The Applicability of Feedback of Virtual Speech app metrics in a Presentation Technique Course

SEFI 2018 Conference

R.G.Klaassen
Programme Coordinator/Researcher
4TU Centre for Engineering Education – TU Delft
Delft, the Netherlands
E-mail: r.g.klaassen@tudelft.nl

P. de Vries [1]
Research Scientist
Delft University of Technology
Delft, The Netherlands
E-mail: Pieter.devries@tudelft.nl

Danielle S. Ceulemans[2]
Educational Research/Designer
4TU Centre for Engineering Education – TU- Delft
Delft,The Netherlands
E-mail: D.S.Ceulemans@tudelft.nl

Conference Key Areas: Engineering Education, Educational Innovation

Keywords: Engineering Skills, Innovative Teaching and Learning Methods

---

[1] P. de Vries
p.devries@tudelft.nl
[2] D.S. Ceulemans
d.s.ceulemans@tudelft.nl

1 INTRODUCTION

Industry 4.0 is up and running, the next industrial revolution of artificial intelligence and virtual reality systems are likely to profoundly change our world in general and our world of work in particular [1]. Virtual Reality systems might be a game changer as one of the emerging technologies in Education. Students have none or have very limited access to the new artificial and virtual reality systems techniques in the current educational system. Which causes students to have a (a shortage of skills that are in high demand) skills gap, if not confronted at an early stage with the techniques that will await them in the labour-market. Introducing VR/AR in education, is not that simple however, also because at present it is unclear to what extent VR/AR can contribute to a different kind of learning [2]. This paper is about the experiences with a practical experiment to better understand the consequences and value of such systems for education. The primary focus here is on feedback from a VR-system and reflection and their effects on the relevance for learning presentation skills. We experimented to find out the impact of a VR speech app as an additional practice tool in presentation techniques courses. The focus is on the organisation of the study and the outcomes including students' experiences in terms of relevant feedback for student learning and possible implications for learning activities in VR.

1.1 Presentation skills

The Academic Skills course "Presenting in Engineering" is considered highly relevant in the Engineering program (Meijers criteria). Academic presentation skills are included in both the Meijer's criteria and the Dublin descriptors [3]. Embedding soft skills into regular courses makes the acquisition of soft skills easier as a) content is provided, b) there is a purpose for practicing and c) it is just in time for application. This usually results into 1 or 2 ECTS courses running over the course of a semester parallel and integrated into a content course in one of the faculties. The number of participating students typically exceeds 50 student and runs up to 300 or more. The limited staff capacity and the time for guided practice do not suffice to coach all these students to a level of presentation skills considered acceptable within Engineering. As quite a few studies have shown immersive virtual reality learning environments (1) may boost learning outcomes and (2) reduce cognitive load due to fewer distractors[4],[5]. (3) It provides a representative practice in simulated situations or in other words; a sensing of what it will be like in real life [6]. So VR may offer an extra opportunity to practice presentation skills and boost learning outcomes.

1.2 Feedback in VR- speech research

This paragraph addresses what might be gained from using VR in practice situation. It is discussed what the findings are in Education, in using a three dimensional (3D) virtual reality systems for presentation skills.

Virtual reality speech applications and their relevance for education have been researched by quite a few large projects such as Metalogue, Rhema and Cicero [12],[13], [14],[15],[16]. The VR-trainers in these projects are specifically focused on screening non-verbal behaviours as these are easy to measure with available algorithms. Metalogue is a European project investigating the perceived benefits of a stand-alone presentation trainer when practicing elevator pitches. The presentation trainer provided immediate feedback on speech features such as loudness, speech rate, pitch/intonation and pauses. As well as body postures such as body and hand-movements, eye tracking. [12], [13]. Rhema focuses on objective non-verbal cues from the speaker and the immediate (in action) feedback of these cues via google glass. The interface in the Rhema project allows for real time detection of speech volume and speaking rate and it provides feedback during real time delivery [14]. The researchers have tried to establish the optimum mode of feedback, differentiating between sparse feedback, continuous feedback and no feedback during the presentation. The same authors [15] studied body postures and the effect of immediate feedback during presentations on patterns in body postures

Cicero is a platform and project in which the observed objective data (flow of speech, vocal variety, eye-contact, intonations, arm-hand movement) are correlated with qualitative assessments of experts to determine "good" feedback[16].

