7th International Conference on Higher Education Advances (HEAd'21)
Universitat Politècnica de València, València, 2021
DOI: http://dx.doi.org/10.4995/HEAd21.2021.13093

Linking teachers' facial microexpressions with student-based evaluation of teaching effectiveness: A pilot study using FaceReader™

Maximilian Sailer, Ruben Schlag
Department of Education, University of Passau, Germany.

Abstract

This study seeks to investigate the potential influence of facial microexpressions on student-based evaluations and to explore the future possibilities of using automated technologies in higher education. We applied a non-experimental correlational design to investigate if the number of videotaped university lecturers' facial microexpressions recognized by FaceReader™ serves as a predictor for positive results on student evaluation of teaching effectiveness. Therefore, we analyzed five videotaped lectures with the automatic facial recognition software. Additionally, each video was rated by between 8 and 16 students, using a rating instrument based on the results of Murray's (1983) factor analysis. The FaceReader™ software could detect more than 5.000 facial microexpressions. Although positive emotions bear positive influence on the "overall performance rating", "emotions" is not predicting "overall performance rating", $\beta = .05$, $t(37) = .35$, $p > .05$. The study demonstrates that student ratings are affected by more variables than just facial microexpressions. The study showed that sympathy as well as the estimated age of the lecturer predicted higher student ratings.

Keywords: *Teaching effectiveness; facial expression recognition software; student evaluation; higher education.*

This work is licensed under a Creative Commons License CC BY-NC-ND 4.0

1. Introduction and facial expression recognition software

In order to move teaching evaluations beyond content-based questionnaires, researchers and lecturers in higher education can look towards other design-based fields that traditionally seek to observe or influence human behavior. In both advertising (Lewinski, Fransen & Tan, 2014) and food sciences (Danner, Sidorkina, Joechl & Duerrschmid, 2014), among others, facial expressions have been incorporated into experimental research for years as an indicator of subjects' emotional states. In these studies, the face itself is treated as a dependent variable (Ekman & Rosenberg, 1997). Of particular interest are what Haggard and Isaacs call "micromomentary facial expressions" (1966) which last only fractions of a second and are therefore difficult to recognize for untrained individuals (Yan, Wu, Chen, Liang & Fu, 2013). As a result, automated facial recognition technologies may help in analyzing these involuntary, spontaneous expressions.

This study seeks to investigate the potential influence of facial microexpressions on student-based evaluations and to explore future possibilities of using automated technologies for experimental studies in higher education. The study was conducted using the commercial, automatic facial recognition software FaceReader™ 6 (Noldus, 2014), which utilizes the Active Appearance Model (Cootes, Edwards & Taylor, 2001) to describe the face's texture and the key points on it. A trained neural network then continually examines and classifies the recognized face into various categories. These take form as the six universal emotions which exhibit pan-cultural qualitites—happy, angry, sad, surprised, scared, disgust (Ekman, 1970)—as well as "neutral". The accuracy of this tool has been validated early in various studies (den Uyl & van Kuilenburg, 2005; Lewinski & Butler, 2014), as has as an agreement between facial data and self-reporting when similar emotions were grouped together appropriately.

There are numerous potential applications of FaceReader™ technology in higher education, including the possibility of creating additional, custom expressions within the respective software (Sarkol-Teulings, 2021) to allow further specialization and adaptation of research. In online learning environments especially, facial recognition has been thought of as a possibility concerning student engagement detection (Dewan, Murshed & Lin, 2019; Liu, Wang, Yang & Wang, 2021).

2. Background

2.1. Emotion in the (virtual) classroom

In recent decades, increased attention has been given to the role that emotions play in cognitive processes in a (virtual) classroom (Titsworth, Quinlan & Mazer, 2010; Zembylas & Schutz, 2016). There is still a lack of tangible results through classroom research on

emotion in learning environments (King, Ritchie, Sandhu & Henderson, 2015); this applies even moreso to emotional expressions of teachers/lecturers. There have, however, been descriptions of reciprocal multilevel interactions between the emotions of students and their teachers/lecturers (Frenzel, Becker-Kurz, Pekrun, Goetz & Lüdtke, 2018). Additionally, most research on the topic has focussed the school level, as opposed to higher education (Mendzheritskaya & Hansen, 2019).

