Received: 25 November 2022 | Accepted: 5 April 2023

DOI: 10.1111/bjet.13326

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

The first steps for adapting an artificial intelligence emotion expression recognition software for emotional management in the educational context

Jorge Fernández Herrero[1] | **Francisco Gómez Donoso**[2] | **Rosabel Roig Vila**[1]

[1]Department of General and Specific Didactics, Area of Didactics and School Organization, University of Alicante, San Vicente del Raspeig, Alicante, Spain

[2]Department of Computer Science and Artificial Intelligence (DCCIA), University of Alicante, San Vicente del Raspeig, Alicante, Spain

Correspondence
Jorge Fernández Herrero, Department of General and Specific Didactics, University of Alicante, Carretera San Vicente del Raspeig s/n - 03690 San Vicente del Raspeig, Alicante, Spain.
Email: j.ferher@ua.es

Funding information
Grant CIPROM/2021/17 funded by the Prometeo program from Conselleria d'Innovació, Universitats, Ciència i Societat Digital of Generalitat Valenciana (Spain). Redes de Investigación Docente Universitaria Program of the Instituto de Ciencias de la Educación de la Universidad de Alicante (call 2021-22). Ref: 5484.

Abstract

To test the suitability of an automatic system for emotional management in the classroom following the control-value theory of achievement emotions (CVT) framework, the performance of an emotional expression recognition software of our creation is evaluated in an online synchronous context. Sixty students from the Faculty of Education at the University of Alicante participated in 16 educational activities recording close-ups of their faces and completing the AEQ emotional self-report, as well as detailed reports from the subsequent review of their videos. In addition, they completed the VCQ-36 test to measure their volitional competencies and relate their influence on their emotional response. The results indicate a high coherence between the emotional expressions detected by the automatic system and the detailed emotional self-reports, but insufficient precision to meet the CVT requirements. On the other hand, both the AEQ test results and the emotion expression recognition software suggest students' preference for participative activities as opposed to passive ones. Meanwhile, statistical analysis results indicate that volitional competencies seem to influence the emotional response of students in the educational context, although the AI system does not show sufficient sensitivity in this field. Implications and limitations of this study for future work are discussed.

KEYWORDS

architecture for educational technology system, human–computer interface, improving classroom teaching, teaching/learning strategies

Practitioner notes

What is already known about this topic

- Student motivation and involvement in the learning process are highly related to appropriate emotional regulation, which can be associated with particular educational activities, strategies and methodologies.
- Deep learning technology based on convolutional neural networks feeds automatic systems focused on facial expression recognition from image analysis.

What this paper adds

- There is high coherence between the emotional expressions detected by the AI system and the students' emotional self-reports, but the AI system provides just emotional valences, insufficient to meet the CVT framework.
- Both emotional self-reports and the emotion recognition software suggest students' preference for active educational activities as opposed to passive ones.
- Volitional competencies seem to influence the emotional response of students in the educational context.

Implications for practice and/or policy

- It is possible to use automatic systems to effectively monitor the emotional response of students in the learning process.
- Only if sensitivity improved, a real-time, easy-to-interpret emotional expression recognition software interface could be implemented to assist teachers with the emotional management of their classes within the CVT framework, maximizing their motivation and engagement.

INTRODUCTION

Basic emotion theory and the control-value theory of achievement emotions

Basic emotion theory (BET) posits that emotions are distinct and brief states which include physiological, subjective and expressive components that allow humans to respond in ways that are typically adaptive concerning evolutionarily significant problems. This theory has been central to the study of emotional expression and its relative universality (Keltner et al., 2019), although its robustness is questioned (Durán et al., 2017). Approaches such as behavioural-ecological theory (Fridlund, 1991), which give preponderance to motivational or socialization aspects as the cause of emotional expressions over inner emotion, should be considered when trying to interpret the emotional states of individuals from the actions of their facial muscles.

Thus, the process of emotional expression is complex, multimodal in nature (Keltner & Cordaro, 2017) and responds to dynamic patterns of behaviour (facial action, vocalization, gaze, touch, posture, etc). The perception of emotion in others depends on cultural (Masuda et al., 2008), contextual (Hess & Hareli, 2017) as well as motivational and intentional aspects. In any case, we can understand BET as a sort of emotional grammar of socialization that somehow coordinates social interactions (van Kleef, 2016).

The control-value theory of achievement emotions (CVT) (Pekrun, 2006) proposes an integrative framework that considers antecedents and effects of the so-called achievement emotions in academic contexts, through estimates of subjective value and level of emotional control in both prospective and retrospective outcomes, considering as well the nature of the activity performed. This framework predicts variations in emotional outcomes between learning contexts, but generality in the mechanisms that relate antecedents and outcomes to emotions (Loderer et al., 2020).

To assess various achievement emotions experienced by students in academic settings, the Achievement Emotions Questionnaire (AEQ) (Pekrun et al., 2011) was designed. The AEQ differentiates between activity, prospective and retrospective emotions, measures positive and negative valences and activating and deactivating emotions, resulting in four categories of emotions: positive activating (enjoyment, hope and pride), positive deactivating (relief), negative activating (anger, anxiety and shame) and negative deactivating (hopelessness and boredom).

Furthermore, the CVT validity has been tested through the meta-analysis of emotional research within diverse technology-based learning environments (TBLE) (Loderer et al., 2020), suggesting that it is a valid framework for the management of the fundamental mechanisms of emotions in various learning environments, capable of assessing the robustness of their emotional design.

Students' motivation and involvement in the learning process are strongly related to appropriate emotional regulation (Camacho-Morles et al., 2021; Iqbal et al., 2021; Thomas & Allen, 2021). Developing the ability to regulate one's emotions and channelling them into positive emotional states favours self-directed learning, concentration and socialization in a collaborative work context (Hill et al., 2021; Järvenoja et al., 2019; Zhoc et al., 2018), an aspect that affects both students and faculty (Moè & Katz, 2021).

Thus, as suggested by the CVT framework, certain emotional outcomes can be associated with particular educational designs, strategies and methodologies, influencing students' motivation, resilience and performance (Daumiller et al., 2020). So-called active educational methodologies, which encourage student participation and interaction, are associated with positive mental states and higher levels of motivation and engagement, in contrast to passive didactic methodologies, in which students are limited to mere recipients of information (Labrecque et al., 2021; Morrell et al., 2021; Shaziya & Zaheer, 2021). Attributional feedback promotes enjoyment of the learning process and reduces frustration and negative emotional states, as opposed to performance-based feedback (Schrader & Grassinger, 2021; Stiller et al., 2019).

Additionally, students' volitional competencies are a factor closely linked to their emotional response in the learning context and, consequently, to their academic performance (Ramzi & Saed, 2019). High levels of self-regulation in the form of integrative emotional regulation (Benita et al., 2020; Roth et al., 2019) are associated with positive mental states, well-being and learning motivation (Grund et al., 2018), while a high degree of self-control is related to good academic performance (Troll et al., 2021).

Motivation, ie, the reasons that lead to pursuing a goal, plays a relevant role in the development of levels of self-regulation that does not require willpower, the latter being closely linked to self-control (Werner & Milyavskaya, 2019). High levels of self-control that are not complemented by sufficient levels of self-regulation can result in negative emotional states (Forstmeier & Ruddel, 2005), as self-control by itself does not always work, either because it is temporarily depleted or because of alterations in motivation, emotional state and attention influence its effectiveness (Englert, 2019).

Automatic systems to identify emotional expression in the learning context

Therefore, identifying students' emotional states and having didactic strategies that favour their modulation towards optimal states for learning can be a relevant educational improvement (Foutsitzi et al., 2019). The involvement and engagement of students have been studied, on the one hand, utilizing automatic systems, based on image analysis, the use of body sensors and the processing of log files. On the other hand, semi-automatic systems have been used based on the monitoring of response and success in certain tests. Finally, manual methods have also been employed, using observational checklists and self-report forms (Dewan et al., 2019).

Automatic systems focus on facial expression recognition from images analysed using deep learning technology based on convolutional neural networks (Krizhevsky et al., 2017), showing high accuracy levels within training and inference tests using predefined datasets (Chowdary et al., 2021; El Hammoumi et al., 2018; Srinivas & Pragnyaban, 2021). These approaches, however, align with a static understanding of emotions, typical of BET.