The virtual audiences in speech-apps are used to provide non-verbal feedback to the presenter. All the systems made use of different audiences. The adaptation of the audience or audience feedback varies in two ways: 1) orchestrated by prior coding of the machine for a type of audience or 2) adapted on the basis of the ongoing analysis of cues in the system. The non-verbal feedback signals (1) elevated attention, (2) lack of interest or (3) disagreement with the speaker.

According to the analysed results using the presentation trainer did yield significant perceived benefits as opposed to a control group, whom did not train the elevator pitch with the presentation trainer (Metalogue). In the Rhema project the perceived learning did greatly improve on the basis of parsed immediate feedback during the presentation on speech features. In the second Rhema study body postures were analysed: the presenters as opposed to a control group did become more aware of their body postures. In the Cicero project is was established pacing, vocal variety/intonation and eye contact either positively or negatively correlated with the assessments realised by performance experts. Whether the performance objectively improved could actually not be established as perceptions of participants and observers are the closest means of evaluations to learning outcomes.

We used a virtual speech app from Virtualspeech.com, that supports "at home practice" with a smart phone and google cardboard and free body movement. The combination of these tools help students to immerse into the experience of being in front of a real-life audience and receiving immediate feedback. The app provides (objective) immediate feedback on the following parameters: voice loudness, filler words like uh, speech-rate (time used per slide), eye-tracking (looking around across the room) and pauses, such that the students could become individual drivers of their own learning. It allows for a choice of multiple presentation settings, such as a classroom, a ted talk and others. The virtual speech app is used among other at Oxford university and according to the Virtual speech company with "good" results. Details are not available.

Virtual reality in this study is defined as an immersive, interactive, experience based learning environment, allowing practice in a targeted artificial environment representing a realistic situation in real time [7], [8], [4]. According to Dillenbourg [6], the effectiveness of virtual 3D interventions is a careful balance between immersion, engagement and reflection of students. If the environment is too realistic or immersive it will make learners uncomfortable (uncanny valley). If the engagement through gamification or exiting experiences is too much they will be engaged but the learning curve is not increasing. If the learner is comfortable and engaged with a little cognitive distance of the activity it will allow reflection and integration of knowledge into their own knowledge base. Before learning and reflection can take place, however, feedback is dearly needed [9]). '*Students have to be able to judge the quality of what they are producing and be able to regulate what they are doing during the doing of it'* [10] . Strijbos [11] points out that feedback is used when it is perceived as useful, acceptable, creates positive affect and induces a willingness to change behaviour on the part of the learner. Reflection may thus result from a combination of objective metrics, the attitude with respect to the feedback (given the right learning environment) and the expectations of being able to act on these variables.

The research questions focused on:

- To what extent does feedback from VR systems compared with other sources induce behaviour (reflection) supportive of learning presentation techniques.
- To what extent does VR at this moment offer a reasonable additional practice for presentation skills practice

2 METHODS AND TOOLS

The quasi- experiment was applied in a Bachelor course conducted in the faculty of Applied Physics, the life science track, and is called Biopharmaceutical Technology. The subgroups have all been questioned with a 5 point Likert scale survey prior to the final presentation on the relevance of the feedback used, the level of reflection that was used and the expectations of improved performance. A reflection on the VR system and what the students did with it was included as a separate assignment.