2.2. Evaluation of Teaching Effectiveness

In Europe especially, efforts toward achieving greater degrees of standardization and comparability of educational outcomes in higher education have increased following the Bologna process (Seyfried & Pohlenz, 2018). One assessment method frequently used by institutions of higher education is a student evaluation of both the courses offered and the teachers/lecturers involved. Many of the categories used therein—enthusiasm, rapport, and clarity, among others—were established in a factorial analysis by Murray (1983). While the approach of applying findings from personality psychology onto research about teachers and their (perceived) effectiveness has been perpetuated throughout the decades, it has never become dominant in the field of education (Göncz, 2017).

3. Research Questions

The following research questions (RQs) were developed in order to explore the potential uses of facial recognition software in higher education:

RQ1. What kind of emotional patterns does FaceReader™ software detect when analyzing non-calibrated facial microexpressions of different videotaped university lecturers?

RQ2. Is there a connection between video-taped lecturers' facial microexpressions recognized by FaceReader™ and the results of student evaluation of teaching effectiveness?

Hypothesis 1 (H1). The number of video-taped university lecturers' facial microexpressions recognized by FaceReader™ serves as a predictor for positive results on student evaluation of teaching effectiveness.

4. Methods

We applied an non-experimental correlational design to investigate the research questions, analyzing five video-taped international guest lectures, originally held at a Bavarian university. The lecturers seen in the video were all male and well known in the scientific community. The topics of their guest lectures ranged from finance to health and were located in the field of social sciences. Each video was between 80-100 minutes in length and the lecturers consented to the recordings being used for academic purposes. Non-calibrated

FaceReader™ software was used to analyze the microfacial expressions shown by the lecturers in the video at a framerate of 25fps. A general face recognization model was used for the analysis.

$N = 43$ learning science students voluntarily participated in the study to evaluate the video-taped lectures. Each video was rated by between 8 and 16 students, using a rating instrument based on the results of Murray´s (1983) factor analysis. We developed a 10-item rapport scale with items chosen for example from the students' perception of teaching effectiveness (SPTE) rapport, such as "communicates with students in a respectful way" (Jackson et al., 1999). The enthusiasm scale consisted of a total of 11 items, mostly drawn from the Enthusiasm Awareness Index (EAI; Rosenshine, 1970). The clarity scale likewise consisted of 11 items, such as "gives various examples to agreed with those actually taught" and was developed based on the results of several studies, mainly on Murray's (1983) factor analysis of the Teacher Behaviors Inventory. All values of the three scales were added up and divided by the number of items in order to calculate the average mean of the overall performance rating. Additionally, the rating instrument measured the stduent's level of sympathy towards their lecturer, as well as the lecturer's professional appearance, English language skills, estimated age, and the complexity of their lecture. Rapport, clarity and enthusiasm were measured on a 10-point Likert scale. The additional variables were measured categorically: Cronbach's α for Rapport is $\alpha = .93$; Cronbach's α for Enthusiasm is $\alpha = .96$; and Cronbach's α for Clarity is $\alpha = .86$. We used SPSS 26 for the analysis of the RQs. To answer RQ2, we used a linear regression model.

5. Results

Table 1 shows the descriptive results of the analysis via FaceReader™. The software could detect a total of 5,133 different facial microexpressions from the analysis of five guest lecturers. Expectedly, the categories "unknown" and "neutral" were detected most often. This can be explained by the fact that real-life lectures include movement on the part of the lecturer. Lecturers do not always look straight into the camera. They turn to the board, move their heads, nod, and do many other things. Therefore, the software was not able to analyze every frame of the recorded session. All emotions besides "scared" and "disgusted" were detected in the videos; "Happy" was recognized most frequently, 351 times. Facial microexpressions of single emotions lasted slightly more than 2 seconds on average.

Table 1. Frequency of facial microexpressions detected by FaceReader™.

