shown that video-annotation allows for more specific feedback, as it relates to specific moments in a video of task-performance (Leung & Shek, 2021; McFadden et al., 2014). As a result, it is not surprising that both groups had relatively high scores for specificity of comments. Students apparently were able to easily relate their feedback to specific moments in the video. The comparable suggestion score might be attributed to the fact that students were not directly prompted to include this in their comments. The markers prompted students to include questions or positive feedback, and the initial prompt "because" would invoke students to write a justification, but not necessarily an elaborate description of the consequences and alternatives. We could very well expect that a more complex prompt could invoke further elaborations and suggestions from students. For example, instead of only a sentence opener, follow-up sentences could be prompted too: "Not good, because…" "This could lead to…/A risk is…" "Next time, you could try to…" or "It would be better if…". The prompt in our study was only focused on the prompted statement 'Not good, because'. We suspect that this might not have been specific enough, and that some students were more inclined to complete that sentence in a way that related to the specific topic. For example, the sentence "Not good, because… you didn't wear gloves." only explains what was wrong, but not why it was wrong, nor did it really explain the risks or alternatives. In this sense, the prompt we introduced might actually have supported specificity or reasoning, depending on how students interpreted it.

Another notable trend to discuss is the acceptance of the feedback by the feedback receiver (Alqassab et al., 2023; Panadero & Alqassab, 2019; Strijbos et al., 2010), which was not a part of the current study. However, we noticed that students sometimes assume some form of shared knowledge that the feedback receiver, as fellow participant in the domain, will understand. Students often try to justify their feedback using sentences such as 'as it should be done', or 'using the correct procedure'. They seem to assume that the feedback receiver will understand what this ideal way of working is. When the researcher asked students what this 'correct way' meant, they were able to explain verbally why something should be done in a certain way, and what the consequences of not doing so are. However, this often is not visible in their comments. We do not expect that feedback receivers will always be able to understand the correct way of working, as sociocultural practices and personal experiences may differ greatly across sites within a domain (e.g. Roth, 2014). Therefore, we would suggest that it might be very beneficial for student learning to be guided in sense making of the various perspectives on practice that simultaneously exist, or that feedback activities are performed between peers who know each other and their work contexts. Of course, this finding can also be caused by the context of this study, in which students do not know the recipient of their feedback, or perhaps also felt pressure to find things to comment on, that they would not have mentioned in a natural setting.

In addition to this argument, we also need to discuss the necessity of reasoning. As not all topics of performance really needed to be justified or elaborated in this study. A commonly commented aspect of the performance was the workplace organisation. Stating that an environment was a mess did not really need to be elaborated on, as it was clearly visible in the video. As we did not provide rubrics to support students to pay attention to specific aspects of performance, sometimes more general comments about work environments were to be expected, as these performance aspects stand out most. Of course, the context of this study could also be of influence here.

5.2 Limitations

A limitation to this study is the combined nature of the support. By combining both the sentence openers and evaluative markers, the effects of the isolated supports remain inconclusive. Further, in this study students were asked to provide feedback on five short instances of diverse task performance. While it is a strength of this study that all students provided feedback on the exact same performance, a more focused approach towards one specific task could allow for a more detailed analysis of peer feedback. Finally, this study is performed in a simulation. The peer feedback providers did not know the actual performers, nor would the feedback reach them. Anonymity is of influence on peer feedback provision, leading to more critical evaluations (Panadero & Alqassab, 2019). We also expect that due to the simulated setting, participants probably paid a little less attention to how the feedback would be received than in a realistic setting. Similarly, the presence of a researcher during the process may have influenced the feedback output somehow, but as this was the case for all participants, the groups remained comparable. In relation to comparable participant groups, the data sample was collected at various VET locations. Because of the randomized approach to assign participants to either condition, the possible bias of how these VET contexts influence the results is ruled out as much as possible.

5.3 Conclusion and Implications for Future Research

This study has shown that it is possible to influence student feedback using various prompts in a VET setting. While at first sight, the study results mostly did not indicate significant results between both groups, we can still conclude that a prompt-based support, combining evaluative markers and sentence openers, can improve peer feedback quality. Providing students in VET, and perhaps also outside of VET, with additional supports like sentence openers and evaluative markers may help them to formulate more effective peer feedback. This is especially true for students who are new to providing peer feedback. Using these supports as scaffolds and examples for types of feedback and ways to formulate feedback can help them understand how to formulate quality feedback. But as has been noted in previous research, providing experienced feedback providers with these supports might greatly reduce their motivation, as they lose some freedom to formulate the feedback in a way they want (Gielen et al., 2010). Thus, we conclude that sentence openers and evaluative markers can both help to support peer feedback quality for VET students.

