Feedback is a gift

Do Video-enhanced rubrics result in providing better peer feedback than textual rubrics?

Ackermans, Kevin; Rusman, Ellen; Nadolski, Rob; Brand-Gruwel, Saskia; Specht, Marcus

DOI
10.7275/hk9e-8d82

Publication date
2021

Document Version
Final published version

Published in
Practical Assessment, Research and Evaluation

Practical Assessment, Research, and Evaluation

Volume 26 Article 17

2021

Feedback is a gift: Do Video-enhanced rubrics result in providing better peer feedback than textual rubrics?

Kevin Ackermans
The Open University of the Netherlands

Ellen Rusman
The Open University of the Netherlands

Rob Nadolski
The Open University of the Netherlands

Saskia Brand-Gruwel
Zuyd University of applied science

Marcus Specht
Delft University of Technology

Recommended Citation

Ackermans, Kevin; Rusman, Ellen; Nadolski, Rob; Brand-Gruwel, Saskia; and Specht, Marcus (2021) "Feedback is a gift: Do Video-enhanced rubrics result in providing better peer feedback than textual rubrics?," *Practical Assessment, Research, and Evaluation*: Vol. 26 , Article 17.
DOI: https://doi.org/10.7275/hk9e-8d82
Available at: https://scholarworks.umass.edu/pare/vol26/iss1/17

Feedback is a gift: Do Video-enhanced rubrics result in providing better peer feedback than textual rubrics?

Cover Page Footnote

The Viewbrics-project is funded by the practice-oriented research program of the Netherlands Initiative for Education Research (NRO), part of The Netherlands Organization for Scientific Research (NWO) [grant number: 405-15-550]. The ethics procedure was approved by the ethics committee of the author's institute. Consent was obtained from the participants. Data generated during and/or analyzed during the current study are available in the Open Science Framework.

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 26 Number 17, June 2021 ISSN 1531-7714

Feedback is a gift: Do Video-enhanced rubrics result in providing better peer feedback than textual rubrics?

Kevin Ackermans, *The Open University of the Netherlands*
Ellen Rusman, *The Open University of the Netherlands*
Rob Nadolski, *The Open University of the Netherlands*
Saskia Brand-Gruwel, *Zuyd University of Applied Science*
Marcus Specht, *Delft University of Technology*

High-quality elaborative peer feedback is a blessing for both learners and teachers. However, learners can experience difficulties in giving high-quality feedback on complex skills using textual analytic rubrics. High-quality elaborative feedback can be strengthened by adding video-modeling examples with embedded self-explanation prompts, turning textual analytic rubrics (TR) into so-called 'video-enhanced analytic rubrics' (VER). This study contrasts two experimental conditions (TR, n = 54; VERs, n = 49) with their version of the anonymized online tool (used to collect the given feedback in 'Tips for improvement and Tops identifying strengths'). Peer feedback quality (concreteness and consistency) was evaluated using Natural Language Processing. As expected, the video-enhanced rubrics condition resulted in a higher quantity of words used and a lower amount of naive wording compared to the textual rubric condition. Contrary to our assumptions, it did not lower the amount of non-constructive wording nor improved the amount of behavioral and process-related feedback. Possibly, the transition from providing more feedback to delivering more accurate behavioral and process-related feedback has not yet been made in the time set for the study.

Background

Many secondary education schools struggle to teach and assess 21st-century skills in a methodologic, structured way (Thijs et al., 2014). It can be hard to move to efficient and systematic ways of project education that include teaching complex 21st-century skills. Feedback and textual rubrics are frequently implemented to tackle the problem of formatively assessing 21st-century skills in education (Rusman et al., 2014; Thijs et al., 2014). An added advantage of rubrics is that they appear to facilitate more valid peer feedback (Panadero et al., 2013). When learners develop an accurate and consistent mental model of (aspects of) 21st-century skills, this allows them to provide more valid, elaborate, and high-quality feedback (Shute, 2008). In this study, we investigate if the developed rich mental models (confirmed in our previous work) lead to higher quality feedback.

Connecting Rich Mental Models and Feedback Quality

Our previous study established that learners developed a richer mental model for the complex skills of information literacy and collaboration when using video-enhanced rubrics (VER) than learners using textual rubrics during formative assessment through the Viewbrics online tool (Ackermans, Rusman, Nadolski et al., 2019). A rich mental model is rich in concepts (i.e., it contains a multitude of concepts), has a linear structure (like a Fishbone or cause and effect diagram), contains hierarchies, and a variety of complex relationships (Besterfield-Sacre et al., 2004; Novak, 1985). These research outcomes were in line with Ilgen, Fisher, and Taylor's (1979) work on the stages of feedback. Ilgen and colleagues found that learners' improved mental models benefit their peer and self-feedback quality. They also (Ilgen et al., 1979) discovered that the mental model accuracy of the feedback provider is positively related to feedback acceptance of the receiver. Accepting feedback is essential for feedback to be actionable (Wiggins, 2012). To summarize, we expect a rich mental model to benefit peer and self-feedback quality and the acceptance of feedback by the receiver. To understand this expected effect, we further define the value and importance of high-quality feedback.

The value and importance of high-quality peer feedback

Quality peer feedback is a valuable asset in the teacher's toolbox when supporting students' complex skill development. Implementing peer feedback can save valuable teacher time, and providing peer feedback can be a valuable learning exercise for the feedback-giver as well as the feedback-recipient (Hattie & Timperley, 2007; Shute, 2008). In this paper, we explore whether the quality of peer feedback benefits from using a video-enhanced rubric-format within the Viewbrics online formative assessment methodology and tool when acquiring complex generic skills, instead of a textual rubric format. To explore the difference between using textual rubrics and video-enhanced rubrics (both within online formative assessment methodology and tool), we first need to understand the concept of feedback. Feedback is intended to help learners pinpoint and reflect on (aspects of) their complex skill performance that can be improved. Feedback consists of an assessment (performance-related) with an explanation (content-specific information) (Jaehnig & Miller, 2007). The assessment part of feedback provides insight into the current strengths and weaknesses of the learner. When learners are aware of their strengths and weaknesses, they know how competent they are and how they can still grow (Weaver, 2006). The explanation (content-specific information) part of feedback, therefore, contains content-related solutions and advice for the future, providing insight into achievable goals known as "feed-up" (Hattie & Timperley, 2007). A way of expressing content-specific feedback commonly used in schools is formulating tips (feedback used for improvement) and tops (feedback that identifies strengths). The resulting combination of content-specific and performance-related feedback, known as elaborative feedback, can only be effective if the learner can also process it (Gibbs et al., 2003). A commonly used way of processing feedback in schools is a learner-formulated learning goal. A recent study by Mattheiss, Alexander, & Graves (Mattheiss et al., 2018) suggests that elaborative feedback's enhanced effectiveness may result from activating the reward-related and task-relevant brain regions. Peer feedback is thought to be especially useful in a formative assessment setting. It encourages students to develop a clear concept of complex skills and a sense of ownership for their peers' learning. Ownership for their peer's learning results in critical, independent, accurate, and fair peer-feedback (Gielen, Tops et al., 2010; Hovardas et al., 2014). In terms of acceptance of peer-feedback by the learner, the feedback must be received mindfully (Gielen, Peeters, et al., 2010). Peer feedback has the added quality of increasing the frequency, extent, and speed of feedback for learners while keeping teachers' workload under control (Gielen, Peeters, et al., 2010).

Rubrics

The formative assessment methodology is designed to provide feedback on frequent and ongoing moments in the learning cycle, identify learning needs, and adjust teaching appropriately (Bennett, 2011; Irons, 2007; Reddy & Andrade, 2010). Various formative assessment approaches have been developed, featuring rubrics as a method for scaling progress and providing feedback onto the learner's current skill level and future goal level (feed-up). From a learner's point of view, the transparency of knowing their skill level and goal may aid the feedback process. A rubric allows the learner to review the received feedback and provides or inspires to give (self-and peer) feedback. The transparency of a rubric may also enable low-achieving learners to strategically reach a passing grade by providing valuable insight into the minimum requirements per constituent skill (Mertler, 2001; Panadero & Jonsson, 2013). Analytic textual rubrics mainly contribute to developing complex skills on a cognitive level, providing rich feedback, anxiety-reducing transparency, and performance-enhancing insight into the performance levels of a complex skill (Ackermans et al., 2017). A rubric is an analytical assessment instrument that supports providing feedback on individual task-aspects. A rubric differs from the traditional grading system by providing insight into the process of mastery of skills through clear descriptions of performance levels. This focus on learning makes rubrics a useful tool for formative assessment purposes. Rubrics have been shown to improve scoring performance and self-assessment accuracy (Panadero & Romero, 2014). We implemented expert, peer, and self-assessment via validated rubrics for collaboration, information literacy, and oral presentation. Kerkhoffs, Stark, & Zeelenberg (2006) developed the collaboration and information literacy rubric for the Dutch National Expertise Centre for Curriculum Development (SLO) in the Combo project. The rubrics, within the Viewbrics online formative assessment method and accompanying tool, were partly based on the Combo project rubrics, while they were designed and validated for Dutch pre-university learners. Van Ginkel, Gulikers, Biemans, & Mulder (2015) developed the oral presentation rubric for Dutch higher education. Van Ginkel et al. 's oral presentation rubric was further refined towards pre-university education. Learners, teachers, and researchers were involved in iterative revisions of all three rubrics to ensure a learner-understandable, detailed textual description of four complex skill mastery levels in an ecologically valid rubric (Rusman & Dirkx, 2017). The final ecologically validated versions of the collaboration, oral presentation, and information literacy rubrics were embedded in this study's Viewbrics online tool.