Emotion	Frequency	Frequency in %	Average duration of emotion in seconds	Detected in how many videos
Unknown	2,262	44.07%	13.00	5
Neutral	2,086	40.64%	4.47	5
Sad	184	3.58%	1.95	5
Surprised	133	2.59%	2.06	5
Angry	41	.80%	2.20	5
Happy	351	6.84%	2.17	5
Scared	23	.45%	2.06	3
Disgusted	53	1.03%	2.16	3
Total	***5,133***			

Regarding the connection between teachers' facial microexpressions and the student-based evaluation of teaching effectiveness, the variables "overall performance rating" and "emotions" were found to be moderately negatively correlated, $r(38) = -.34$, $p = .019$. Since the categories "unknown" and "neutral" were predominant but did not provide additional information to the RQ, we decided to exclude those categories from further calculations. As Table 2 shows, we labeled the remaining emotions in two categories: positive and negative. Lecturer 5 received the best overall performance rating ($M = 7.57$); still, the recognized emotions were mostly negative. For the linear regression model, "overall performance rating" served as the dependent variable and "appropriate professional appearance," "level of sympathy for the lecturer," "estimated age of the lecturer," and "emotions (without neutral and unknown)" were included as independent variables. Although positive emotions have a positive influence on the "overall performance rating", "emotions" does not predict "overall performance rating" ($\beta = .05$, $t(37) = .35$, $p > .05$). "Appropriate professional appearance" also does not predict the "overall performance rating" ($\beta = -.142$, $t(37) = -1.19$, $p > .05$). "Level of sympathy for the lecturer" significantly predicts the dependent variable ($\beta = .651$, $t(37) = 5.34$, $p < .001$). The regression model also showed that students tend to rate younger lecturers (estimated age) more positively ($\beta = -.320$, $t(37) = -2.11$; $p < .05$). The independent variables also explained a significant proportion of variance in the overall rating ($R^2 = .57$, $F(4, 37) = 10.73$, $p < .001$).

Table 2. Connection between facial microexpressions and overall performance rating.

Lecturer	No. of observations	Frequency of emotions detected in %	Emotion (without neutral and unknown)	Overall performance rating (Average)
1	9	▪ Unknown (92.09%) ▪ Neutral (39.01%) ▪ Positive (.41%) ▪ Negative (1.08%)	Mostly negative	4.59
2	9	▪ Unknown (77.33%) ▪ Neutral (6.42%) ▪ Positive (9.29%) ▪ Negative (1.76%)	Mostly positive	6.97
3	9	▪ Unknown (60.64%) ▪ Neutral (35.94%) ▪ Positive (1.39%) ▪ Negative (2.03%)	Mostly negative	6.93
4	16	▪ Unknown (59.21%) ▪ Neutral (39.01%) ▪ Positive (1.53%) ▪ Negative (.26%)	Mostly positive	4.64
5	8	▪ Unknown (43.43%) ▪ Neutral (49.53%) ▪ Positive (3.34%) ▪ Negative (3.76%)	Mostly negative	7.57

6. Discussion

This study sought to investigate the potential of facial recognition software for higher education, particularly the connection of facial microexpressions detected by FaceReader™ with student-based evaluations of teaching effectiveness. The results showed a moderate negative correlation between lecturers' emotions and their overall performance ratings. Lecturers who show emotions in lectures do not necessarily receive more favorable effectiveness ratings from students. The expression of mostly negative emotions does not exclude the possibility of favorable student ratings. This might be explained by the lack of consideration of the content presented in the videos. Students' ratings are affected by more variables than just facial microexpressions. The study showed that sympathy and lower estimated age of the lecturer predicted higher student ratings. FaceReader™ software could

detect more than 5,000 facial microexpressions, but the leading category was "unknown," which is a limitation to the study. Further research should focus on the multidimensional connections between content, emotions, and student evaluation in higher education.

References

Cootes, T. F., Edwards, G. J. & Taylor, C. J. (2001). Active Appearance Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23*(6), 681-685. DOI: 10.1109/34.927467

Danner, L., Sidorkina, L., Joechl, M. & Duerrschmid, K. (2014). Make a face! Implicit and explicit measurement of facial expressions elicited by orange juices using face reading technology. *Food Quality and Preference, 32B,* 167-172. DOI: 10.1016/j.foodqual.2013.01.004

den Uyl, T. M. & van Kuilenburg, H. (2005). *The FaceReader: Online facial expression recognition*. Retrieved February 23, 2021, from https://www.vicarvision.nl/pub/fc_denuyl_and_vankuilenburg_2005.pdf

Dewan, M. A. A., Murshed, M. & Lin, F. (2019). Engagement detection in online learning: a review. *Smart Learning Environments, 6*(1). DOI: 10.1186/s40561-018-0080-z

Ekman, P. (1970). Universal Facial Expressions of Emotion. *California Mental Health Research Digest, 8*(4), 151-158. DOI:

Ekman, P. & Rosenberg, E. L. (1997). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS).* New York: Oxford University Press

Frenzel, A. C., Becker-Kurz, B., Pekrun, R., Goetz, T., & Lüdtke, O. (2018). Emotion transmission in the classroom revisited: A reciprocal effects model of teacher and student enjoyment. *Journal of Educational Psychology, 110*(5), 628–639. DOI: 10.1037/edu0000228

Göncz, L. (2017). Teacher personality: a review of psychological research and guidelines for a more comprehensive theory in educational psychology**.** *Open Review of Educational Research, 4*(1), 75-95. DOI: 10.1080/23265507.2017.1339572

Haggard, E. A. & Isaacs, K. S. (1966). Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy. *Methods of Research in Psychotherapy,* 154-165. DOI: 10.1007/978-1-4684-6045-2_14

Jackson, D. L., Teal, C. R., Raines, S. J., Nansel, T. R., Force, R. C. & Burdsal, C. A. (1999). The Dimensions of Students' Perceptions of Teaching Effectiveness. *Educational and Psychological Measurement, 59*(4), 580-596.

King, D., Ritchie, S., Sandhu, M. & Henderson, S. (2015). Emotionally Intense Science Activities. *International Journal of Science Education, 37*(12), 1886-1914. DOI: 10.1080/09500693.2015.1055850

Lewinski, P., den Uyl, T. M. & Butler, C. (2014). Automatic facial coding: validation of basic emotions and facs aus recognition in noldus facereader. *Journal of Neuroscience, Psychology, and Economics, 7*(4), 227-236. DOI: 10.1037/npe0000028

Lewinski, P., Fransen, M. L., & Tan, E. S. H. (2014). Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli. *Journal of Neuroscience, Psychology, and Economics, 7*(1), 1–14. DOI: 10.1037/npe0000012

Liu, T., Wang, J., Yang, B. & Wang, X. (2021). Facial expression recognition method with multi-label distribution learning for non-verbal behavior understanding in the classroom. *Infrared Physics & Technology, 112,* 103594. DOI: 10.1016/j.infrared.2020.103594

Mendzheritskaya, J. & Hansen, M. (2019). The role of emotions in higher education teaching and learning processes. *Studies in Higher Education, 44*(10), 1709-1711. DOI: 10.1080/03075079.2019.1665306

Murray, H.G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. *Journal of Educational Psychology, 75*(1), 138–149. DOI: 10.1037/0022-0663.75.1.138

Noldus. (2014) FaceReader: Tool for automatic analysis of facial expression: Version 6.0. Wageningen, the Netherlands: Noldus Information Technology B.V.

Rocha, C., Lima, R. C., Moura, A. P., Costa, T. & Cunha, L. M. (2019). Implicit evaluation of the emotional response to premium organic herbal infusions through a temporal dominance approach: Development of the temporal dominance of facial emotions (TDFE). *Food Quality and Preference, 76,* 71-80. DOI: 10.1016/j.foodqual.2019.04.001

Rosenshine, B. (1970). Enthusiastic teaching: A research review. *School Review, 78,* 499-514.

Sarkol-Teulings, D. (2021). *Creating a custom expression for Engagement: A validation study with FaceReader*. Retrieved February 23, 2021, from https://www.noldus.com/blog/custom-expression-engagement-validation-study

Seyfried, M. & Pohlenz, P. (2018). Assessing quality assurance in higher education: quality managers' perceptions of effectiveness. *European Journal of Higher Education, 8*(3), 258-271. DOI: 10.1080/21568235.2018.1474777

Titsworth, S., Quinlan, M. M. & Mazer, J. P. (2010). Emotion in Teaching and Learning: Development and Validation of the Classroom Emotions Scale. *Communication Education, 59,* 431-452. DOI: 10.1080/03634521003746156

Yan, W.-J., Wu, Q., Chen, Y.-H., Liang, J. & Fu, X. (2013). How Fast Are the Leaked Facial Expressions: The Duration of Micro-Expressions. *Journal of Nonverbal Behavior, 37*(4), 217-230. DOI: 10.1007/s10919-013-0159-8

Zembylas, M. & Schutz, P. A. (2016). *Methodological Advances in Research on Emotion and Education.* Zurich: Springer. DOI: 10.1007/978-3-319-29049-2