Furthermore, learning contexts involve complexity that does not always allow for the adequate recording of participants' faces, which can limit the reliability of such systems (Hirt et al., 2018). The use of hybrid convolutional neural networks, combining facial expressions with hand gestures and body postures, can minimize this issue (Ashwin & Ram Mohana Reddy, 2020), although the correct interpretation of an emotional expression, as seen before, poses a complex process involving multiple parameters.

Research goals and questions

The ultimate goal, of which this is the first step, is to develop an automatic system capable of assisting the teacher in the emotional management of the classroom in the learning process (Arroyo et al., 2014; Holstein et al., 2019; Kim et al., 2018), providing reliable information on the emotional state of the students in the classroom and assisting in the design of solid emotional educational strategies to facilitate desirable outcomes.

This work presents the system we designed based on face recognition, tested in an online synchronous learning context, which facilitated the capture of the participants' faces. To determine its strengths and weaknesses in the process of developing a comprehensive semi-automated classroom emotional management system, we pose the following research questions:

- What is the level of correspondence and reliability at this stage of development of the automatic system in identifying facial emotional expressions concerning students' self-reported achievement emotions in a synchronous online teaching context?
- Can we identify associations between the nature of the educational activities experienced and both the self-reported emotional states and those identified by the automatic system?

- Can we identify correlations between the volitional competencies of the participating students and both the self-reported emotional states and those identified by the automatic system?
- What aspects of the automated system should be modified, adapted or implemented to facilitate emotional classroom management consistent with a comprehensive framework such as CVT?

For this purpose, three levels of analysis are established. On the one hand, in addition to testing the AI system's theoretical accuracy based on its training with predefined datasets, the results derived from the analysis of the videos recorded during the educational activities set are contrasted with emotional self-reports made by the participants themselves. Secondly, to identify educational strategy preferences among the students, their emotional states during active educational activities are opposed to those shown in passive educational activities, contrasting the emotional self-reports with the results of the automatic system. Finally, the relationship between the participants' emotional states and their levels of self-control and self-regulation is analysed, verifying the coherence between the emotional reports and the automatic software results.

MATERIALS AND METHODS

Research model and procedure

This research implements a mixed-method case study (Plano Clark, 2019) with a group of students from the Faculty of Education of the University of Alicante during the 2020–21 academic year, under a synchronous online teaching modality. To analyse the emotional state of the students during the learning process, a close-up of the participants' faces was recorded on video during a series of predesigned educational activities. To ensure adequate quality of the images collected, necessary for proper processing by the AI program, participants were asked to use their own mobile devices, with precise instructions for proper framing and lighting, and set to record in at least HD quality (1280 × 720 pixels).

The determination of the emotional state of the participants in these educational activities was carried out using three parallel strategies. On the one hand, participants completed the Achievement Emotions Questionnaire (AEQ) (Pekrun et al., 2011) for each educational activity. Secondly, the video material was analysed by an automatic emotion expression recognition system that uses machine learning AI based on convolutional neural networks. Finally, subjects were asked to make detailed reports of their emotional state based on the analysis of their videos.

As a starting point, a preliminary study was carried out with a small sample of 10 students completing a total of 7 educational activities (Fernández-Herrero et al., 2021) in which detailed emotional reports were not yet carried out. It was decided to add these emotional self-reports to those collected by the AEQ to provide additional information that could be directly contrasted with that emitted by the automatic system. At the same time, the students completed the VCQ-36 test (Forstmeier & Ruddel, 2008) to measure their volitional competencies and subsequently analyse possible relationships between these and their emotional reactions during educational activities.

Statistical paired-sample analysis studies the association between the emotional response and the set of passive and active activities experienced. In addition, the correlations between the volitional competencies of the students and their emotional responses to the different educational contexts are analysed.

Participants and educational activities

The present investigation included a total of 60 participants aged between 19 and 22 belonging to the 2nd year of the Degree in Primary School Education at the Faculty of Education of the University of Alicante. All of them signed a participant information sheet, as well as an informed consent form, according to the Ethics Committee of the University of Alicante provided format. As shown in Table 1, 16 educational activities were designed with an approximate duration of 5 to 10 min, differentiating between participatory educational activities, in which the participation and interaction of the students are required, and passive activities, in which the student's intervention is limited to that of a mere spectator who receives information.

Instruments

Emotion expression recognition system

In this work, we propose an automatic emotion expression recognition system. The system is based on machine learning AI technology based on convolutional neural networks to state

TABLE 1 Educational activities.

Educational activities		Description	Type
A1	Master class	A lesson where students mainly listen to the teacher's lecture	Pas
A2	Test	Multiple-choice test on theoretical concepts of the course	Act
A3	Task correction	Group correction of practical assignments	Pas
A4	Kahoot game	Online collaborative game on theoretical aspects of the course	Act
A5	Exposition	Presentation of practical work in groups	Act
A6	Exposition correction	Group correction of the presentation of practical assignments	Pas
A7	Teamwork self-evaluation	Internal self-assessment of teamwork	Act
A8	Task correction	Group correction of practical assignments	Pas
A9	Task correction	Group correction of practical assignments	Pas
A10	Test	Multiple-choice test on theoretical concepts of the course	Act
A11	Exposition	Presentation of practical work in groups	Act
A12	Teamwork	Teamwork on practical assignments	Act
A13	Teamwork self-evaluation	Internal self-assessment of teamwork	Act
A14	Task correction	Group correction of practical assignments	Pas
A15	Video lesson	Viewing videos on theoretical concepts	Pas
A16	Video lesson	Viewing of videos on practical assignments	Pas

Own elaboration based on principles established by Labrecque et al. (2021).

whether a person is showing a neutral, positive or negative emotion. Although there are some tools available similar to that developed by us, the possibility of modifying its source code in a process of customization, evolution and adaptation to the CVT framework led us to the decision to develop a tool of our own.

Pre-processing method

To feed the system with samples of the highest possible quality, a pre-processing step is carried out. First, a face detector based on the Haar cascades classifier (Viola & Jones, 2001) is used to detect the faces present in the input image. As a result, bounding boxes of all faces present in the image are returned. The biggest face is chosen as the face of interest. Then, we normalize the pose of the face to be completely vertical.

For this, we use Dlib's face key point detector (see Appendix A), which provides a range of facial landmarks such as the position of the eyes, the nose, the boundaries of the mouth and others. This method also reportedly provides good accuracy while keeping the computational cost at bay (Pool, 2018). As shown in Figure 1, the pre-processing carried out to normalize the poses of the faces involves landmark detection and the transformation of the image so that the imaginary line that joins the eyes is completely horizontal.

Data selection and split method

To train our convolutional neural network, we mixed the samples of the AffWild2 (Kollias & Zafeiriou, 2019; Zafeiriou et al., 2017) and AffectNet datasets (Mollahosseini et al., 2019), compatible with our colour-based approach, as the images captured and analysed come in colour. These datasets are among the most comprehensive in existence and claim to include expressions of all types of minorities, races and diverse cultures. AffectNet is by far the largest database of facial affect in still images which covers both categorical and dimensional models, and CVT could be classified as one of the latter. Once we had the images for both datasets, we unified the labels to neutral, which was already labelled in the datasets; positive, which represents happy; and negative, which are anger, disgust, fear and sad emotions.

Then, we shuffled the images within each category and dataset and limited them to the lowest count, as a standard procedure so the system is not biased towards any category. We chose a random 30% of each category as the test split, which is also a standard figure for this matter. Ultimately, our dataset comprises more than 50,000 samples for training and 2000 samples for testing, equally distributed in the three categories mentioned before.

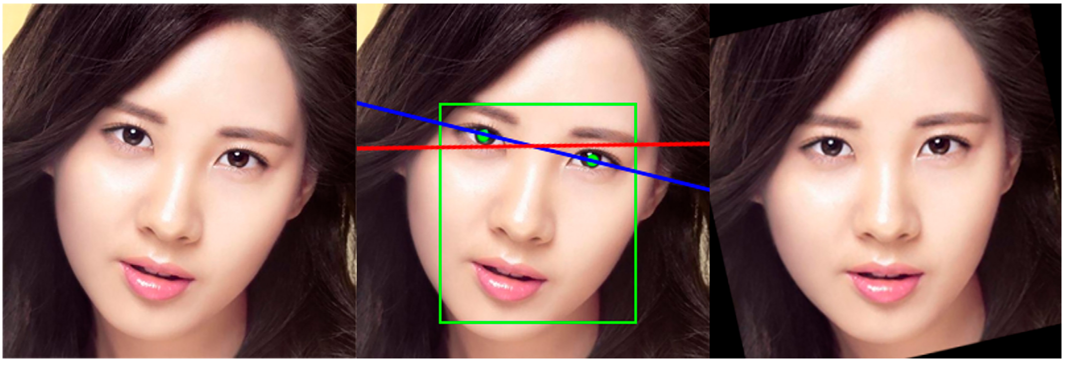

FIGURE 1 Steps of the normalization process.