2.1 Sample

The presentation lab at Applied physics holds 6 project groups learning about ethics in which a total N-66 participants are taking part and follow an embedded and supportive presentation techniques course. Each presentation technique group of around 10 participants with an individual lecturer is sub-divided in groups of 3 participants. The groups are distributed across two control groups, two VR groups and two Peer feedback groups. In each of these conditions they practice presentation techniques by themselves, with peers or with the virtual app The Peer group (N= 20 students) intervention will not be discussed here, due to the limited space.

- Two control groups- N =20: practice independently in what-ever way, whenever convenient. (non vr/non peer feedback groups)
- Two VR groups – N = 20 Practice with a virtual reality speech app, with Google cardboard with peers. The app provides immediate feedback (speech analysis) to the participants, if you give your presentation in English. The virtual speech app can be accessed here https://virtualspeech.com/

2.2 The Questionnaire

The acceptance of feedback questionnaire is validated by Strijbos [11] and was used to obtain information about the likelihood of the feedback received, from the lecturer, the peers during class sessions and the Virtual Reality app, being accepted. It is also used as an indicator of engagement.

- The Feedback is evaluated for usefulness, acceptance, the willingness to do something with it and its positive/negative affect (engagement with the material)

The extent to which the VR app,the lecturer and peers in class, induced reflection was measured through the RISE- model of Wray [17] and the perceived performance improvement on performance indicators.

- Reflection consists of questions about: the extent students expected to improve their grades, did awareness of outcomes increase reflection, did they take up suggestions and do things differently, did they identify strengths and weaknesses and would they re-use of this practice mode.

In each condition they received information on how to present. The groups are divided on the basis of a preliminary questionnaire. At that moment they were also informed about the pending interventions

2.3 Preliminary Survey Results

A preliminary survey was used for group division where students could choose for one of the options communicated as extra practice opportunities no extra practice/peer group/VR app. The 43 participants (those whom were enrolled in the supportive online learning environment) were asked to fill out a questionnaire, the others were allotted to the available places left.

The questionnaire addressed their experience and preferred mode of practice.

The response rate was N= 33. From this group 12 students had little experience with giving presentation, 17 had moderate experience and 4 students had a lot of experience. Most of them practiced giving presentation in secondary education (100%), some with friends 39% and some in front of the mirror (30%). When asked to give their preferred practice mode 58% voted for the peer group option, 27% for the control group and only 15% for the VR mode. When asked to give the second preferred practice 52% chose the control group and 21% peer group and 27% the virtual option. Showing that in practice this group would not necessarily use virtual reality, if it were not offered in a course.

3 RESULTS

The reliability of the overall questionnaire at the end of the course was Cronbach's alpha 0.79. The reliability of each of the constructs used are included in table 1 measured on the entire group N= 68. The sub-constructs are all sufficiently high to be considered for further analysis. The mean scores show the control group (N = 20) and VR group (N=24) means on these constructs.

Table 1. overview reliability/mean for feedback/Reflection constructs

Constructs	**Sub-scales**	**Reliability Cronbach's alpha**	**Mean scores**	
Feedback			Control (sd)	VR (sd)
	Usefulness of feedback	.80	4.3 (.60)	3.9 (.71)
	Acceptance of feedback	.83	1.47 (.88)	1.9 (1.3)

	Willingness to use feedback	.79	3.94 (.99)	3.8 (.59)
	Positive/negative affect	.82	3.60 (1.07)/1.3 (.97)	2.29 (1.25) /1.6 (1.14)
Reflection				
	Rise model questions	.76	3.42	

3.1 Useful, Helpful and supportive Feedback

The Usefulness construct consisted of the question "*I consider this feedback as useful, helpful and supportive to future learning*", differentiated to feedback received from lecturers, peers and VR. Results on this scale show there are no significant differences between the experimental groups on the perceived feedback from lecturers peers in class. The VR group (N=24) was specifically asked whether they felt the VR speech app, was useful, helpful and supportive. The majority of the students felt it was not useful, not helpful and certainly not supportive to use a VR speech app.