Room for the improvement of rubrics. There is room for improvement when using rubrics for the formative assessment of complex generic skills. A learner can only provide high-quality feedback on complex skills if (s)he knows how the expected performance looks (Scheeler et al., 2004). We found three problems with applying textual rubrics for the formative assessment in our previous work (Ackermans et al., 2017). First, textual rubrics provide a fragmentary textual framework because a rubric describes a complex skill using a subdivided set of constituent (sub)skills identified by experts. Identifying subdivided sets may result in insufficient attention to the necessary integration of constituent skills during task execution (Van Merriënboer & Kester, 2005; Van Merriënboer & Kirschner, 2007). Second, a textual rubric lacks contextual information needed to convey the real-world

attributes and natural context of skills' performance and representation of dynamic information (such as gesturing in the complex skill of presenting) (Matthews et al., 2010). Third, as complex skills contain several constituent (sub)skills, the learner's priority, sequence, and physical performance need to be observed by the learner to supplement the textual assessment criteria with context and dynamic information (Matthews et al., 2010). Many aspects of desired behavior are hard to put into words, such as body posture or voice during a presentation (de Grez et al., 2014; O'Donovan et al., 2004). We assume that video-enhanced rubrics can provide a solution for these three limitations (Ackermans et al., 2018). With a video-enhanced rubric, we combine a textual rubric with video modeling examples, and self-explanation prompts, which support observational learning of the desired behavior via a role model, a previously proven method (Rohbanfard & Proteau, 2013; van Gog & Rummel, 2010).

Video-enhanced rubrics: integrating video modeling examples with rubrics and self-explanation prompts. Video modeling examples show the complex skills in this study with a perceivable level of mastery by professional (peer-aged) actors. We deliver the videos to the learner through means of the Viewbrics online formative assessment tool. Video can support the development of a mental model of learners for complex skills, while a video is remembered better, contains more and different information provides more cues to aid retrieval from long-term memory, attracts more attention of learners, and increases learner engagement (Matthews et al., 2010). The video-enhanced rubric supports explicitly developing the three complex skills in this study: information literacy, collaboration, and presentation (Ackermans, Rusman, Brand-Gruwel, et al., 2019). Research into information literacy by Frerejean, van Strien, Kirschner, and Brand-Gruwel (2016) indicates that video positively supports performance. Research into collaboration by Kim and McDonough (2011) has found similar results, with video supporting collaboration performance. Finally, also research into oral presentation by De Grez, Valcke, and Roozen (2014) corroborates the beneficial effects of video on performance. However, in these previous studies, the specific format of integrated, video-enhanced rubrics was not developed and studied.

Video Enhanced Rubrics in our Anonymized Online Formative Assessment Tool

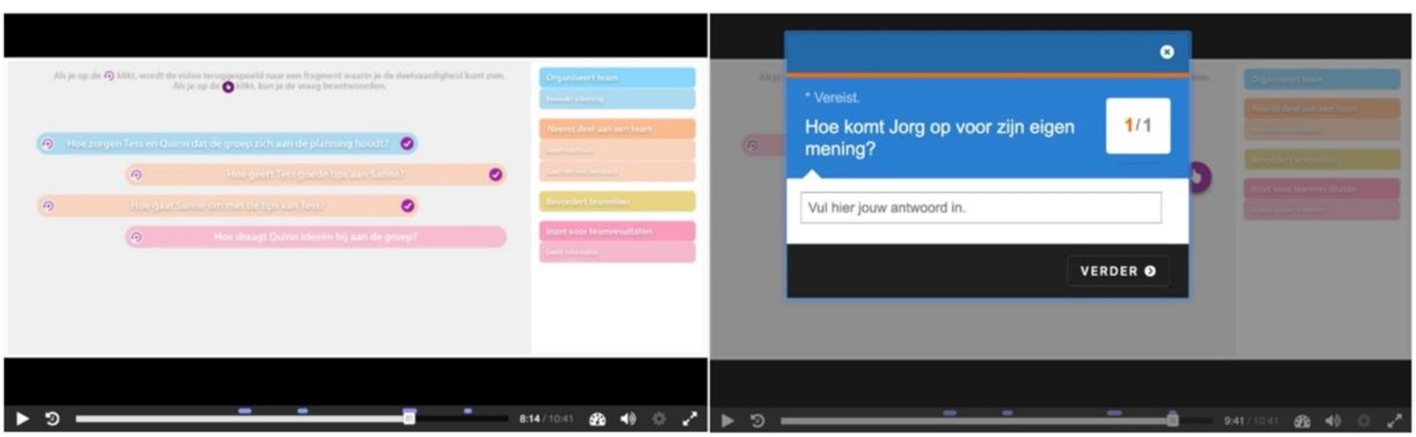
The Viewbrics online tool is a digital 360-degree formative assessment instrument, which embeds a 5-step formative assessment method. Learners themselves, their peers, and their teacher use the Viewbrics online formative assessment tool (Ackermans et al., 2021).

First, it helps learners to get a clearer picture of what it is they are meant to be learning. Learners watch video-enhanced rubrics (VERs) with video-modeling examples and self-explanation prompts in the Viewbrics online tool. The learners watch and process the complete video modeling examples using self-explanation prompts to link the video to the highest performance level description of a sub-skill in the rubrics (as depicted in Figure 1). Learners then proceed to the screen where they can watch the video modeling examples in fragments associated with a sub-skill as defined in the rubrics and review the complete video as illustrated in Figure 2.

Second, learners go 'into the real world' to practice a skill with the impression of skilled behavior they formed by looking at the VER or TR. This (project-or problem-based learning) activity is provided to them by a teacher.

Third, learners self-assess their performance using the rubrics in the Viewbrics online formative assessment tool shown in Figure 3. A skill cluster contains sub-skills that are divided into four performance level descriptors. Only after completing the self-assessment, learners can look at the 360-degree feedback of peers and the teacher (who assess a learners' performance while practicing by scoring the rubrics and providing tips and tops per skills' cluster). This written feedback has the form of tips for improvement and tops identifying strengths.

Figure 1. Self-Reflection Questions and Prompt Received[1]



[1] The image to the left shows the self-reflection questions in order of the appearance of a sub-skill in the video and the color of the corresponding skill-cluster. The image to the right shows the prompt the learner receives when a self-reflection question is selected.

DOI: https://doi.org/10.7275/hk9e-8d82

Figure 2. TR and VER versions of the Viewbrics online tool[2]

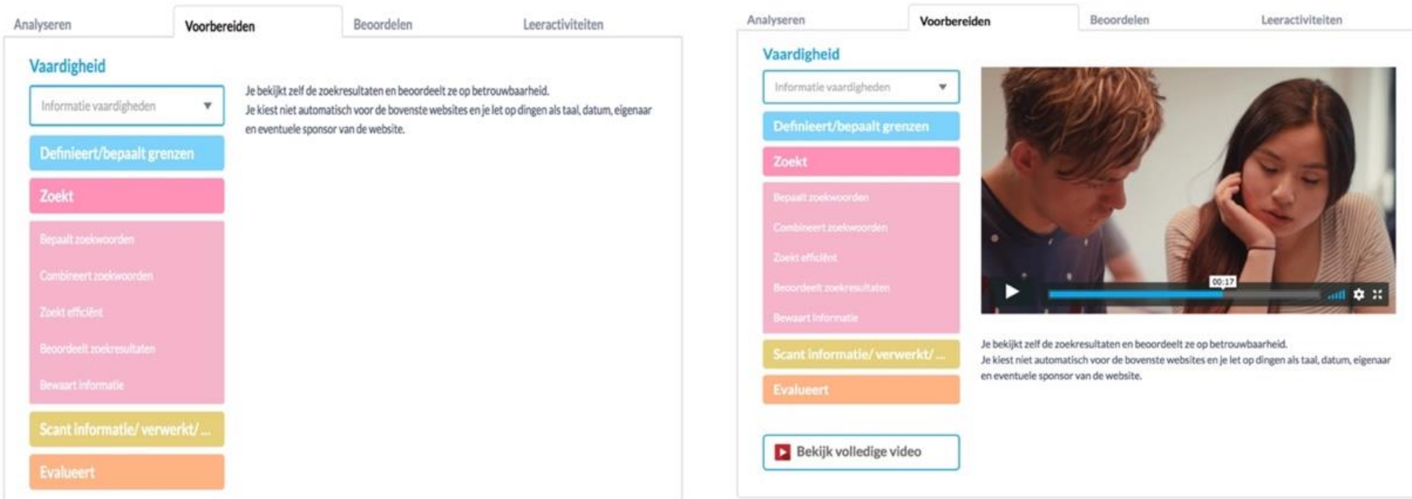
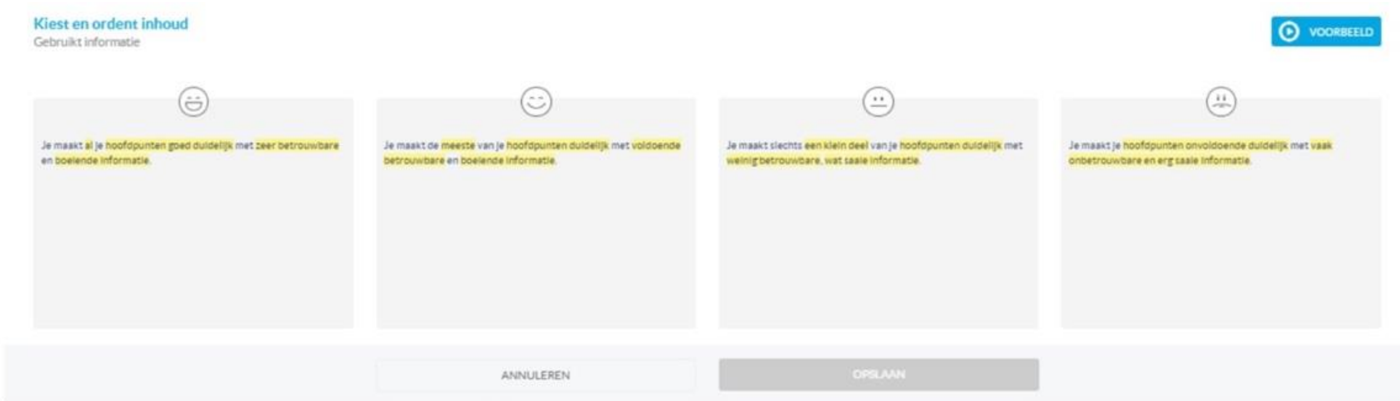