Architecture definition and training

The pipeline of neural layers is shown in Figure 2. The input images are resized to 48×48 pixels regardless of the native resolution. The output of the architecture displays three different values that correspond to the probability of the face showing a neutral, positive or negative emotion. This pipeline features convolutional layers for feature extraction at different scales, pooling layers for dimensionality reduction and dense layers for classification. Dropout layers are also used to improve the generalization capabilities of the ensemble.

The optimization of the training split was carried out by Adam's method (Kingma & Ba, 2015), which used the categorical cross-entropy as a loss function. The weights were randomly initialized, so no fine-tuning was performed. We trained the system from scratch. A testing stage was performed every 10 training epochs. We used early stopping criteria to halt the training process by manually reviewing the training and testing accuracy and loss metrics over time to avoid over-fitting and provide the best test accuracy.

The system achieved 82.26% training accuracy and 74.95% test accuracy after the training process. While other similar proposals claim training accuracy over 95% (Chowdary et al., 2021; El Hammoumi et al., 2018), the training strategy and the datasets used, which are key to the outcome (Ge et al., 2022), are not the same. Consequently, we understand that the accuracy obtained aligns with the state of the art in this field (Chowdary et al., 2021; El Hammoumi et al., 2018).

Emotion recognition software

To automatically analyse the emotions displayed by a set of students during their online lessons using our emotion expression recognition system, a custom piece of software was designed. As shown in Figure 3, the program takes as input videos of the students during the online lessons. The software outputs a video in which the detected emotions are plotted, and several statistics of each video are also computed: frames and time with no face detected, as well as frames and time with neutral, negative and positive emotions.

When processing a document, the software generates a video that displays, next to the original video and synchronously, a graphic that plots the detected emotional state. Figure 4 shows a frame generated as an example.

Questionnaires

Self-Report Emotional Questionnaire

Participants completed the Achievement Emotions Questionnaire (AEQ) (Pekrun et al., 2011) after every educational activity. This questionnaire presents versions for regular classroom activities versus exams or performance tests, which were used accordingly. It answers questions based on a Likert scale ranging from 1 to 5 (strongly disagree to strongly agree) and offers high sensitivity to emotions associated with the educational event as a tool adapted to the CVT conceptual framework (Pekrun, 2006). Given that the automatic system emits only emotional valences (positive, negative or neutral), comparable values to these are extracted

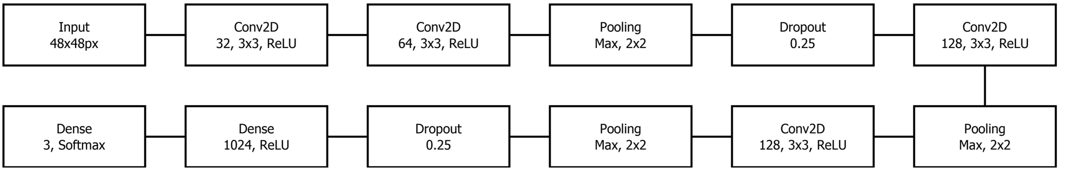

FIGURE 2 The convolutional neural network architecture proposed for emotion recognition.

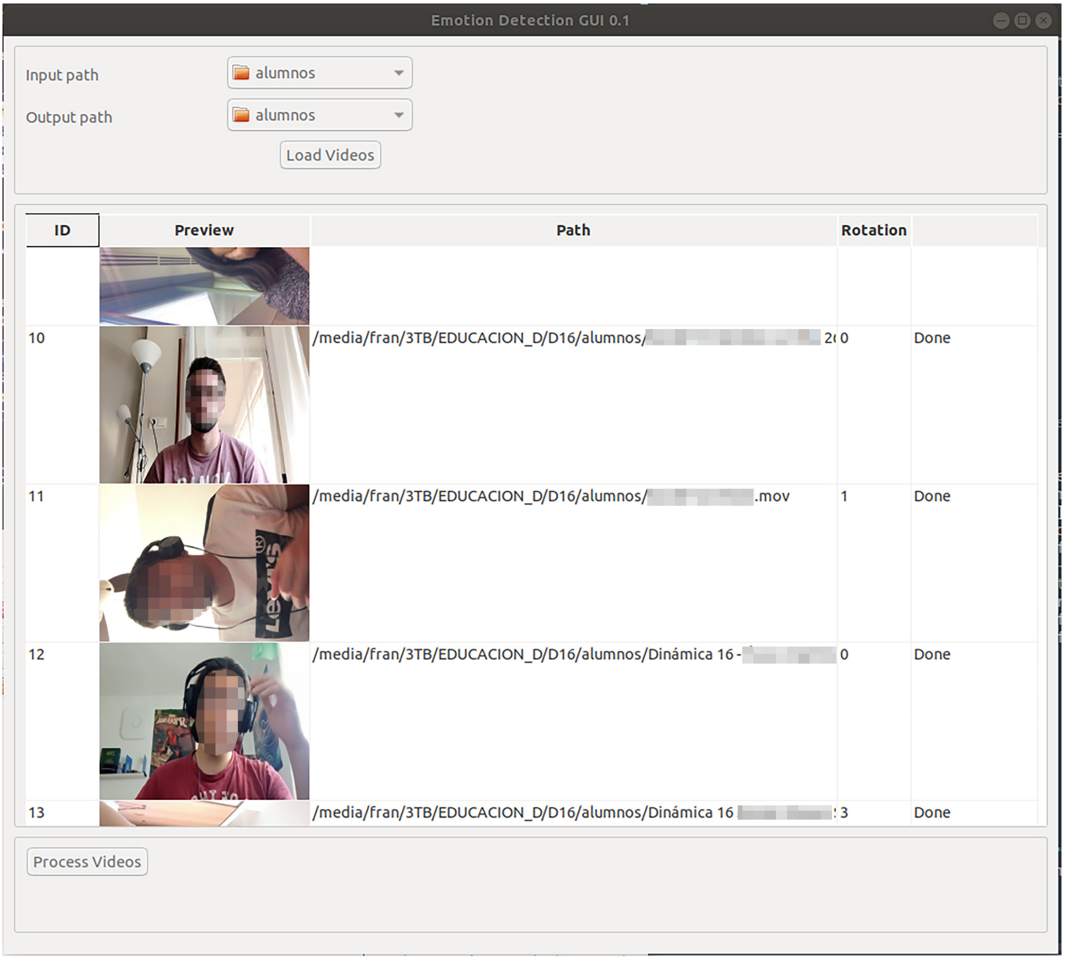

FIGURE 3 The main window of the emotion expression recognition software.

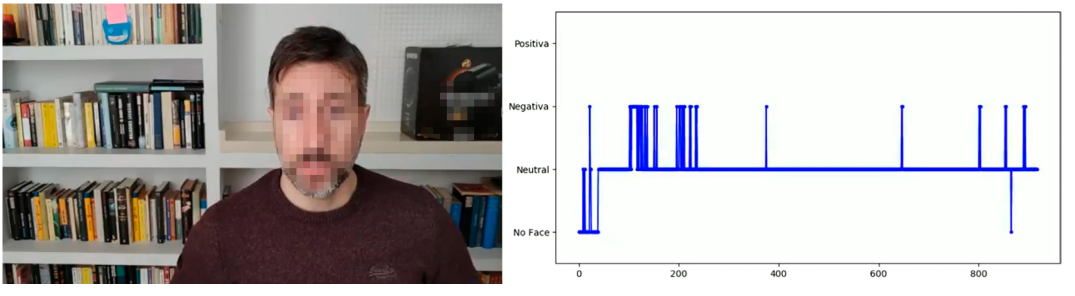

FIGURE 4 A frame of the synchronized footage with its AI software processed data.

from the results of the AEQ, with the aim of understanding to what extent this kind of AI resource can be effectively integrated into the CVT.