Table 2. means scores usefulness constructs part 1. of the Feedback scale

Usefulness construct	VR-Mean (sd)	Frequencies
Useful	2.83 (1.59)	50% score 2 or lower
Helpful	2.67 (1.63)	54% scored 2 or lower
Supportive	2.42 (1.50)	62 % scored 2 or lower

The virtual reality group in general felt the Lecturer and peer feedback in class as less **useful**, to a significance level of .004 (F2.28) and slightly less **helpful** than the control group .010 (F2.36). It is not clear whether this is caused by the lecturer, the peers or the vr or the atmosphere in the group.

3.2 Acceptance of Feedback

Acceptance of feedback is discussed as in "*did the students dispute or reject the feedback given by the lecturer, their peers or the VR system*".

Table 3. means scores acceptance constructs part 2. of the Feedback scale

Acceptance of Feedback	Control	VR Group
Reject/Dispute		
Feedback from lecturer	1.30 (.80)	1.29 (1.04)
Feedback from peer	1.20 (.83)	1.75 (1.7)
Feedback from VR	n.a.	2.79 (1.9)
Feedback from lecturer	1.80 (1.4)	2.17 (1.7)
Feedback from peer	1.60 (1.3)	2.46 (1.8)
Feedback from VR	n.a.	3.21 (1.8)

Likert scale 1 (strongly disagree to 5.strongly agree). A low score meant no dispute or no rejection

No significant differences were found between the control group or VR group with respect to disputing or rejecting feedback from the lecturer and peers in class. On average the VR group was a little more likely to dispute the feedback from the lecturer, peers in class and especially the feedback from the VR.

3.3 Willingness to accept feedback

The "Willingness" to use the received feedback was covered by two questions: (1)*I am willing to improve my performance on the basis of the feedback provided by the lecturer, peers, individual experience*. (2) *I am willing to invest a lot of effort in revisions on the basis of* *received feedback*

Table 4. means scores “willingness to use feedback” construct part 3 feedback scale

Willingness to use feedback	Control Mean (sd)	VR Mean (sd)
To Improve performance		
Lecturers	4.50 (1.1)	4.58 (.58)
Peers in class	4.25 (1.2)	3.83 (.92)
Individual experience	4.05 (1.2)	3.38 (.88)
Virtual reality app		2.83 (1.77)
Revision effort based on feedback		
Lecturers	3.70 (.98)	4.17 (.76)
Peers in class	3.20 (1.28)	3.46 (1.10)
Based on VR		2.83 (1.74)

1 being strongly disagree and 5 strongly agree

The first question showed that each group was very ready to change their performance based on the feedback of the lecturer. They were equally ready to change their performance on the basis of peer feedback in class and on the basis of their individual experience. Note that none of the groups have significant within group-difference on this question. When we ask whether the students are willing to put effort in revision on the basis of the feedback from the lecturer, peers in class or VR practice, the lecturer is emerging as the undisputed authority, especially in the VR- group. (almost significant .08). VR was not really triggering a willingness to change.

3.4 Positive/Negative affect from Feedback

The last item on feedback concerns the positive or negative affect of the extra practice opportunity. The Control Group felt significantly more Satisfied, Confident and Successful on the basis of the feedback received in class (control group) than the VR practice (VR Group) felt about the feedback received from the vr speech app.(table 5).

Table 5. Mean scores/significance “Positive affect”

Positive affect	Control Mean (sd)	VR Mean (sd)	Control/vr significance
Satisfied	3.70 (1.08)	2.17 (1.27)	.000 (F1.82)
Confident	3.75 (1.30)	2.33 (1.31)	.001 (F1.31)
Successful	3.35 (1.13)	2.38 (1.27)	.011 (F1.00)

All calculations have been realised with an independent sample t-test. With respect to the Negative affect, no significant differences were found between any of the groups. Overall the mean scores were very low – indicating there was not an issue with being particularly angry, offended or frustrated.