Figure 3. Four performance level descriptors from the left (best) to the right (most possibilities for improvement)



Fourth, a 'skill performance feedback wheel' visualizes teacher and peer assessments. The skill performance feedback wheel represents the learners' performance score on the sub-skills of a complex skill in blue. The visualization allows learners to see at a glance what skills they may still improve and what skills went well. The skill performance feedback wheel shown in figure 4 visualizes growth or shrinkage between assessment moments in performance levels highlights (red for decrease, green for growth). Below the wheel, the top 3 skills that went either well or less well are shown. All tips (for improvement) and tops (identifying strengths) are summarized in a feedback report. During this step, learners analyze this information, determine what went well, and what sub-skills may still need improvement.

Finally, learners describe pursued learning goals in the online formative assessment tool based on their analysis, to support their processing of feedback and determine where to focus on during their next practice session. This learning goal, shown in figure 5, completes the Viewbrics formative assessment report of one specific assessment moment. A distinction is

[2] The image on the left illustrates the TR version of the Viewbrics online tool, whereas the image on the right shows the VER version of the Viewbrics online tool. The skill of information literacy is selected from a drop-down menu, and the skill cluster 'searching' is displaying its five sub-skills (pink). The TR version shows the textual description found in the highest performance level of the textual analytic rubric, whereas the VER version shows the same text supported with a video fragment illustrating the appropriate sub-skill.

Figure 4. The Full Viewbrics Formative Assessment Report[3]

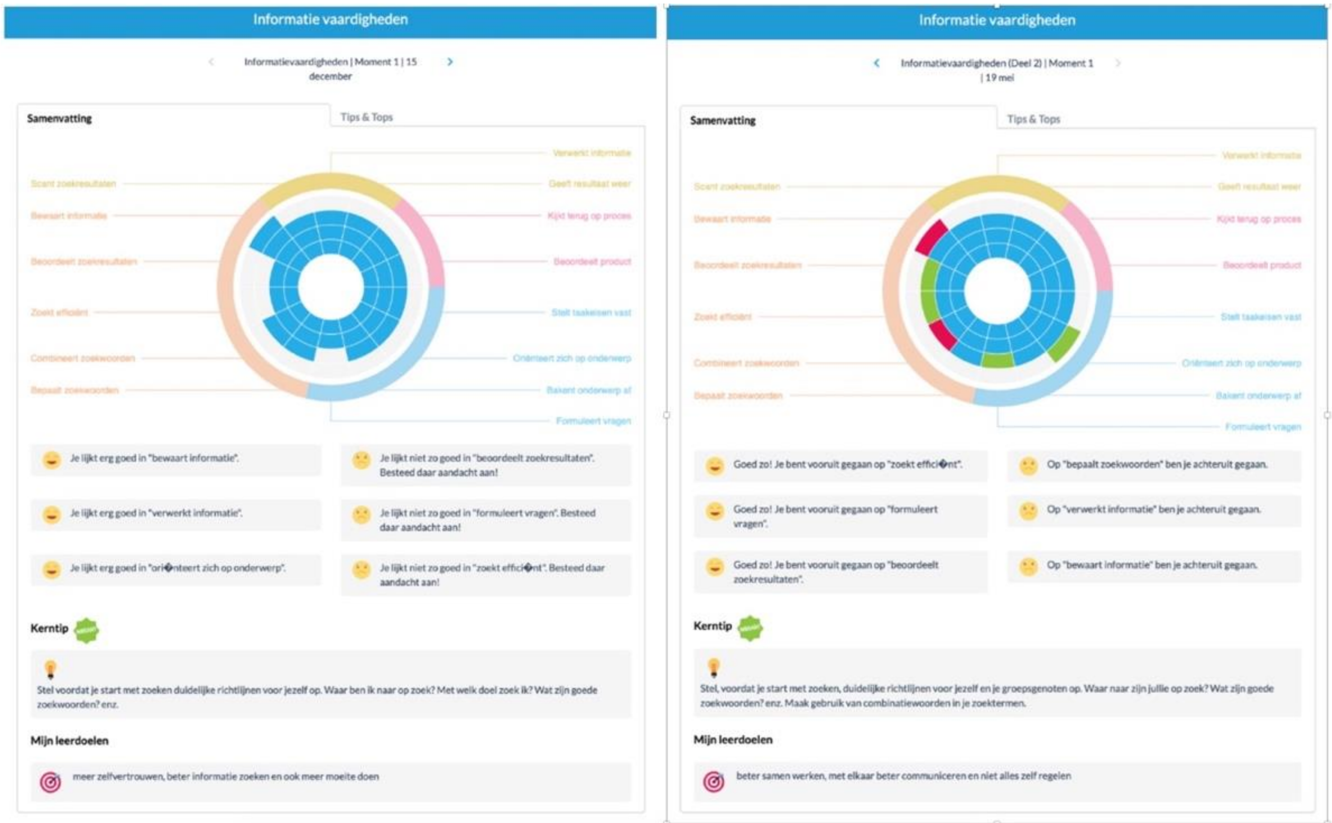
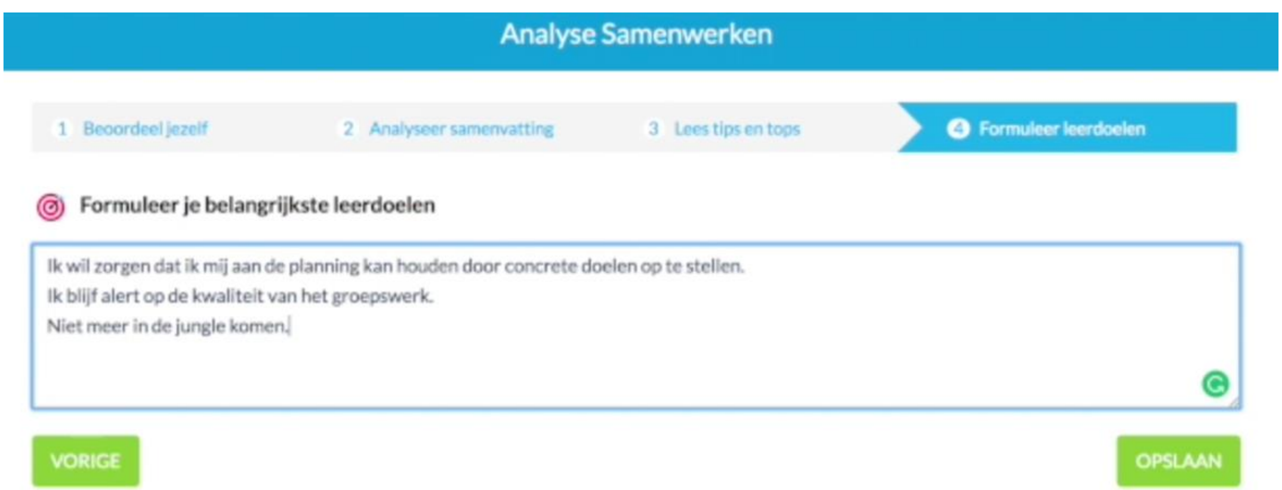


Figure 5. To-be Pursued Learning Goals, Described by the Learner



[3] The image on the left shows a dashboard of a students' first performance of information literacy in blue segments. The image on the right shows this same students' progress (the second performance compared to the first performance). The second performance shows added green (improvement) and red (deterioration) segments and the learners' goal after reflection (depicted as a bulls-eye). This subskill-specific improvement or deterioration is also used as a basis for automatically generated feedback next to the emoticons.

made for learners using the textual rubric version of the Viewbrics online formative assessment tool. The Viewbrics online tool in the textual version contains no videos but is otherwise identical.