In addition, the questionnaire included two short-answer questions that inquired, on the one hand, about any external factors that could have influenced the student's emotional state and, on the other hand, whether they had found the activities experienced motivating and appropriate to promote learning. The questionnaire can be found in the corresponding link in Appendix B.

Since the AEQ design and extension do not allow for an emotional assessment for every frame recorded, as the AI system does, the students were additionally asked to develop detailed reports after reviewing the videos of each educational activity. These reports were used to contrast the coincidence between what they self-reported and what was detected by the automatic system. For this purpose, as described in Table 2, the basic emotional states considered and coherent with the AEQ test were grouped into negative, neutral and positive states to meet the information emitted by the AI system. To adequately contrast the data offered by these self-reports with those given by the automatic system, an additional option referring to a neutral state was added.

Volitional questionnaire

At the same time, the students who took part in the study completed the VCQ-36 test of volitional skills (Forstmeier & Ruddel, 2008) to analyse possible relationships between these competences and their emotional reactions during the educational activities. As Forstmeier and Ruddel (2005) point out, self-regulation skills are related to more positive emotions and fewer negative emotions, while self-control competencies exhibit the opposite pattern.

Statistical analysis

We start by contrasting the coherence among emotional self-reports and emotional expression detection given by the automatic AI system. Also, applying paired sample statistics, synergies between individual emotional self-report data with their volitional results are considered. The contrast between active and passive educational activity outcomes is analysed as well. Coefficients marked with an asterisk indicate that there are significant differences between the pairs compared (*0.05; **0.01; ***0.001).

To complement these analyses and measure the strength and direction of the linear relationship between variables, we study Pearson's correlation (Benesty et al., 2009). Values always range between −1 (strong negative relationship) and +1 (strong positive relationship). Coefficients marked with an asterisk indicate that the correlation is significant.

TABLE 2 Emotional states covered in the detailed self-report.

	NEGATIVE	NEUTRAL	POSITIVE
Enjoyment			X
Excitement			X
Pride			X
Relief			X
Annoyance	X		
Anxiety	X		
Shame	X		
Despair	X		
Boredom	X		
Neutral		X	

RESULTS

AEQ test results

Table 3 summarizes of the aggregate mean of the AEQ test results for each of the educational activities experienced.

Nevertheless, to translate the results into values analogous to those emitted by the automatic system, the proportional percentages corresponding to each reported emotion are calculated, interpreting intermediate values on the Likert scale as neutral. Figure 5 plots the proportion of activation and deactivation emotions reported by the participants.

Correction coefficients are applied since the AEQ does not define the same number of questions to report each emotion. Despite this, by defining a different questionnaire for tests than for the rest of the educational activities, the percentage of positive deactivating emotions (relief), only present in these cases, is very low.

Consistency between AI analysis and self-reporting

As summarized in Table 4, about 100 h of video material were collected throughout the 16 educational activities, of which almost 60% were satisfactorily analysed by the AI software. In turn, coherence ranging between 58.8% and 81.2% is identified between what is detected by the automatic system and what is self-reported by the participants, resulting in an overall cumulative coincidence of 71.1%. No significant variations are observed in this field when analysing active or passive experiences separately.

Analysis of emotional AI and self-reporting through educational activities

Figure 6 contrasts the AEQ test results with those identified by the automatic AI system. The cumulative corresponding to the AI system is obtained by calculating the percentage of aggregate emotional valences identified when analysing each frame of the set of educational activities. This results in a greater number of cases where the expression is identified as neutral.

The AEQ, completed after every activity, implies an assessment of the educational experience as a whole, resulting in a more balanced proportion between emotional valences. Also, its design differentiates between emotions before, during and after the educational event, and partial results are shown. An upwards trend of positive emotions gaining on negative ones can be identified. In any case, the accumulated results show a balanced proportion among positive, negative and neutral emotions, while the automatic system reveals a marked preponderance of neutral emotions, with negative emotions more than doubling the positive ones.

In addition, paired variables are studied to determine whether there are significant statistical differences among levels of negative, neutral or positive emotions if active or passive educational activities are involved. Table 5 shows their mean, standard deviation and standard error statistic results.

Table 6 confirms that there are statistically significant differences in all combinations, confirming that participants reported lower levels of negative emotions and higher levels of positive emotions in the set of participatory educational activities compared to the passive ones.

TABLE 3 AEQ test: Aggregate mean[a].

Ed. Activities	Enjoyment	Hope	Pride	Relief	Anger	Anxiety	Shame	Hopelessness	Boredom
A1	3.0	2.8	3.2	–	1.4	2.6	2.0	1.4	1.4
A2	3.4	3.1	3.9	3.7	1.7	2.6	1.6	1.5	–
A3	3.4	3.1	3.6	–	2.0	2.7	1.9	2.0	2.1
A4	3.8	3.4	3.8	–	1.8	2.9	1.4	1.6	1.0
A5	4.0	3.8	3.9	–	1.8	2.2	1.8	1.6	1.5
A6	3.9	3.7	3.8	–	1.8	2.5	1.8	1.8	1.5
A7	3.9	3.8	4.01	–	1.6	1.5	1.7	1.8	1.5
A8	3.6	3.5	3.7	–	1.6	2.6	1.8	2.0	2.0
A9	3.5	3.5	3.7	–	1.8	3.1	2.4	2.6	1.9
A10	3.6	3.5	3.9	3.7	1.8	2.4	1.6	1.4	–
A11	4.0	3.7	4.2	–	1.4	2.1	1.5	2.0	1.9
A12	3.8	3.6	4.0	–	1.6	2.3	1.8	2.3	1.8
A13	4.1	3.9	4.1	–	1.4	2.0	2.0	2.0	1.7
A14	3.7	3.6	3.8	–	1.6	2.2	1.8	2.4	2.1
A15	4.1	3.9	4.0	–	1.6	1.9	1.9	2.1	1.7
A16	3.8	3.8	3.8		1.6	2.1	1.4	2.2	2.4

[a](1 = Strongly disagree to 5 = Strongly agree).

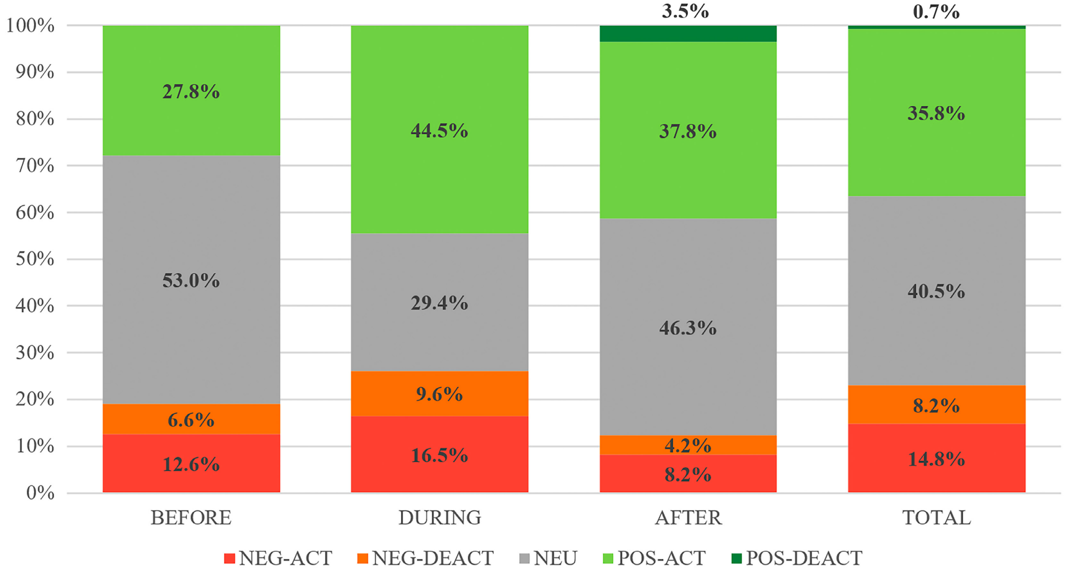

FIGURE 5 AEQ results: Percentage of activating and deactivating emotions.

TABLE 4 Accumulated and processed recorded video and consistency between AI analysis and self-reporting.