3.5 Reflection Questions

The reflection questions pertained to (1) the expectations with respect to grade improvement.

(2) the effect on their changing their behaviour to become better and (3) the effect on subparts of the assessment metrics used for the presentation techniques course.

With respect to the 1st question "*has the practice experience improved your grade*", the students overwhelmingly in both groups felt the practice in class helped the most (control mean = 4.60(.59)/vr mean = 4.25(.99). Practice with peers was also highly rated (control mean = 4.50(1.28))/vr mean = 4.08(1.74). The individual practice was not really experienced as grade increasing activity (control mean = 3.15 (.81))/vr mean = 3.42 (1.53). Finally, the practice in VR was hardly considered as grade improving (vr mean = 2.58(1.53).

With respect to the 2nd questions: "*Did students change their behaviour on the basis of in class, at home without VR or VR practice?"*. Behaviour was changed most based on in class practice M=4.00 (89), somewhat as a result of at home practice without VR M=2.54(1.24), and very little as a result of practice with VR M=2.08(1.5). The control group took their "at home" practice a little more serious (in class M=3.50(143) and at home M=3.00(1.4)). None of these results show significant difference. When we asked whether students were interested to use the VR practice mode again, most students answered they would not use it again. And when we asked by which practice mode they felt most motivated the VR group said they were most motivated by in class practice activities. With respect to the 3rd Question: Did you improve performance on ……? We found the following results.

Table 6. means related to improvement on performance indicators

Performance indicators improvement	Control Mean (sd)	VR Mean (sd)
Preparation	2.45 (1.45)	2.75 (.76)
Content	2.70 (1.18)	3.13 (.90)
Presentation techniques	3.80 (1.20)	3.88 (.74)
Questions Rounds	3.90 (1.07)	3.67 (1.05)

When we look at the differences between the two groups, the VR group improved most on presentation techniques and the control group on Question Rounds, this seem logical as the VR speech app drew more attention to presentation techniques and the physical class more towards interaction. Yet none of this significant, differences are very small and based on perception. Even if they were it can by no means be simply concluded that using VR would be a root cause. The VR group as compared to the control group improved more on preparation and on Content according to their perception. It is unclear to what this can be attributed.

3.6 Reflection assignment

The included reflection assignment on what might be done with the feedback, did not spark useful insights into presentation performance. This resulted from 1. a lack of technological access, causing students not to have any feedback at all (no immersion or engagement) with the VR speech app. 2. A lack of relevant feedback metrics in the system (no engagement and no reflection). 3. A lack of bridging the information gap between standards and system. Teachers already pointed out: "what is "good" varies in different contexts and therewith cannot be gauged by objective metrics from a system". Despite the fact that teachers knew this, they have not actively been involved in guiding the students using the VR, due to amongst others lack of time. The lack of integration between classroom activities and the additional practice may have caused students' to make limited investments.

4 CONCLUSIONS AND DISCUSSION

The purpose of this research was to apply an experiment to better understand the value and the consequences of the use of VR in education. This study is about the use of a VR app to support presentation skill development. We have seen that the feedback provided in class including standardized assessment criteria is the best possible mode of feedback at this moment. VR in this experiment only did little to support perceived learning from extra practice.

Yet there were conditions in this quasi-experiment that could not be controlled for sufficiently to warrant solid conclusions. E.g. We do not know whether the results are due to intergroup differences of the VR group with the control group and of the different lecturers which have been teaching the group. And we do not know the interaction effect of the intervention with the perceived opinion of the feedback from lecturers and peers in class. It is presumed however that there has been an interaction effect, altering the opinions of the feedback. The conclusion we may draw, however, is that the readiness to work with VR both from a systems, teacher and students perspective is still too limited at the time of this experiment and needs further study to create the best extra practice options in which VR is embedded as a serious tool for feedback in improving presentations techniques.