A general walkthrough of the Viewbrics online formative assessment tool and a learner's self-assessment (textual rubric version) process is available at this Vimeo link, the learner interface (video-enhanced rubric) with peer feedback is available at this Vimeo link, and a general walkthrough of creating an activity and expert assessing a learner in the teacher interface is available at this Vimeo link. A prototype of the Viewbrics online tool was usability-tested with learners and teachers and found handy, usable, helpful, and feasible for learning complex skills (E. Rusman et al., 2019). A distinction is made for learners using the textual rubric version of the Viewbrics online tool. The Viewbrics online tool in the textual version contains no videos but is otherwise identical.

Defining Feedback Quality

For this study, we define the concept of elaborative feedback quality as feedback concreteness (the feedback is actionable) and consistency (the feedback is stable, accurate, and trustworthy) (Wiggins, 2012).

Concreteness is measured through a combination of four indicators we can assume influence the actionability of feedback as a whole. The "tips and tops" peer feedback is gathered and analyzed using Natural Language Processing to quantify the following four variables.

As the first indicator for the actionability of feedback, we assume a higher number of words per tips and tops is better. This assumption does not guarantee that the feedback provided is more useful, according to Newton, Wallace, and McKimm (2012). However, more specific feedback, related to the performance feedback indicators as defined in the rubrics' descriptions gives learners more feedback to act upon (Liu & Carless, 2006). More words are also a measure of feedback complexity (Schrire, 2006). The remaining variables (amount of non-constructive feedback, amount of non-specific wording, amount of behavioral and process-related feedback, and consistency between teacher and peers feedback) will address evaluating the quality aspects of actionable feedback. Therefore, the number of words used by peers for tips and tops is measured. The length of feedback indicates its actionability and complexity, but even if feedback is verbose, it also needs to be constructive to be of value.

The second indicator is the amount of non-constructive feedback used by the peers. The amount of non-constructive feedback must be minimized to ensure acceptance of the learner's feedback (Ilgen et al., 1979; Newton et al., 2012). This is in line with the feedback system developed by Newton et al. (2012) and Brown and Glover (2006). Newton et al.'s (2012) feedback system ranged from non-constructive (descriptive) to constructive and notes that the differences in constructiveness are highly relevant to achieving quality feedback. A lower percentage of non-constructive feedback is better.

The third indicator is the amount of non-specific wording used by the peers. The amount of non-specific wording should be kept to a minimum (such as good, fun, fine, bad, better) (Gigante et al., 2011). A lower percentage of non-specific wording is better.

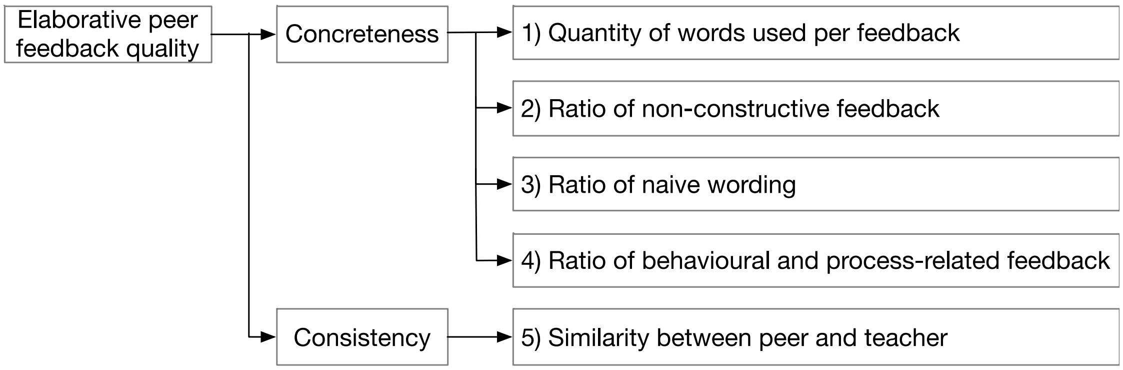
The fourth indicator is the peers' amount of behavioral and process-related feedback. The amount of behavioral and process-related feedback should be high (van der Pol et al., 2008). Behavioral feedback regards how well a task is being accomplished. In contrast, process-related feedback considers the (affective and cognitive) understanding of the gap between what is understood and what is aimed to be understood (Hattie & Timperley, 2007). A higher amount of behavioral and process-related feedback is better.

Consistency is recognized as when teachers and peers are on the same page about quality feedback (Wiggins, 2012). We measure consistency as the similarity between the feedback given by teachers and peers (Stone & Stone, 1985). Figure 6 visualizes how we define feedback quality in this study and which variables were used to determine feedback quality.

Present Study

In our previous work, we established that learners using video-enhanced rubrics through the Viewbrics online formative assessment tool develop a richer mental model for the complex skills of information literacy and collaboration (Ackermans, Rusman, Nadolski, et al., 2019). This study uses the Viewbrics

Figure 6. The definition and variables used to determine feedback quality in this study



online tool to explore the difference between the learners' peer feedback quality using video-enhanced rubrics and learners using textual rubrics. This led to the following research question:

> *Do video-enhanced rubrics, opposed to textual rubrics, applied within the same (online) formative assessment methodology, improve the quality of peer feedback on the performance of a complex skill among learners in secondary education?*

Using our definition of feedback quality (see figure 6), we derived the following five hypotheses for each of the three complex skills. Learners in the VER condition provide more tips and tops than learners in the TR condition (H1). In the VER condition, learners use less non-constructive feedback than in the TR condition (H2). The VER condition learners use less non-specific wording than the learners in the TR condition (H3). The VER condition allows learners to use more behavioral and process-related feedback than learners in the TR condition (H4). The VER condition's peer feedback is more similar to teacher feedback than the similarity between teacher feedback and peers in the TR condition (H5).

Design

The study was a two condition, between-subjects design and evaluated the effect on learners' feedback (quality and quantity) between the TR and the VER condition. The VER and TR conditions used their specific version of the Viewbrics online tool that supports the Viewbrics online formative assessment methodology. One class per school worked within the VER condition (n=49), and one class per school worked within the TR condition (n=54).

Method

Participants

The learners (n = 103) were a convenience sample of four existing bilingual 1st-year classes from two Dutch schools for higher general secondary and pre-university (gymnasium) education (53 female, 50 males; M = 12.48 years, SD = 0.54; range: 12-13 years). All learners were native Dutch speakers, following a bilingual English curriculum. All classes were made up of learners who received a combined general higher secondary/pre-university advice when finishing primary education. School 1 and school 2 selected a convenience sample of two existing classes of bilingual Higher General Secondary Education/pre-university education (HAVO-VWO) for the TR condition and two existing classes of bilingual Higher General Secondary Education/pre-university education (HAVO-VWO) for the VER condition. School 1 has 1331 learners, is located in a municipality with a population of 86.915 inhabitants. School 2 has 1109 learners, is located in a municipality with a population of 122.397 inhabitants.

Setting

The standard curriculum of both schools offered 3 hours of project-based education per week for 24 weeks. School 1 provided the projects on the subject of Humanity and Nature (Mens en Natuur). School 2 provided the projects in the subject of Scientific training and formation. These subjects were chosen by the schools to accommodate project-based education. The projects contained outdoor activities, arts and crafts assignments. Their teacher formatively assessed the complex skills of information literacy, collaboration, and oral presentation during these projects. Both conditions continued this standard curriculum but used the Viewbrics online formative assessment tool instead of the standard curriculum's formative assessment.

Materials

The study materials comprised the Viewbrics online tool, a condition-specific introductory workshop for teachers, and an introductory workshop for learners.

Introductory workshop for teachers. In preparation for the initial workshops, teachers were asked to describe their school curriculum in detail. On a school level, this description included the didactical vision. On a project level, this had work forms, roles a learner is asked to fulfill, the composition of groups and how long these are maintained, the didactical instruments used in the project, the learning environment the learner is limited to, the organization, the learner, is limited to (amount of time available to the learner, scheduled frontal instruction), the amount, type and frequency of guidance and feedback provided by the teacher, and the way formative and summative assessment are currently organized. We took this description as input for the third part of the workshop.

The teacher workshops consist of four parts during one day. The activities in this workshop were formulated actively; promoting discussion, collaboration, and facilitating practice. First, a PowerPoint presentation was used to explain the Viewbrics formative assessment methodology. Second, teachers practiced using the specific version (TR or VER) of the Viewbrics online tool using dummy accounts under included guidance from the research team on the learners' enrollment into the Viewbrics online tool. The third part of the workshop used the preparatory descriptions to discuss which practical constraints to maintain during their project-based education. We followed the discussed rules to synchronize both schools. For instance: all conditions should be as equal as possible across schools; the creation of subgroups for collaborative assignments (comparable in size and gender distribution) and week planning (the number of weeks the learners used the Viewbrics online tool between assessments is kept equal, also taking vacations and exams into account). Peer assessment processes and peer groups' formation were kept equal across classes and schools but were slightly different for the three complex skills. For example, all peers within a class provided feedback on the given presentation. Each group member received feedback from their three peers for collaboration; each learner received feedback from one peer for their information literacy (peer duos). Teachers also assessed learners' performance by giving feedback. During the final part of this workshop, teachers received all necessary informed consent forms for recruiting participants. The ethics committee of the authors' institution approved the informed consent procedure and forms. All participants consented.