Future research on supporting feedback provision in VET is needed, as still a lot is unknown. We invite researchers to try and alter the studied supports to fit other contexts and approaches. We postulate that results might have been different if other sentence openers, or even worked out examples of feedback were given. For example, we could have prompted students to explain noticed issues and give suggestions for alternatives. Such additional prompts might provide the required guidance for students to formulate higher-quality feedback. Additionally, instead of only sentence openers, a support could also embed other parts of sentences, or prompt certain sentence structures. Further research on such supportive elements might provide additional and complementary insight on the effects of student feedback quality in digital environments.

6 References

Aamri, K. S. A. (2018). Teacher-Researchers' Reflective Narratives on Their Experiences at the Language Center Research Committee, Sultan Qaboos University, Oman. 10, 15.

Alqassab, M., Strijbos, J.-W., Panadero, E., Ruiz, J. F., Warrens, M., & To, J. (2023). A Systematic Review of Peer Assessment Design Elements. *Educational Psychology Review, 35*(1), 18. https://doi.org/10.1007/s10648-023-09723-7

Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education, 4*. https://www.frontiersin.org/articles/10.3389/feduc.2019.00087

Barkhuizen, G., & Wette, R. (2008). Narrative frames for investigating the experiences of language teachers. *System, 36*(3), 372–387. https://doi.org/10.1016/j.system.2008.02.002

Boldrini, E., & Cattaneo, A. (2014). Scaffolding Collaborative Reflective Writing in a VET Curriculum. *Vocations and Learning, 7*(2), 145–165. https://doi.org/10.1007/s12186-014-9110-3

Cattaneo, & Motta, E. (2020). "I Reflect, Therefore I Am… a Good Professional". On the Relationship between Reflection-on-Action, Reflection-in-Action and Professional Performance in Vocational Education. Vocations and Learning. https://doi.org/10.1007/s12186-020-09259-9

Colasante, M., & Douglas, K. (2016). Prepare-Participate-Connect: Active Learning with Video Annotation. *Australasian Journal of Educational Technology, 32*(4), 68–91.

Evi-Colombo, A., Cattaneo, A., & Bétrancourt, M. (2020). Technical and Pedagogical Affordances of Video Annotation: A Literature Review. *Journal of Educational Multimedia and Hypermedia, 29*(3), 193–226.

Farkas, K. (2019). Reflective templates in language teacher education: A pilot study on Hungarian and Turkish pre-service EFL teachers' beliefs. 12.

Gavota, M. C., Cattaneo, A., Arn, C., Boldrini, E., Motta, E., Schneider, D., & Betrancourt, M. (2010). Computer-supported peer commenting: A promising instructional method to promote skill development in vocational education. *Journal of Vocational Education & Training, 62*(4), 495–511. https://doi.org/10.1080/13636820.2010.529499

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304–315. https://doi.org/10.1016/j.learninstruc.2009.08.007

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education, 36*, 101–111. https://doi.org/10.1016/j.tate.2013.07.008

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

Hulsman, R. L., & van der Vloodt, J. (2015). Self-evaluation and peer-feedback of medical students' communication skills using a web-based video annotation system. Exploring content and specificity. *Patient Education and Counseling, 98*(3), 356–363. https://doi.org/10.1016/j.pec.2014.11.007

Jossberger, H., Brand-Gruwel, S., Boshuizen, H., & van de Wiel, M. (2010). The challenge of self-directed and self-regulated learning in vocational education: A theoretical analysis and synthesis of requirements. *Journal of Vocational Education & Training, 62*(4), 415–440. https://doi.org/10.1080/13636820.2010.523479

Jossberger, H., Brand-Gruwel, S., van de Wiel, M. W. J., & Boshuizen, H. P. A. (2015). Teachers' Perceptions of Teaching in Workplace Simulations in Vocational Education. *Vocations and Learning, 8*(3), 287–318. https://doi.org/10.1007/s12186-015-9137-0

Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction, 20*(4), 344–348. https://doi.org/10.1016/j.learninstruc.2009.08.005

Kori, K., Pedaste, M., Leijen, Ä., & Mäeots, M. (2014). Supporting reflection in technology-enhanced learning. *Educational Research Review, 11*, 45–55. https://doi.org/10.1016/j.edurev.2013.11.003

Lai, C.-Y., Chen, L.-J., Yen, Y.-C., & Lin, K.-Y. (2020). Impact of video annotation on undergraduate nursing students' communication performance and commenting behaviour during an online peer-assessment activity. *Australasian Journal of Educational Technology, 36*(2), Article 2. https://doi.org/10.14742/ajet.4341

Lavoué, É., Molinari, G., Prié, Y., & Khezami, S. (2015). Reflection-in-action markers for reflection-on-action in Computer-Supported Collaborative Learning settings. *Computers & Education, 88*, 129–142. https://doi.org/10.1016/j.compedu.2015.05.001

Leung, K. C., & Shek, M. P. (2021). Adoption of video annotation tool in enhancing students' reflective ability level and communication competence. *Coaching: An International Journal of Theory, Research and Practice, 0*(0), 1–11. https://doi.org/10.1080/17521882.2021.1879187

McFadden, J., Ellis, J., Anwar, T., & Roehrig, G. (2014). Beginning Science Teachers' Use of a Digital Video Annotation Tool to Promote Reflective Practices. *Journal of Science Education and Technology*, 23. https://doi.org/10.1007/s10956-013-9476-2

Mikkonen, S., Pylväs, L., Rintala, H., Nokelainen, P., & Postareff, L. (2017). Guiding workplace learning in vocational education and training: A literature review. *Empirical Research in Vocational Education and Training, 9*(1). https://doi.org/10.1186/s40461-017-0053-4

OECD. (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. OECD. https://doi.org/10.1787/9789264204256-en

Ortoleva, G., & Bétrancourt, M. (2016). Supporting productive collaboration in a computer-supported instructional activity: Peer-feedback on critical incidents in health care education. *Journal of Vocational Education & Training*, *68*(2), 178–197. https://doi.org/10.1080/13636820.2015.1133693

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00422

Panadero, E., & Alqassab, M. (2019). An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading. *Assessment & Evaluation in Higher Education*, *44*(8), 1253–1278. https://doi.org/10.1080/02602938.2019.1600186

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. https://doi.org/10.1016/j.edurev.2013.01.002

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*. https://doi.org/10.1016/j.edurev.2020.100329

Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, *35*, 100416. https://doi.org/10.1016/j.edurev.2021.100416

Prilop, C. N., Weber, K. E., Prins, F. J., & Kleinknecht, M. (2021). Connecting feedback to self-efficacy: Receiving and providing peer feedback in teacher education. *Studies in Educational Evaluation*, *70*. https://doi.org/10.1016/j.stueduc.2021.101062

Roth, W.-M. (2014). Rules of bending, bending the rules: The geometry of electrical conduit bending in college and workplace. *Educational Studies in Mathematics*, *86*(2), 177–192.

Schaap, H., Baartman, L., & de Bruijn, E. (2012). Students' Learning Processes during School-Based Learning and Workplace Learning in Vocational Education: A Review. *Vocations and Learning*, *5*(2), 99–117. https://doi.org/10.1007/s12186-011-9069-2

Schaap, H., Khaled, A., Faber, M., Bijlsma, N., & Bruijn, E. de. (2023). Vocational knowledge development during teacher-student interactions: Exploring vocational conversations. *Pedagogische Studiën*, *100*(1), Article 1. https://doi.org/10.59302/ps.v100i1.13990

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

Strijbos, J.-W., Narciss, S., & Dünnebier, K. (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. https://doi.org/10.1016/j.learninstruc.2009.08.008

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: Research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior*, *21*(4), 671–688. https://doi.org/10.1016/j.chb.2004.10.039

Yiong-Hwee, T., & Churchill, D. (2007). Using Sentence Openers to Support Students' Argumentation in an Online Learning Environment. Educational Media International, 44(3), 207–218. https://doi.org/10.1080/09523980701491658

Acknowledgements

This work was funded by the Dutch Organisation for Scientific Research (NWO).