Educational activities	Recorded time (hh:Mm:Ss)	% AI processed	Time AI processed (hh:Mm:Ss)	% coherence AI/ self-report
A1	5:39:27	66.9%	3:47:12	72.9%
A2	4:32:19	50.5%	2:17:30	74.8%
A3	7:42:55	65.2%	5:01:56	75.8%
A4	6:18:15	61.9%	3:54:02	73.9%
A5	6:34:09	52.9%	3:28:40	81.2%
A6	8:55:09	67.1%	5:59:18	69.5%
A7	5:32:19	68.8%	3:48:42	68.6%
A8	7:04:32	61.0%	4:18:56	67.5%
A9	5:21:45	52.6%	2:49:10	73.3%
A10	6:17:27	55.8%	3:30:48	64.9%
A11	6:52:40	58.2%	4:00:03	58.8%
A12	5:18:48	58.5%	3:06:30	70.2%
A13	6:08:26	56.8%	3:29:14	74.6%
A14	7:18:54	57.3%	4:11:27	70.2%
A15	4:16:37	55.3%	2:21:49	69.4%
A16	4:50:23	54.2%	2:37:18	75.9%
Total	98:44:05	59.5%	58:42:35	71.1%

On the other hand, as shown in Tables 7 and 8, the results provided by the automatic system indicate that there are statistically significant differences in the last pair. This suggests that the emotional expression recognition software identifies a greater number of positive expressions during active educational activities than during passive ones.

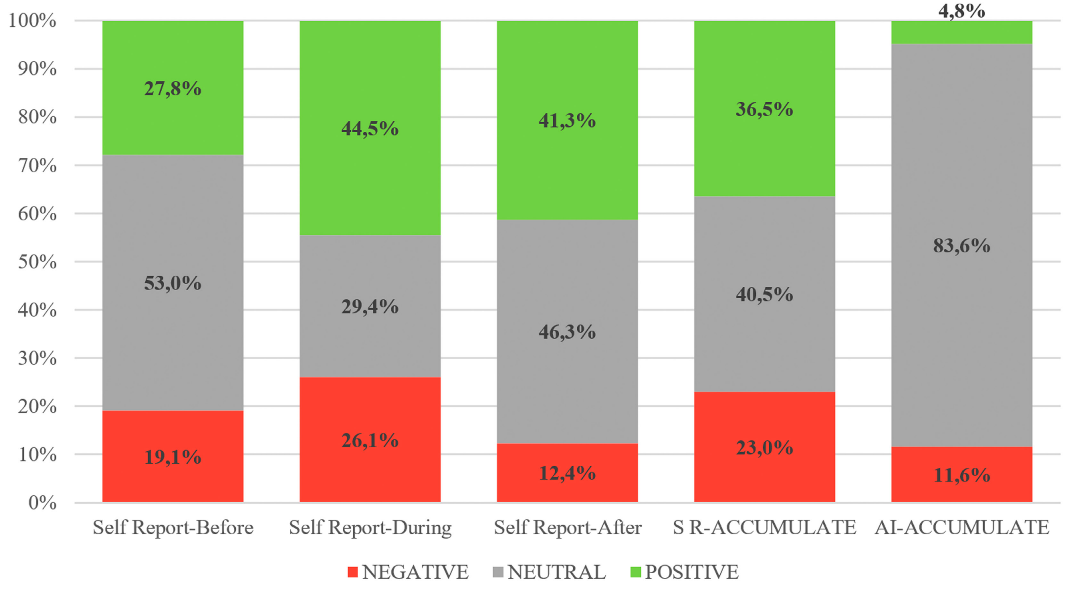

FIGURE 6 AEQ results versus AI results.

TABLE 5 Emotional self-report—Paired sample statistics (Active vs. Passive educational activities).

N = 60		Mean	Standard deviation	Average error deviation
Pair 1	NEG Emotional self-report/Act	21.22	12.227	1.578
	NEG Emotional self-report/Pas	24.32	12.741	1.645
Pair 2	N Emotional self-report/Act	27.63	8.870	1.145
	N Emotional self-report/Pas	29.67	9.735	1.257
Pair 3	POS Emotional self-report/Act	20.45	9.305	1.201
	POS Emotional self-report/Pas	16.08	8.516	1.099

TABLE 6 Emotional self-report—Paired sample statistics (Active vs. Passive educational activities).

		Paired differences							
					95% confidence interval of the difference				
		Mean	Standard deviation	Average error deviation	Inferior	Superior	t	Gl	Sig. (bilateral)
Pair 1	NEG Act-Pas	−3.100	9.025	1.165	−5.431	−0.769	−2.661	59	0.010**
Pair 2	N Act-Pas	−2.033	7.465	0.964	−3.962	−0.105	−2.110	59	0.039*
Pair 3	POS Act-Pas	4.367	8.501	1.097	2.171	6.563	3.979	59	0.000***

The asterisk numbers indicate statistically significant results.

Volitional levels and emotional outcome

As we can observe in Table 9, subjects with high self-regulation and low self-control report higher levels of positive emotions than those belonging to the group with low self-regulation and high self-control.

Extracting from these sets those individuals with self-regulation levels above 40 and self-control levels below 30, and vice versa, there are significant differences in the mean scores corresponding to the positive emotion variable, as shown in Table 10.

However, when analysing the results emitted by the emotion expression recognition software, the correlation with the participants' levels of volition is not so clear. If we extract individuals with self-regulation levels above 40 and self-control levels below 30, and vice versa, there are no significant differences in the percentage scores given in the group of variables (negative, neutral and positive emotions), as shown in Table 11.

When repeating these analyses but differentiating between active and passive educational activities, we obtain that, for both types, the emotional self-reports of individuals with high self-regulation and low self-control show higher levels of positive emotions than those belonging to the group with low self-regulation and high self-control, as there are significant differences in the mean scores corresponding to this variable. However, the strength of these results is higher for active educational activities (Sig. (bilateral) = 0.009**) compared to the passive ones (Sig. (bilateral) = 0.020*).

In turn, the results offered by the automatic system, like the complete set of educational activities, do not show significant differences in the percentage scores given in the group of variables, neither for the set of active activities nor for the passive ones.

Pearson's correlation coefficient results

Applying Pearson's correlation coefficient to the variables considered, different results are obtained. First, there is a significant positive correlation between levels of self-regulation and self-control (0.645***). In addition, a negative correlation is identified between high volitional levels and self-reported negative emotions, this being stronger when self-regulation is involved (−0.442***). A positive correlation is also detected between high self-regulation levels and self-reported positive emotions (0.364**).

Likewise, a slight negative correlation can be identified between levels of self-regulation and negative emotions identified by the AI system (−0.074). There is also a slight positive correlation between levels of self-regulation and neutral emotions identified by the automatic system (0.069). In any case, none of these are statistically significant results. Finally, no correlation was detected between levels of self-regulation and positive emotions identified by the software, or between levels of self-control and any emotions detected by the system. Figure 7 summarizes these results.

TABLE 7 AI analysis—Paired sample statistics (Active vs. Passive educational activities).

		Mean	Standard deviation	Average error deviation
Pair 1	NEG Emotional analysis/Act	12.82	14.005	1.808
	NEG Emotional analysis/Pas	12.62	13.415	1.732
Pair 2	N Emotional analysis/Act	80.39	16.394	2.116
	N Emotional analysis/Pas	83.00	14.956	1.931
Pair 3	POS Emotional analysis/Act	6.70	6.640	0.857
	POS Emotional analysis/Pas	4.37	7.231	0.934

TABLE 8 AI analysis—Paired sample statistics (Active vs. Passive educational activities).

		Paired differences							
					95% confidence interval of the difference				
		Mean	Standard deviation	Average error deviation	Inferior	Superior	t	Gl	Sig. (bilateral)
Par 1	NEG Act-Pas	0.202	8.469	1.093	−1.986	2.390	0.185	59	0.854
Par 2	N Act-Pas	−2.617	12.375	1.598	−5.814	0.579	−1.638	59	0.107
Par 3	POS Act-Pas	2.331	8.040	1.038	0.254	4.408	2.246	59	0.028*

The asterisk number indicate statistically significant results.

TABLE 9 Emotional self-report: Self-regulation/Self-control statistics.