Introductory workshop for learners. Learners in TR and VER conditions received an introductory workshop lasting two hours from project members of the Viewbrics project. This workshop was given four times (once for every participating condition, per school) with the teachers' support. During this workshop, learners received a PowerPoint presentation, a video walkthrough, and feedback instructions. The PowerPoint presentation explained the Viewbrics formative assessment methodology and the steps in the formative assessment cycle. Learners were then shown a video walkthrough, practically going through the formative assessment cycle steps using the Viewbrics online tool (condition-specific). Finally, learners received instruction on quality criteria for peer feedback. For peer feedback, the learner received specific tips on formulating feedback, such as defining the behavior that leads to observation, *"I see you are shaking; this makes you seem anxious"* instead of merely making an observation *"you looked nervous."* We also teach the learner to give specific feedback, such as *"I found your presentation difficult to follow because it lacked an index"* instead of *"your presentation was vague."* Learners did not practice with the Viewbrics online tool or giving feedback during the workshop.

Procedure

Both conditions started up their projects according to their school's standard curriculum. All teachers wrote a similar accompanying letter that provided context to the informed consent forms. Parents and learners of all four chosen classes received their school-specific letter and informed consent forms and were asked to return signed consent forms. Then, the workshops were organized for learners and teachers. With the start of the projects, the Viewbrics online tool guided the learner through the Viewbrics methodology's formative assessment and reflection cycle. For the complex skill of oral presentation; Each learner gave feedback on the presentation of all other classmates twice, received feedback from every classmate twice, completed a self-assessment twice (including formulating learning goals), and received expert feedback twice. For the complex skill of collaboration; Each learner gave three teammates feedback (classmates they formed a team with) twice, received feedback from three teammates twice, completed a self-assessment twice (including formulating learning goals,) and received expert feedback twice. For the complex skill of information literacy; Each learner gave feedback to one teammate (a classmate they formed a team with) twice, received feedback from one teammate twice, and completed a self-assessment twice (including formulating learning goals) received expert feedback twice. The participants invested 14 out of the 24 weeks on each project. 6 of the 24 weeks accommodate vacation, exams, and staff meetings. Four weeks were required to complete all oral presentations.

Analysis

Our research question is concerned with the difference in feedback quality between video-enhanced rubrics and textual rubrics applied within the same (online) formative assessment methodology. Peers and teachers write feedback in the form of tips and tops to the learner. Our five hypotheses focus on the quantity of feedback (H1), the quantity of constructive feedback (H2), the quantity of non-specific wording (H3), and used behavioral and process-related feedback (H4) and Consistency between peer and teacher feedback (H5). The analysis used for these hypotheses is detailed in the following paragraphs. The coded data for hypotheses one, two, three, four, and six four were analyzed quantitively using tidytext in RStudio Version 1.3.1073 (Fay, 2018). Tidytext output was then hypothesis tested using Bayesian Paired Samples T-Testing using the BayesFactor package in JASP version 0.14 (Morey et al., 2015; Rouder et al., 2009). The choice for Bayesian Paired Samples T-Testing reduces exaggeration of the strength of a significant effect or p-value (Kubsch et al., 2021). Research spanning 855 t-tests by Wetzels et al. (2011) found Bayes factors and p-values often disagree on the strength of the effect. 70% of p-values in the .01-.05 interval yield evidence that is only “anecdotal”. The coded data for hypothesis five was analyzed using word frequency analysis in NVivo version 12.4.0 (QSR International Pty Ltd, 2015).

Feedback Quantity

Quantity is operationalized as count data of the number of words used for tips and tops (H1). The tips and tops are logged in a secure SQL database for this analysis and exported to Microsoft Excel (version 16.26). For this measurement, a formula counts the average amount of words in tips and tops and the total amount of tips and tops given per condition (while removing extra spaces and empty cells). Information literacy (where one peer gave feedback to the learner) yielded the least amount of feedback—followed by collaboration and presentation. Three or more learners were asked to provide feedback to the learner. The TR condition entered peer-feedback consisting of tips and tops 1745 times. Of these 1745 entries, 154 entries regard information literacy (containing a total of 1505 words), 1078 entries regarding presentation (containing a total of 6305 words), and 512 entries consider collaboration (containing a total of 4960 words). The VER condition entered peer-feedback consisting of tips and tops 1035 times. Of these 1035 entries, 166 entries regard information literacy (containing a total of 1670 words), 381 entries regard presentation (containing a total of 2664 words), and 487 entries consider collaboration (containing a total of 4532 words).

Constructive Feedback

The quantity of constructive feedback given by peers was analyzed using two validated instruments (H2). First, the amount of non-constructive feedback is counted using Newton et al.'s (2012) feedback system. Second, constructive feedback is also identified by asking Wiggins's (2012) question: "What specifically should I do more or less of next time, based on this information?" (p. 14). Example from our data includes: "you've *got cool shoes*," "*I don't know why, but I want to write something*," "*you should wear your hair in a bun*" or "*nice chicken-dance.*" The amount of non-constructive feedback is calculated as a percentage of the total amount of tips and tops per condition.

Non-specific Wording

Quality was thirdly operationalized as the extent to which non-specific wording was used in peer feedback utterances (H3). For this measure, naïve words such as fun, good, fine, bad, and better were counted. The amount of non-specific wording is calculated as a percentage of the total amount of words within the tips and tops per condition.

Behavioral and Process-Related Feedback

Quality was operationalized via rule-based Natural Language Processing of behavioral and process-related feedback in the feedback utterances (H4). To measure behavioral and process-related feedback, we used the validated rubrics detailed in the background section. The current rubric's behavioral and procedural indicators (that allow an assessor to differentiate the four performance levels) are highlighted for the assessor's convenience. The stemmed rubric's highlights found at this Open Science Framework link serve as a corpus for rule-based Natural Language Processing (Hirschberg & Manning, 2015). A case-insensitive formula uses our corpus to cross-reference our tips and tops database. The amount of behavioral and process-related feedback is presented as a percentage of the total amount of tips and tops per condition.

Feedback Consistency

Consistency between peer and teacher feedback was measured through a word frequency analysis in NVivo (version 12.4.0). NVivo listed the top 20 words used in teacher and peer feedback for comparison. The resulting lists were used to determine if there is overlap between the top 20 words used in teacher feedback and the top 20 words used in peer feedback.

Results

The difference between the TR and VER conditions is tested with a Bayesian Paired Samples T-Test using the BayesFactor package in JASP version 0.14 (Morey et al., 2015; Rouder et al., 2009). All our hypotheses follow our argued assumption that the VER condition results in higher feedback quality than the TR condition. Bayes factor hypothesis decisions are expressed in a range from anecdotal to extreme support for a hypothesis. Bayes factor hypothesis classifications are 'anecdotal', 'weak', 'moderate', 'strong', 'very strong', and 'extreme' (Jarosz & Wiley, 2014).

Feedback Quantity

The average learner in the TR condition wrote 4.14 words of feedback per top and 3.19 words of feedback per tip. The average learner in the VERs condition wrote 4.4 words of feedback per top and 4.17 feedback per tip. We found the evidence for our hypothesis (H1) to be extreme for tips and tops (in Tables 1, 2, and figure 7). The VERs condition used significantly more words per tip and top to provide peer feedback compared to the TR condition.

Constructive Feedback

Learners in the TR condition provided non-constructive feedback on 193 instances. The total amount of constructive and non-constructive feedback consists of 11% non-constructive remarks regarding information literacy, 8.53% non-constructive remarks

Table 1. Descriptive statistics of Feedback Quantity

					95% Credible Interval	
	N	Mean	SD	SE	Lower	Upper
Quantity-tip_VER	1034	4.170	5.539	0.172	3.832	4.508
Quantity-tip_TR	1744	3.189	4.702	0.113	2.968	3.410
Quantity-top_VER	1034	4.404	4.200	0.131	4.148	4.661
Quantity-top_TR	1744	4.135	4.268	0.102	3.935	4.336

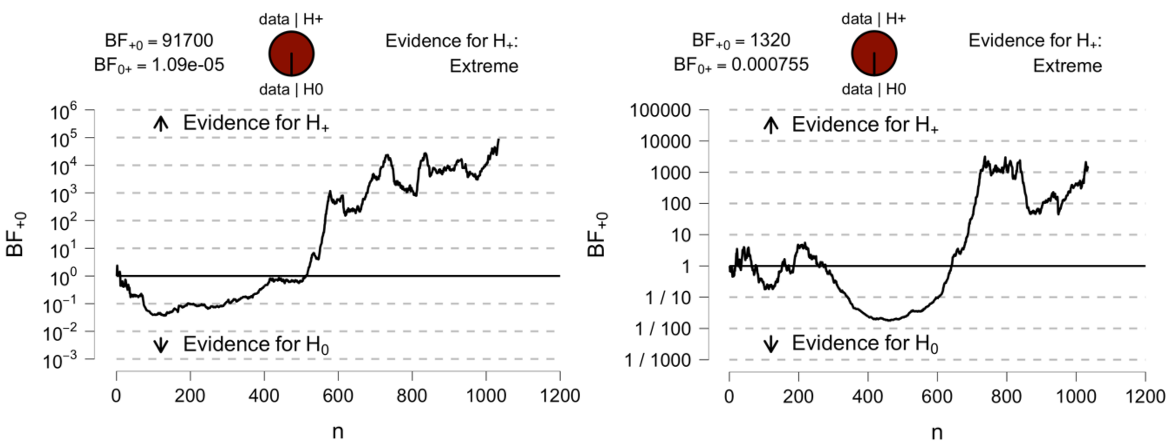
Table 2. Bayesian Paired Samples T-Test of Feedback Quantity

Measure 1		Measure 2	$Log(BF_{+0})$	error %	
Quantity-tip_VER	-	Quantity-tip_TR	11.426	NaN	a
Quantity-top_VER	-	Quantity-top_TR	7.188	~ 2.454e -6	

Note. For all tests, the hypothesis specifies that Measure 1 is greater than Measure 2. For example, Quantity-tip_VER is greater than Quantity-tip_TR.