Acknowledgement: A special thank you goes to Jolien Strouss, Angeniet Kam and lecturers supporting this quasi-experiment at TU Delft

5 REFERENCES

[1] Brettel, M, Friederichsen, N., Keller, M en Rosenberg, M. (2014) How Virtualization, Decentralization and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective, World Academy of Science, Engineering and Technology, *International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering*, Vol:8, No:1

[2] Klaassen R.G., de Vries, P. , Ioannides, M. G. & Papazis, S. (2017), Tipping your toe in the "Emerging Technologies" pond from an educational point of view, *Proceedings of the 45th SEFI Annual Conference 2017: Education Excellence for Sustainability, 18-21 Sept. 2017, Azores, Portugal.* 7 p.

[3] Dublin descriptors (2004), http://www.qaa.ac.uk/crntwork/nqf/nqf.htm

[4] Webster, R. (2017) Declarative knowledge acquisition in immersive virtual learning environments, *Interactive Learning Environments*, 24:6, 1319-1333, DOI: 10.1080/10494820.2014.994533

[5] Wadee S. Alhalabi (2016) Virtual reality systems enhance students' achievements in engineering education, *Behaviour & Information Technology*, 35:11, 919-925, DOI: 10.1080/0144929X.2016.1212931

[6] Dillenbourg, P., VR, AR, AW after wow, Presentation at the 4TU Centre for Engineering Education Virtual Onboarding Day, October 5, 2017, https://www.4tu.nl/cee/en/events/onboarding_day_using_virtual_and_augmented_reality_in_education/

[7] Mikropoulos, T. A., & Strouboulis, V. (2004). Factors that influence presence in educational virtual environments. CyberPsychology & Behavior, 7(5), 582–591. doi:10.1089/1094931042403109

[8] Sartain, J. (2012). Virtual reality: More virtual than reality. Retrieved from http://www.networkworld. com/news/2012/092412-virtual-reality-262519.html

[9] Hounsell, D., McCune,V, Hounsell,J. & Litjens, J. (2008). The quality of guidance and feedback to students, *Higher Education Research & Development*, 27:1, 55-67, DOI: 10.1080/07294360701658765

[10] Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144.

[11] 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, 291-303.

[12] van Helvert, J., Petukhova, V., Stevens, C., de Weerd, H., Börner, D., van Rosmalen, P., Taatgen, N. (2016). Observing, coaching and reflecting: Metalogue - A multi-modal tutoring system with metacognitive abilities. *EAI Endorsed Transactions on Future Intelligent Educational Environments*, *2016*(6), [e6]. DOI: 10.4108/eai.27-6-2016.151525 , http://www.metalogue.eu

[13] Schneider,J, Borner, D. Rosmalen, P. Specht, M, (2016), Can You Help Me with My Pitch? Studying a Tool for Real-Time Automated Feedback, IEEE *Transactions on Learning Technologies*: Volume 9 Issue 4, October 2016, Publisher: IEEE Computer Society Press

[14] Iftekhar Tanveer, M., Lin, E. and Hoque, M (2015) Rhema: A Real-Time In-Situ Intelligent Interface to Help People with Public Speaking

[15] Iftekhar Tanveer, M., Ru Zhao, Kezhen Chen, Zoe Tiet and Mohammed (Ehsan) Hoque, (2016), AutoManner: An Automated Interface for Making Public, Speakers Aware of Their Mannerisms, DOI: http://dx.doi.org/10.1145/2856767.2856785

[16] Batrinca, L, Stratou, G. , Shapiro, A. , Morency, L.P., Scherer, S. (2013), Cicero-Towards a Multimodal Virtual Audience Platform for Public Speaking Training, in R. Aylett et al. (Eds.): IVA 2013, LNAI 8108, pp. 116–128, 2013.c_Springer-Verlag Berlin Heidelberg 2013

[18] Wray, E. , RISE model , accessed. RISE Model by Emily Wray is licensed under a Creative Commons, http://www.emilywray.com/rise-model/ first accessed, in 2016