		Mean	Standard deviation	Average error deviation
NEG	Low SR/High SC (*n* = 11)	54.82	27.440	8.274
	High SR/Low SC (*n* = 6)	37.33	7.711	3.148
N	Low SR/High SC (*n* = 11)	56.27	17.170	5.177
	High SR/Low SC (*n* = 6)	61.83	8.424	3.439
POS	Low SR/High SC (*n* = 11)	27.73	13.972	4.213
	High SR/Low SC (*n* = 6)	48.33	7.789	3.180

When contrasting active and passive educational experiences, the correlations previously described are reproduced as a whole. However, the negative correlation between self-reported negative emotions and levels of self-regulation is significantly higher within passive educational activities (−0.443***) compared to active activities (−0.393**). On the other hand, this correlation applied to levels of self-control is only significant within active activities (−0.279*). On the contrary, the positive correlation between levels of self-regulation and self-reported positive emotions is only significant in passive activities (0.393**).

Although they are not significant, the active experiences show a slight negative correlation between levels of self-regulation and negative emotional expressions detected by the automatic system (−0.030), being nonetheless more intense for the set of passive ones (−0.092). A weak positive correlation between positive emotional expressions detected by the software and levels of self-regulation is also identified for the passive activities (0.097).

Student's perceptions

Finally, Table 12 summarizes the participants' comments on two aspects related to the activities experienced. On the one hand, they indicate whether some factor external to the learning context may have influenced and altered their emotional state, elaborating if that is the case. On the other hand, they state whether the educational experiences in question are motivating and promote learning, briefly explaining their answer.

TABLE 10 Emotional self-report: Self-regulation/Self-control statistics.

	Equal variances assumed	Levene's test for equality of variances		*t*-test for equality of means						
									95% confidence interval of the difference	
		F	Sig.	t	gl	Sig. (bilateral)	Mean difference	Standard error difference	Inferior	Superior
NEG	Yes	3.155	0.096	1.508	15	0.152	17.485	11.593	−7.225	42.195
	No			1.975	12.578	0.071	17.485	8.852	−1.705	36.674
N	Yes	0.994	0.335	−0.738	15	0.472	−5.561	7.531	−21.613	10.492
	No			−0.895	14.950	0.385	−5.561	6.215	−18.812	7.691
POS	Yes	1.520	0.237	−3.311	15	0.005**	−20.606	6.223	−33.871	−7.341
	No			−3.904	14.941	0.001***	−20.606	5.278	−31.860	−9.352

Note: Test for independent samples.

The asterisk numbers indicate statistically significant results.

TABLE 11 AI emotional analysis: Self-regulation/Self-control statistics.

	Equal variances assumed	Levene's test for equality of variances		t-test for equality of means						
									95% confidence interval of the difference	
		F	Sig.	t	gl	Sig. (bilateral)	Mean difference	Standard error difference	Inferior	Superior
NEG	Yes	0.000	0.995	−0.613	15	0.549	−2.355	3.841	−10.542	5.833
	No			−0.610	10.274	0.555	−2.355	3.857	−10.918	6.209
N	Yes	3.075	0.100	1.122	15	0.279	5.134	4.574	−4.616	14.884
	No			0.997	7.569	0.350	5.134	5.149	−6.858	17.126
POS	Yes	5.483	0.033	−0.982	15	0.342	−2.779	2.831	−8.815	3.256
	No			−0.754	5.585	0.481	−2.779	3.686	−11.962	6.404

Note: Test for independent samples.

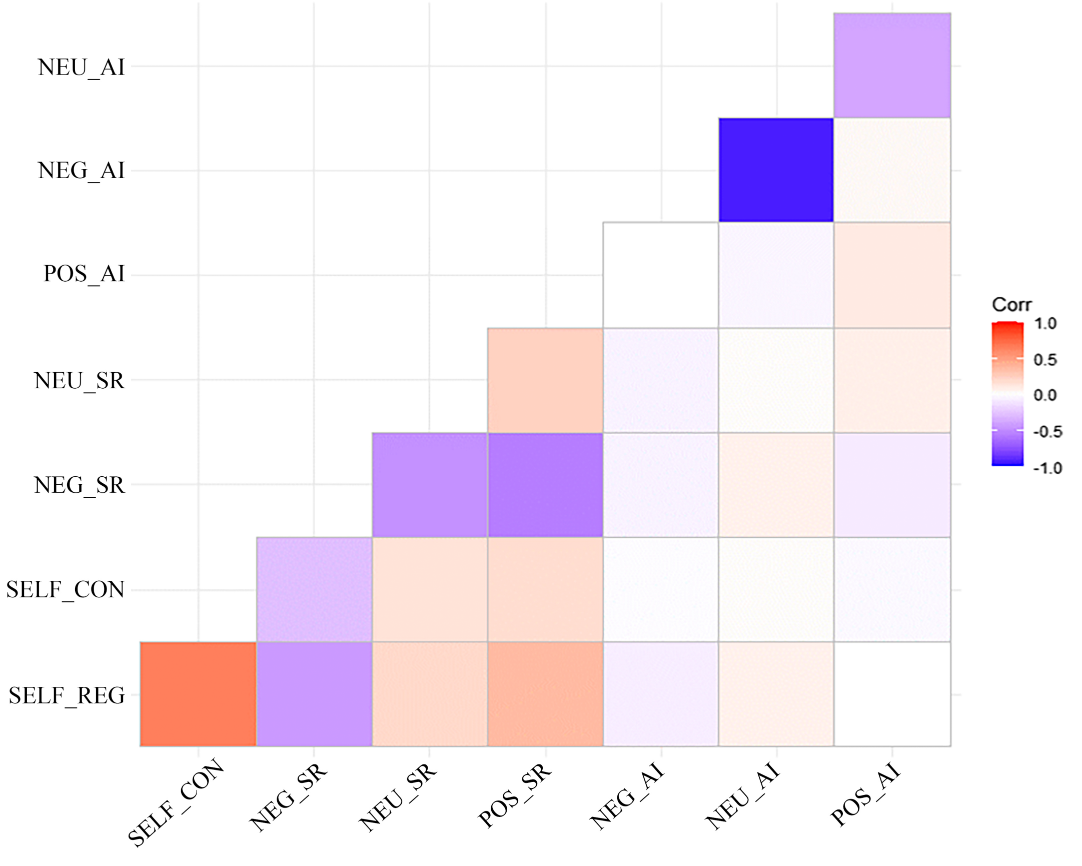

FIGURE 7 Pearson's correlation coefficient results between variables.

DISCUSSION

To answer the first research question, the high degree of consistency between the emotional self-reports and the automatic system results, around 70%, in contrast with Hirt et al. (2018), suggests that the software is reasonably reliable to determine the emotional valence of the students in the learning context. However, only 60% of the video material collected is recognized and processed by the emotional valence recognition software, which emphasizes the need for adequate framed facial images for the AI system to work (Grm et al., 2018).

Moreover, the reduction of the analysis to three emotional valences (negative, positive and neutral) undeniably favours the accuracy outcome but can lead to an oversimplification if more precise and personalized emotional management is required, as the CVT framework implies. The automatic system does not differentiate between activating and deactivating emotions, fundamental aspects of the CVT framework that establishes prospective and retrospective connections and would allow for the emotional management of the learning process associated with the design of educational proposals and experiences (Loderer et al., 2020).

Furthermore, there is an apparent contradiction between the proportion of emotional valences reported by the AEQ test, with a notable balance between neutral, positive and negative, and the automatic system results, with a high predominance of neutral emotional expressions. This is because the AEQ test provides the emotional perceptions of the activity

TABLE 12 Personal opinions of participants and external factors.

Activities		Passive	Active
Emotion is affected by external factors	Yes	25.1%	22.2%
	Reason	Nerves; Cell-phone issues; Stress; Noise; Context; Fatigue; Heavy workload; Technical issues; Health issues	Nerves; Cell-phone issues; Stress; Noise; Context; Fatigue; Heavy workload; Technical issues; Health issues; Teamwork issues
	No	74.9%	77.8%
Favours learning	Yes	79.5%	89.5%
	Reason	Helps to identify errors; Learning from others; Forces concentration; Promotes attention; Orientates	Helps fix teamwork issues; Favours implication; Favours attention; Motivates; Helps improve results
	No	20.5%	10.5%
	Reason	Makes me tense; Does not help fix errors; Boring; Distracting	If you are shy, it makes you nervous; People do not say the truth

as a whole, while the automatic system deals with every emotional facial expression from beginning to end. The calculated coincidence does not occur, therefore, compared to the AEQ test, but with subsequent detailed self-reports made by the students from the viewing of their videos.