[a] t-value is large. A Savage-Dickey approximation was used to compute the Bayes factor but no error estimate can be given.

Figure 7. Sequential Analysis: Quantity-tip_VER - Quantity-tip_TR (on the left) and Quantity-top_VER - Quantity-top_TR (on the right).



regarding presentation, and 16.4% non-constructive remarks regarding collaboration. Learners in the VERs condition provided less non-constructive feedback in 50 instances. The amount of constructive and non-constructive feedback utterances consists of 10.3% non-constructive remarks regarding collaboration and no non-constructive remarks for information literacy or oral presentation. We found the evidence against our hypothesis (H2) to vary from extreme (regarding tops) to anecdotal (regarding tips) (in tables 3, 4, and figure 8).

Non-specific Wording

Learners in the TR condition provided non-specific wording on 900 instances. 900 instances of the following five examples of non-specific wording were found: good (5.22%), fun (1.31%), fine (0.04%), bad (0.03%), better (0.45%). Learners in the VER condition provided less naïve wording when compared to the TR condition, on 608 instances. The 608 instances of non-specific consists out of good (7,52%), fun (1,88%), fine (0,06%), bad (0,05%) and better (0,64%). We found the evidence for our hypothesis (H3) to be anecdotal (in table 5, 6, and figure 9).

Behavioral and Process-Related Feedback

Learners in the TR condition provided behavioral and process-related feedback on 2210 instances. Behavioral and process-related feedback makes up 22% of the feedback in information literacy, 15.4% of the feedback in presentation, and 18.1% of the feedback in collaboration. Learners in the VER condition provided behavioral and process-related feedback on 1837 instances. Behavioral and process-related feedback makes up for 27.2% of the feedback in information literacy, 16.9% of the presentation feedback, and 20,6% of the feedback in collaboration. We found the evidence against our hypothesis (H4) to be strong (in table 7, 8, and figure 10).

Feedback Consistency

The transcripts of the Tip and Tops provided by teachers and students were uploaded to NVivo for a word frequency analysis. NVivo omitted common words, presented raw count data, and calculated a weighted percentage for each of the 20 most used words. The 20 most used words make up between 14% and 18% of Tips and Tops. The generally most used word was "information" (varying from 1.8 to 2.1% of

Table 3. Descriptive Statistics of Constructive Feedback

					95% Credible Interval	
	N	Mean	SD	SE	Lower	Upper
non-constructive-tip_TR	1744	0.084	0.401	0.010	0.065	0.103
non-constructive-tip_VER	1034	1.826	17.488	0.544	0.759	2.893
non-constructive-top_TR	1744	0.138	0.506	0.012	0.114	0.161
non-constructive-top_VER	1034	0.044	0.295	0.009	0.026	0.062

Table 4. Bayesian Paired Samples T-Test of Constructive Feedback

Measure 1		Measure 2	$Log(BF_{+0})$	error %
non-constructive-tip_TR	-	non-constructive-tip_VER	-4.844	~ 0.888
non-constructive-top_TR	-	non-constructive-top_VER	-1.065	~ 0.002

Note. For all tests, the hypothesis specifies that Measure 1 is greater than Measure 2. For example, non-constructive-tip_TR is greater than non-constructive-tip_VER.

Figure 8. Sequential Analysis: non-constructive-top_TR - non-constructive-top_VER (on the left) and non-constructive-tip_TR - non-constructive-tip_VER (on the right).

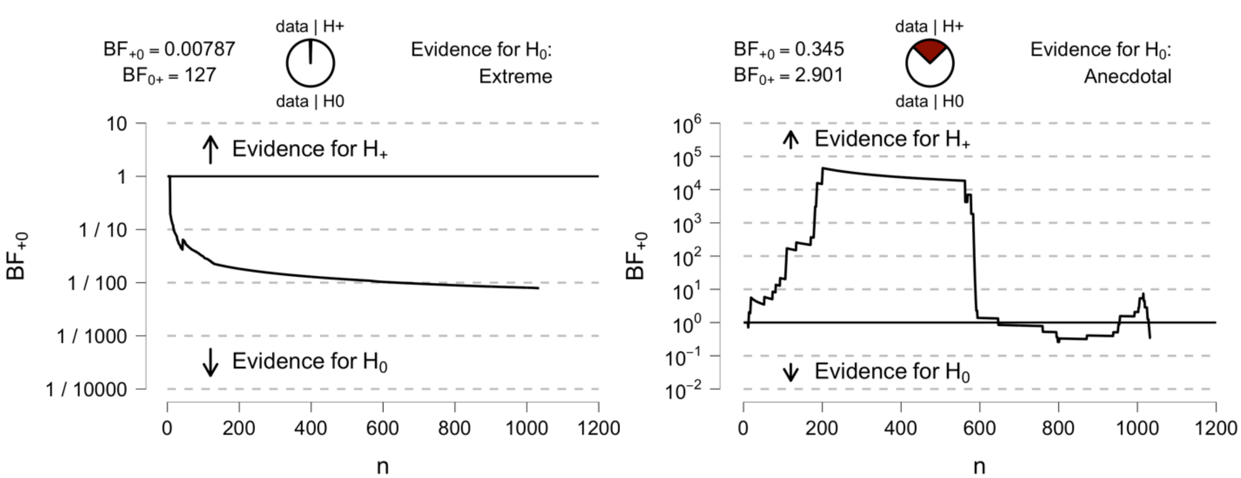


Table 5. Descriptive statistics of Non-Specific Wording

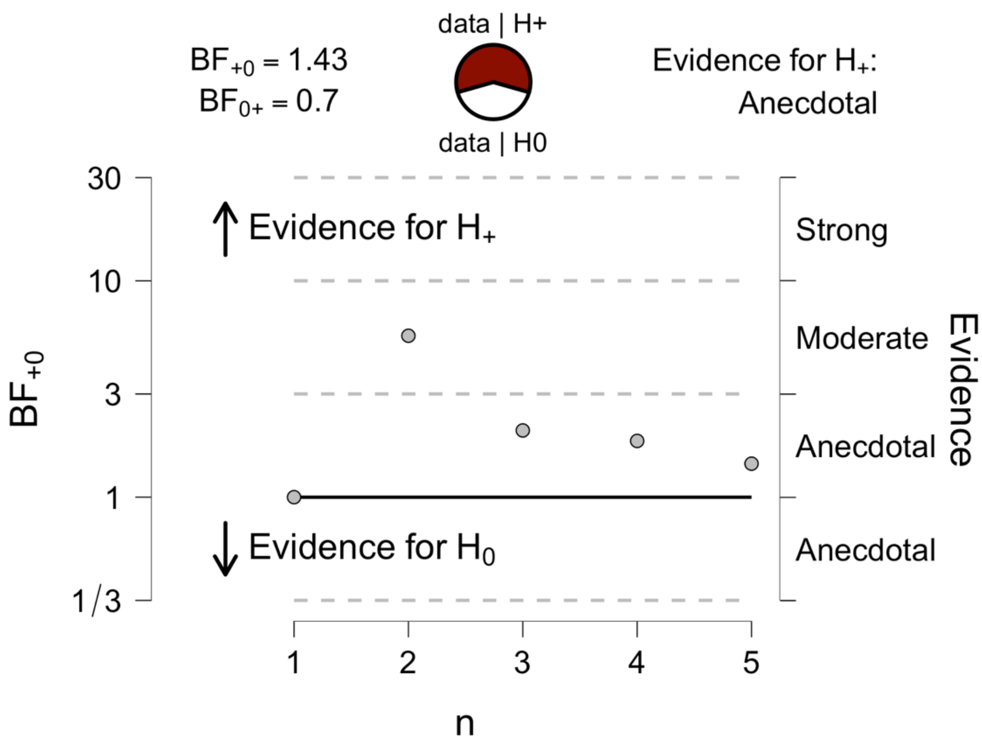
	N	Mean	SD	SE	95% Credible Interval Lower	95% Credible Interval Upper
Naive-wording_TR	5	180.000	280.209	125.313	-167.925	527.925
Naive-wording_VER	5	121.600	224.970	100.609	-157.737	400.937

Table 6. Bayesian Paired Samples T-Test of Non-Specific Wording

Measure 1		Measure 2	BF_{+0}	error %
Naive-wording_TR	-	Naive-wording_VER	1.428	~ 4.739e -5

Note. For all tests, the hypothesis specifies that Naive-wording_TR is greater than Naive-wording_VER.