With regard to the second research question, in line with other works (Labrecque et al., 2021; Morrell et al., 2021), results suggest the students' preference for active educational activities, with statistically significant differences in all combinations. Despite a high predominance of neutral emotions, the automatic system gives consistent results with the self-reporting in this area, since there are statistically significant differences in the positive emotions pair, meaning that the emotional expression recognition software identifies a greater number of positive expressions during the active educational activities than during the passive ones. However, both strategies seem to be perceived by the participants as complementary and useful depending on the contents, design and context.

In any case, the AEQ test results reflect interesting trends. These indicate an increase in positive emotions to the detriment of negative ones as the activities are completed. This probably indicates the participants felt more self-confident as activities progressed, something consistent with other educational experiences (Aziz Hussin, 2018). Interestingly, this phenomenon is more pronounced within active educational experiences, something aligned with the qualitative results, which indicate a greater capacity to promote learning and less influence of external factors for these.

Additionally, regarding the third research question, the students' volitional competencies seem to be related to their emotional response in the learning environment, in coherence with that suggested by others (Benita et al., 2020; Grund et al., 2018; Ramzi & Saed, 2019). Pearson's correlation coefficients indicate that the higher the volitional competencies, the fewer negative emotions and the more positive and neutral ones. However, the influence of self-regulation compared to self-control in this matter does not always match the assumptions made by Forstmeier and Ruddel (2005).

When differentiating between active and passive educational activities, results suggest a higher proportion of positive emotions corresponding to the active activities for those with high levels of self-regulation and low levels of self-control, compared to the passive ones. This moderate contrast might be consistent with the student's preferences for active educational activities previously detected. Nevertheless, only passive educational activities find a significant positive correlation between levels of self-regulation and self-reported positive emotions.

The AI software, on the other hand, does not obtain significant results in this field. Pearson's correlation analysis only finds mild connections between self-regulation levels and negative and neutral automatically detected emotional expressions, something that calls into question the sensitivity of the automatic system in this regard.

Finally, to answer the fourth research question, the current version of the AI system may be considered accurate to determine the emotional valences of the students as long as good quality face captures are provided, giving the emotional states and their evolution and tendencies throughout an educational activity. Nevertheless, it does not give the kind of information the CVT framework requires (Loderer et al., 2020). To meet these, the system should distinguish between activating and deactivating emotional states, making the right correspondence if monitoring regular class or exam activity. Additionally, feeding the system with antecedents and connections with alleged possible outcomes would align with the CVT emotional management strategies.

CONCLUSIONS

We can conclude that the emotional expression recognition software developed is sufficiently accurate to identify the emotional valences of the students during the learning process, only if obtaining adequate close-up images of the participants is possible. Additionally, the automatic system shows a sensitivity capable of reflecting the students' preference between active and passive educational activities.

On the other hand, students' volitional competencies are closely related to their emotional responses in the classroom. Participants with high levels of volitional skills tend to show more positive and neutral emotions and less negative ones. Meanwhile, the automatic system does not show sufficient sensitivity to identify any of these relationships.

At the same time, the AI system suffers from an oversimplification of the information it provides. This makes it difficult to adapt it to a comprehensive framework of emotional classroom management such as CVT (Pekrun, 2006). To do so, it should, first of all, differentiate between activating and deactivating emotions while maintaining sufficient precision. This would complement and simplify the application of CVT. The integration of antecedents and outcomes would allow for establishing correlations with emotional states and facilitating strategies and educational designs for comprehensive emotional management of the learning process.

Thus, future work will develop, in the short term, a real-time, easy-to-interpret emotional expression recognition software interface. One of the main objectives of the medium-term development of this resource will focus on increasing its sensitivity to meet the CVT framework, complementing the video information that feeds the system with other physiological data, such as biomarkers collected by activity bracelets or the volitional competencies of the participants.

Additionally, integrating the CVT framework, a function will be developed to assess relations among achievement emotions, antecedents and outcomes, to help teachers optimize their classes' emotional management. Further development of the proposed software should also involve practising teachers to test and discuss its utility.

The limited size and characteristics of the sample do not allow for a generalization of the results, although they are relevant as a case study. The special circumstances in which the research was carried out, under synchronous distance learning, may have influenced the perceptions and emotional responses of the participants. Its translation to a face-to-face teaching scenario should add new variables that may influence the emotional management of the classroom. Therefore, they should be taken into account and be the object of pertinent study.

ACKNOWLEDGEMENTS

This work has been supported by a grant from the Redes de Investigación Docente Universitaria Program of the Instituto de Ciencias de la Educación de la Universidad de Alicante (call 2021-22). Ref: 5484. This work has been partially funded by the grant CIPROM/2021/17 funded by the Prometeo programme from Conselleria d'Innovació, Universitats, Ciència i Societat Digital of Generalitat Valenciana (Spain).

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to report in this work.

DATA AVAILABILITY STATEMENT

Appendices mentioned in the text can be found in the following links:

Appendix A: http://dlib.net/.

Appendix B: https://forms.gle/xP8oRxfiuZwVxYPc6.

ETHICS STATEMENT

All the subjects participating in this research signed the appropriate informed consent form, following the guidelines of the Ethics Committee of the University of Alicante. All personal data acquired were anonymized changing names for numbers. Videos recorded were kept private at all times and only used for research purposes. Images included as examples of the software developed pixelated faces of participants when necessary.

ORCID

Jorge Fernández Herrero https://orcid.org/0000-0003-1545-8906
Francisco Gómez Donoso https://orcid.org/0000-0002-7830-2661
Rosabel Roig Vila https://orcid.org/0000-0002-9731-430X

REFERENCES

Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, *24*(4), 387–426. https://doi.org/10.1007/s40593-014-0023-y

Ashwin, T. S., & Ram Mohana Reddy, G. (2020). Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. *Education and Information Technologies*, *25*(2), 1387–1415. https://doi.org/10.1007/s10639-019-10004-6

Aziz Hussin, A. (2018). Education 4.0 made simple: Ideas for teaching. *International Journal of Education and Literacy Studies*, *6*(3), 92. https://doi.org/10.7575/aiac.ijels.v.6n.3p.92

Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In I. Cohen, Y. Huang, J. Chen, & J. Benesty (Eds.), *Noise reduction in speech processing* (pp. 1–4). Springer. https://doi.org/10.1007/978-3-642-00296-0_5

Benita, M., Benish-Weisman, M., Matos, L., & Torres, C. (2020). Integrative and suppressive emotion regulation differentially predict well-being through basic need satisfaction and frustration: A test of three countries. *Motivation and Emotion*, *44*(1), 67–81. https://doi.org/10.1007/s11031-019-09781-x

Camacho-Morles, J., Slemp, G. R., Pekrun, R., Loderer, K., Hou, H., & Oades, L. G. (2021). Activity achievement emotions and academic performance: A meta-analysis. *Educational Psychology Review*, *33*(3), 1051–1095. https://doi.org/10.1007/s10648-020-09585-3

Chowdary, M. K., Nguyen, T. N., & Hemanth, D. J. (2021). Deep learning-based facial emotion recognition for human–computer interaction applications. *Neural Computing and Applications*, 1–18. https://doi.org/10.1007/s00521-021-06012-8

Daumiller, M., Stupnisky, R., & Janke, S. (2020). Motivation of higher education faculty: Theoretical approaches, empirical evidence, and future directions. *International Journal of Educational Research*, *99*, 101502. https://doi.org/10.1016/j.ijer.2019.101502

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

Durán, J. I., Reisenzein, R., & Fernández-Dols, J. M. (2017). Coherence between emotions and facial expressions. In J. M. F. Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 107–129). Oxford University Press.