Figure 9. Sequential Analysis: Naive-wording_TR - Naive-wording_VER



the text). For learner feedback, the 2nd most used word was "audience" (1.2-2.0%). For teacher feedback 2nd most used word was "group" (1.3-1.5%). For the TR condition, 8 out of the top 20 most used words in the teacher and learner feedback are equal. For the VER condition, 9 out of the top 20 words used in the TR condition's teacher and learner feedback are similar. There is no significant difference in feedback consistency between both conditions (H5).

Table 7. Descriptive statistics of Behavioral and Process-Related Feedback

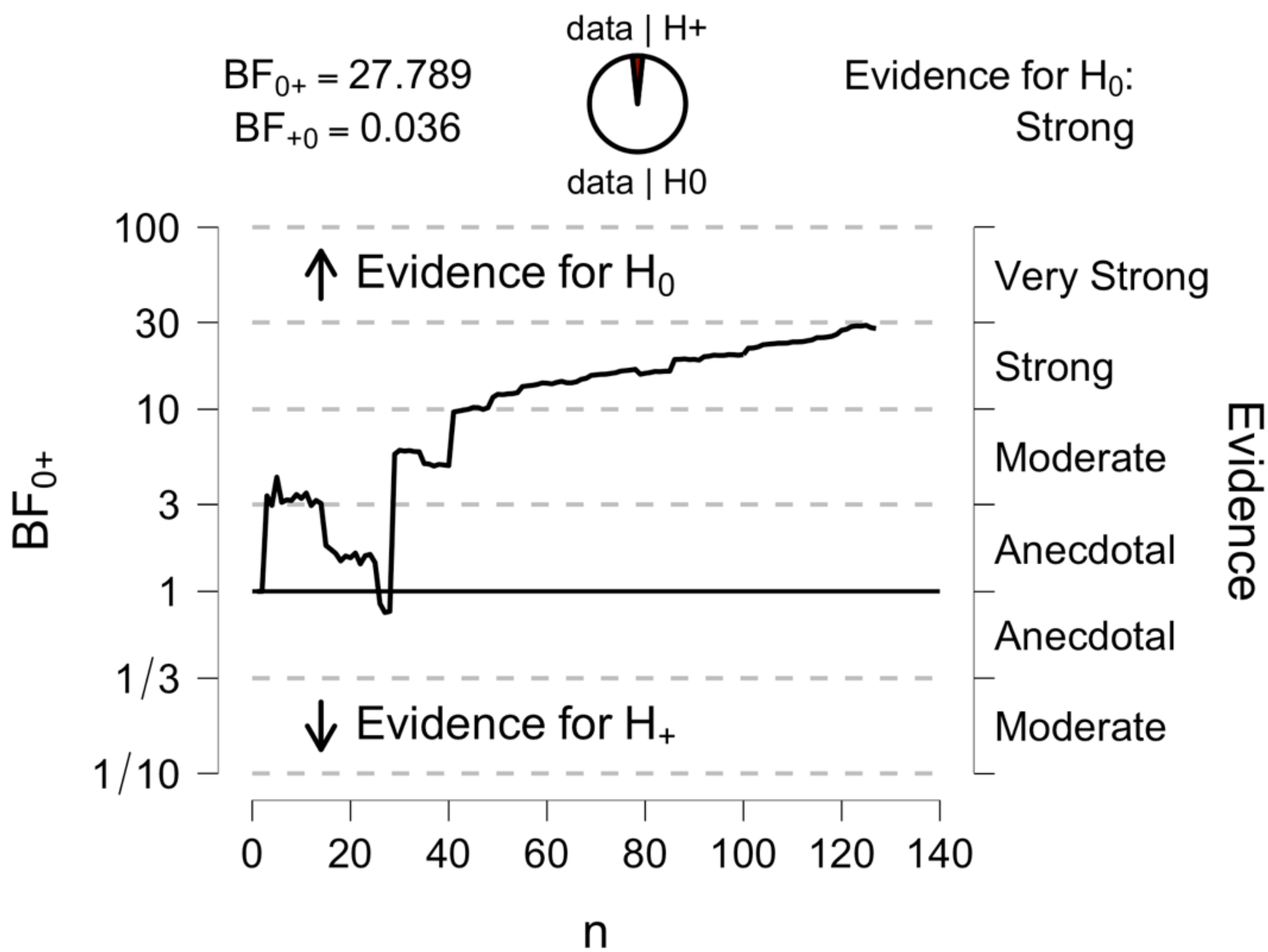
					95% Credible Interval	
	N	Mean	SD	SE	Lower	Upper
Process-related_VER	127	14.465	48.198	4.277	6.001	22.928
Process-related_TR	127	17.402	62.430	5.540	6.439	28.365

Table 8. Bayesian Paired Samples T-Test of Behavioral and Process-Related Feedback

Measure 1		Measure 2	BF_{0+}	error %
Process-related_VER	-	Process-related_TR	27.789	~ 0.185

Note. For all tests, the hypothesis specifies that Process-related_VER is greater than Process-related_TR.

Figure 10. Sequential Analysis: Process-related_VER - Process-related_TR



the text). For learner feedback, the 2nd most used word was "audience" (1.2-2.0%). For teacher feedback 2nd most used word was "group" (1.3-1.5%). For the TR condition, 8 out of the top 20 most used words in the teacher and learner feedback are equal. For the VER condition, 9 out of the top 20 words used in the TR condition's teacher and learner feedback are similar. There is no significant difference in feedback consistency between both conditions (H5).

Conclusion/Discussion

This study investigated whether using a video-enhanced rubric (instead of a textual analytic rubric) within an online formative assessment methodology for learning complex skills that foster a rich mental model resulted in higher peer feedback quality. This study contributed to the scientific field of feedback by further investigating the interaction between rich mental models and feedback quality. Shute (2008) argued that a consistent mental model of (aspects of) 21st-century skills allows learners to provide more valid, elaborate, and high-quality feedback. We found implementing video enhanced rubrics to firstly increase the number of words learners use in feedback, and secondly, it led to less naïve feedback.

Peer feedback quality was evaluated using the tips (for improvement) and tops (identifying strengths). Learners gave each other tips and tops as part of their elaborated peer feedback utterances using a formative assessment methodology (supported via the Viewbrics online tool). The video-enhanced rubrics version of the Viewbrics online tool resulted in peers using a higher number of words (H1). A positive effect of increased feedback quantity is discussed by Ruegg (2014). Ruegg (2014) links increased feedback quantity to increased feedback accuracy. Learners in de video-enhanced rubric condition also lowered the amount of naïve wording (H3). This means learners used less naïve wording for measuring feedback such as "good", "better", "fine", "bad" or "worse".

However, we did not find the hypothesized lower amount of non-constructive feedback (H2), higher amount of process-related feedback (H4), and higher feedback consistency (H5), as expected beforehand. Possibly, the transition from providing more feedback to delivering more accurate behavioral and process-related feedback has not yet been made in the time set for the study (Ruegg, 2014). Another possible explanation is that peer assessors need specific training and scaffolding to produce high-quality process-related feedback (Hovardas et al., 2014). We did not provide the needed scaffolding in the student workshop, which could explain why we did not find high-quality process-related feedback. Our results do not reflect the outcomes of a recent study by Hovardas, Tsivitanidou, & Zacharia (Hovardas et al., 2014), who found a high consistency between written feedback from student and expert assessors.

We conclude that the VERs group produces more (and less naïve) feedback; this higher quantity feedback meets elements of Wiggens' (2012) seven keys to effective feedback. This finding is an exciting confirmation of the interdependencies between a richer mental model and (aspects of) feedback quality, as stated by Gary & Wood (2011). For educational practice, we can confirm the value of fostering a learner's mental model for feedback quality, even though such a mental model's repercussions are not fully covered under our definition of feedback quality.

Limitations and Future Studies

The total number of tips and tops given by the VERs condition is lower than the TR condition (H1). Likely, the VERs condition learners used the "I have no tip or top" button more than the TR condition learners. Although the option "I have no tip/top" did detract from the total amount of data, it stands to reason that not having the "I have no tip/top" can lead to more non-constructive or naïve feedback.

This study differentiates between the quality of feedback relatively. This does not allow us to conclude the objective state of the quality of teacher and learner feedback in general. Therefore, it is entirely possible that teachers and learners were not significantly different (H5) because both teachers and learners were either very good or very bad at giving feedback.

The crossroads between Feedback quality and learner's Mental Models are mainly examined in the field of business, where strategic problems are investigated through managers' mental decision-making process (Capelo & Dias, 2009). These complex interdependencies of learning may be an exciting field for future research.

DOI: https://doi.org/10.7275/hk9e-8d82

References

Ackermans, K., Rusman, E., Brand-Gruwel, S., & Specht, M. (2017). A first step towards synthesizing rubrics and video for the formative assessment of complex skills. In *Communications in Computer and Information Science* (Vol. 653). https://doi.org/10.1007/978-3-319-57744-9_1

Ackermans, K., Rusman, E., Brand-Gruwel, S., & Specht, M. (2018). The dilemmas of formulating theory-informed design guidelines for a video enhanced Rubric. In *Communications in Computer and Information Science* (Vol. 829). https://doi.org/10.1007/978-3-319-97807-9_10

Ackermans, K., Rusman, E., Brand-Gruwel, S., & Specht, M. (2019). Solving instructional design dilemmas to develop a Video Enhanced Rubric with modeling examples to support mental model development of complex skills: the Viewbrics-project use case. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-019-09668-1

Ackermans, K., Rusman, E., Nadolski, R., Specht, M., & Brand-Gruwel, S. (2019). Video-or text-based rubrics: What is most effective for mental model growth of complex skills within formative assessment in secondary schools? *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2019.07.011

Ackermans, K., Rusman, E., Nadolski, R., Specht, M., & Brand-Gruwel, S. (2021). Video-enhanced or textual rubrics: Does the Viewbrics' formative assessment methodology support the mastery of complex (21st century) skills? *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12525

Besterfield-Sacre, M., Gerchak, J., Lyons, M. R., Shuman, L. J., & Wolfe, H. (2004). Scoring concept maps: An integrated rubric for assessing engineering education. *Journal of Engineering Education*. https://doi.org/10.1002/j.2168-9830.2004.tb00795.x

Capelo, C., & Dias, J. F. (2009). A feedback learning and mental models perspective on strategic decision making. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-009-9123-z

de Grez, L., Valcke, M., & Roozen, I. (2014). The differential impact of observational learning and practice-based learning on the development of oral presentation skills in higher education. *Higher Education Research and Development*, *33*(2), 256–271. https://doi.org/10.1080/07294360.2013.832155

Fay, C. (2018). Text Mining with R : A Tidy Approach . *Journal of Statistical Software*. https://doi.org/10.18637/jss.v083.b01

Frerejean, J., van Strien, J. L. H., Kirschner, P. A., & Brand-Gruwel, S. (2016). Completion strategy or emphasis manipulation? Task support for teaching information problem solving. *Computers in Human Behavior*, *62*, 90–104. https://doi.org/10.1016/j.chb.2016.03.048

Gary, M. S., & Wood, R. E. (2011). Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal*, *32*(6), 569–594. https://doi.org/10.1002/smj.899

Gibbs, G., Simpson, C., & Macdonald, R. (2003). Improving student learning through changing assessment-a conceptual and practical framework. *European Association for Research on Learning and Instruction. EARLI Conference*.

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

Gielen, S., Tops, L., Dochy, F., Onghena, P., & Smeets, S. (2010). A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum. *British Educational Research Journal*, *36*(1), 143–162. https://doi.org/10.1080/01411920902894070

Gigante, J., Dell, M., & Sharkey, A. (2011). Getting beyond "good job": How to give effective feedback. In *Pediatrics*. https://doi.org/10.1542/peds.2010-3351

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

Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers and Education*. https://doi.org/10.1016/j.compedu.2013.09.019

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*. https://doi.org/10.1037/0021-9010.64.4.349

Jaehnig, W., & Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *Psychological Record*. https://doi.org/10.1007/BF03395573

Jarosz, A. F., & Wiley, J. (2014). *Journal of Problem Solving Special iSSue What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. 7*. https://doi.org/10.7771/1932-6246.1167

Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research*, *15*(2), 183–199. https://doi.org/10.1177/1362168810388711

Kubsch, M., Stamer, I., Steiner, M., Neumann, K., & Parchmann, I. (2021). Beyond p-values: Using Bayesian Data Analysis in Science Education Research. *Practical Assessment, Research, and Evaluation*, *26*, 2021.

Mattheiss, S. R., Alexander, E. J., & Graves, W. W. (2018). Elaborative feedback: Engaging reward and task-relevant brain regions promotes learning in pseudoword reading aloud. *Cognitive, Affective and Behavioral Neuroscience*. https://doi.org/10.3758/s13415-017-0553-5

Matthews, W. J., Buratto, L. G., & Lamberts, K. (2010). Exploring the memory advantage for moving scenes. *Visual Cognition*, *18*(10), 1393–1419. https://doi.org/10.1080/13506285.2010.492706

Mertler, C. (2001). Designing Scoring Rubrics for Your Classroom. *Practical Assessment Research Evaluation*, *7*(25), 1–10.

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). Package ' BayesFactor .' *R Package Version 0.9.12-2*.

Newton, P. M., Wallace, M. J., & McKimm, J. (2012). Improved quality and quantity of written feedback is associated with a structured feedback proforma. *Journal of Educational Evaluation for Health Professions*. https://doi.org/10.3352/jeehp.2012.9.10

Novak, J. D. (1985). Metalearning and metaknowledge strategies to help students learn how to learn. In *Cognitive structure and conceptual change*.

O'donovan *, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, *9*(3), 325–335. https://doi.org/10.1080/1356251042000216642

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., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 133–148. https://doi.org/10.1080/0969594X.2013.877872

Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195–203. https://doi.org/10.1016/j.stueduc.2013.10.005

QSR International Pty Ltd. (2015). What is NVivo? *Www.Qsrinternational.Com*.

Rohbanfard, H., & Proteau, L. (2013). Live vs. video presentation techniques in the observational learning of motor skills. *Trends in Neuroscience and Education*, *2*(1), 27–32. https://doi.org/10.1016/j.tine.2012.11.001

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. In *Psychonomic Bulletin and Review*. https://doi.org/10.3758/PBR.16.2.225

Ruegg, R. (2014). The effect of assessment of peer feedback on the quantity and quality of feedback given. *Papers in Language Testing and Assessment*.

Rusman, E., Nadolski, R., & Ackermans, K. (2019). Students' and teachers' perceptions of the usability and usefulness of the first viewbrics-prototype: a methodology and online tool to formatively assess complex generic skills with video-enhanced rubrics (VER) in dutch secondary education. In *Communications in Computer and Information Science* (Vol. 1014). https://doi.org/10.1007/978-3-030-25264-9_3

Rusman, Ellen, Martínez-Monés, A., Boon, J., Rodríguez-Triana, M. J., & Villagrá-Sobrino, S. (2014). *Gauging Teachers' Needs with Regard to Technology-Enhanced Formative Assessment (TEFA) of 21st Century Skills in the Classroom* (M. Kalz & E. Ras (eds.); pp. 1–14). Springer International Publishing. https://doi.org/10.1007/978-3-319-08657-6_1

Scheeler, M. C., Ruhl, K. L., & McAfee, J. K. (2004). Providing Performance Feedback to Teachers: A Review. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*. https://doi.org/10.1177/088840640402700407

Schrire, S. (2006). Knowledge building in asynchronous discussion groups: Going beyond quantitative analysis. *Computers and Education*, *46*(1), 49–70. https://doi.org/10.1016/j.compedu.2005.04.006

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

Stone, D. L., & Stone, E. F. (1985). The effects of feedback consistency and feedback favorability on self-perceived task competence and perceived feedback accuracy. *Organizational Behavior and Human Decision Processes*, *36*(2), 167–185. https://doi.org/10.1016/0749-5978(85)90011-1

Thijs, A., Fisser, P., & Hoeven, M. van der. (2014). 21E Eeuwse Vaardigheden in Het Curriculum Van Het Funderend Onderwijs. *Slo*, 128. http://downloads.slo.nl/Repository/21e-eeuwse-vaardigheden-in-het-curriculum-van-het-funderend-onderwijs.pdf

van der Pol, J., van den Berg, B. A. M., Admiraal, W. F., & Simons, P. R. J. (2008). The nature, reception, and use of online peer feedback in higher education. *Computers and Education*.

van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review*, *22*(2), 155–174. https://doi.org/10.1007/s10648-010-9134-7

Van Merriënboer, J. J. G., & Kester, L. (2005). The four-component instructional design model: Multimedia principles in environments for complex learning. *The Cambridge Handbook of …*. https://doi.org/10.1017/CBO9781139547369.007

Van Merriënboer, J. J. G., & Kirschner, P. A. (2007). Ten Steps to Complex Learning. In *In T. L. Good (Ed.), 21st century education: A reference handbook*. Thousand Oaks, CA: Sage. https://doi.org/10.1016/j.compedu.2008.06.001

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment and Evaluation in Higher Education*. https://doi.org/10.1080/02602930500353061

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691611406923

Wiggins, G. (2012). 7 Keys to Effective Feedback. *Educational Leadership*

Citation:

Ackermans, K., Rusman, E., Nadolski, R., Brand-Gruwel, S., & Specht, M. (2021). Feedback is a gift: Do Video-enhanced rubrics result in providing better peer feedback than textual rubrics? *Practical Assessment, Research & Evaluation*, 26(17). Available online: http://pareonline.net/getvn.asp?v=26&n=17.

Corresponding Author

Kevin Ackermans
The Open University of the Netherlands

email: kac [at] ou.nl