El Hammoumi, O., Benmarrakchi, F., Ouherrou, N., El Kafi, J., & El Hore, A. (2018). Emotion recognition in E-learning systems. In *6th International Conference on Multimedia Computing and Systems (ICMCS)*, Rabat, Morocco (pp. 1–6). IEEE. https://doi.org/10.1109/ICMCS.2018.8525872

Englert, C. (2019). The self-regulation of human performance: A critical discussion and future directions for self-control research. *Performance Enhancement & Health*, *6*(3), 156–157. https://doi.org/10.1016/j.peh.2019.04.001

Fernández-Herrero, J., Cazorla Quevedo, M. Á., & Roig-Vila, R. (2021). Inteligencia artificial basada en Deep Learning y Auto Reportes para el análisis emocional del alumnado en el proceso de enseñanza-aprendizaje. In R. P. E. Faubel (Ed.), *Nuevas tendencias en Investigación Educativa/New trends in educational research* (1st ed., pp. 15–25). Ediciones OCTAEDRO, S.L. https://octaedro.com/libro/nuevas-tendencias-en-investigacion-educativa-new-trends-in-educational-research/

Forstmeier, S., & Ruddel, H. (2005). On the superiority of self-regulation over self-control in psychotherapy and psychosomatic rehabilitation. *Verhaltenstherapie*, *15*, 158–166. https://doi.org/10.1169/000087450

Forstmeier, S., & Ruddel, H. (2008). Measuring volitional competences: Psychometric properties of a short form of the volitional components questionnaire (VCQ) in a clinical sample. *The Open Psychology Journal*, *1*(1), 66–77. https://doi.org/10.2174/1874350100801010066

Foutsitzi, S., Asteriadis, S., & Caridakis, G. (2019). An overview of affective models and ICT in education. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, Patras, Greece (pp. 1–8). IEEE. https://doi.org/10.1109/IISA.2019.8900783

Fridlund, A. J. (1991). Evolution and facial action in reflex, social motive, and paralanguage. *Biological Psychology*, *32*(1), 3–100. https://doi.org/10.1016/0301-0511(91)90003-Y

Ge, H., Zhu, Z., Dai, Y., Wang, B., & Wu, X. (2022). Facial expression recognition based on deep learning. *Computer Methods and Programs in Biomedicine*, *215*, 106621. https://doi.org/10.1016/j.cmpb.2022.106621

Grm, K., Štruc, V., Artiges, A., Caron, M., & Ekenel, H. K. (2018). Strengths and weaknesses of deep learning models for face recognition against image degradations. *IET Biometrics*, *7*(1), 81–89. https://doi.org/10.1049/iet-bmt.2017.0083

Grund, A., Fries, S., & Rheinberg, F. (2018). Know your preferences: Self-regulation as need-congruent goal selection. *Review of General Psychology*, *22*(4), 437–451. https://doi.org/10.1037/gpr0000159

Hess, U., & Hareli, S. (2017). The social signal value of emotions: The role of contextual factors in social inferences drawn from emotion displays. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 375–393). Oxford University Press.

Hill, J., Healey, R. L., West, H., & Déry, C. (2021). Pedagogic partnership in higher education: Encountering emotion in learning and enhancing student wellbeing. *Journal of Geography in Higher Education*, *45*(2), 167–185. https://doi.org/10.1080/03098265.2019.1661366

Hirt, F. S., Moser, I., Werlen, E., Imhof, C., & Bergamin, P. (2018). A comparison of students' emotional self-reports with automated facial emotion recognition in a reading situation. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 320–327). Association for Computing Machinery. https://doi.org/10.1145/3284179.3284230

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics*, *6*(2), Article 2. https://doi.org/10.18608/jla.2019.62.3

Iqbal, J., Ashraf, M. A., Perveen, S., Qureshi, N., Imran, Z., & Jin, N. (2021). How emotional intelligence influences cognitive outcomes among university students: The mediating role of relational engagement during the Covid-19 pandemic. *Frontiers in Psychology*, *12*, 4452. https://doi.org/10.3389/fpsyg.2021.711439

Järvenoja, H., Näykki, P., & Törmänen, T. (2019). Emotional regulation in collaborative learning: When do higher education students activate group level regulation in the face of challenges? *Studies in Higher Education*, *44*(10), 1747–1757. https://doi.org/10.1080/03075079.2019.1665318

Keltner, D., & Cordaro, D. T. (2017). Understanding multimodal emotional expressions: Recent advances in basic emotion theory. In J. A. Russell & J. M. Fernandez Dols (Eds.), *The science of facial expression* (Online ed., Vol. 1, pp. 57–76). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190613501.003.0004

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, *43*(2), 133–160. https://doi.org/10.1007/s10919-019-00293-3

Kim, Y., Soyata, T., & Behnagh, R. F. (2018). Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education. *IEEE Access*, *6*, 5308–5331. https://doi.org/10.1109/ACCESS.2018.2791861

Kingma, D. P., & Ba, J. (2015). *Adam: A method for stochastic optimization*. ICLR.

Kollias, D., & Zafeiriou, S. (2019). *Aff-Wild2: Extending the Aff-Wild Database for Affect Recognition* (arXiv:1811.07770). arXiv. https://doi.org/10.48550/arXiv.1811.07770

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Labrecque, L. I., Markos, E., & Darmody, A. (2021). Addressing online behavioral advertising and privacy implications: A comparison of passive versus active learning approaches. *Journal of Marketing Education*, *43*(1), 43–58. https://doi.org/10.1177/0273475319828788

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*, *70*, 101162. https://doi.org/10.1016/j.learninstruc.2018.08.002

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology*, *94*(3), 365–381. https://doi.org/10.1037/0022-3514.94.3.365

Moè, A., & Katz, I. (2021). Emotion regulation and need satisfaction shape a motivating teaching style. *Teachers and Teaching*, *27*(5), 370–387. https://doi.org/10.1080/13540602.2020.1777960

Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2019). AffectNet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, *10*(1), 18–31. https://doi.org/10.1109/TAFFC.2017.2740923

Morrell, B. L. M., Cecil, K. A., Nichols, A. M., Moore, E. S., Carmack, J. N., Hetzler, K. E., Toon, J., & Jochum, J. E. (2021). Interprofessional education week: The impact of active and passive learning activities on students' perceptions of interprofessional education. *Journal of Interprofessional Care*, *35*(5), 799–802. https://doi.org/10.1080/13561820.2020.1856798

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology*, *36*, 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Plano Clark, V. L. (2019). Meaningful integration within mixed methods studies: Identifying why, what, when, and how. *Contemporary Educational Psychology*, *57*, 106–111. https://doi.org/10.1016/j.cedpsych.2019.01.007

Pool, W. (2018). *Comparison of various landmark detecting techniques in the context of forensic facial recognition (public)*. http://essay.utwente.nl/77533/

Ramzi, F., & Saed, O. (2019). The roles of self-regulation and self-control in procrastination. *Psychology and Behavioral Science International Journal*, *13*(3), 1–4. https://doi.org/10.19080/PBSIJ.2019.13.555863

Roth, G., Vansteenkiste, M., & Ryan, R. M. (2019). Integrative emotion regulation: Process and development from a self-determination theory perspective. *Development and Psychopathology*, *31*(3), 945–956. https://doi.org/10.1017/S0954579419000403

Schrader, C., & Grassinger, R. (2021). Tell me that I can do it better. The effect of attributional feedback from a learning technology on achievement emotions and performance and the moderating role of individual adaptive reactions to errors. *Computers & Education*, *161*, 104028. https://doi.org/10.1016/j.compedu.2020.104028

Shaziya, H., & Zaheer, R. (2021). Strategies to effectively integrate visualization with active learning in computer science class. In N. Chaki, J. Pejas, N. Devarakonda, & R. M. Rao Kovvur (Eds.), *Proceedings of international conference on computational intelligence and data engineering* (pp. 69–81). Springer. https://doi.org/10.1007/978-981-15-8767-2_6

Srinivas, P. V. V. S., & Pragnyaban, M. (2021). Prediction of emotional score of the multiple faces of a photo frame through facial emotion recognition using the deep convolutional neural network. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(10), Article 10. https://doi.org/10.17762/turcomat.v12i10.4306

Stiller, A.-K., Kattner, M. F., Gunzenhauser, C., & Schmitz, B. (2019). The effect of positive reappraisal on the availability of self-control resources and self-regulated learning. *Educational Psychology*, *39*(1), 86–111. https://doi.org/10.1080/01443410.2018.1524851

Thomas, C. L., & Allen, K. (2021). Driving engagement: Investigating the influence of emotional intelligence and academic buoyancy on student engagement. *Journal of Further and Higher Education*, *45*(1), 107–119. https://doi.org/10.1080/0309877X.2020.1741520

Troll, E. S., Friese, M., & Loschelder, D. D. (2021). How students' self-control and smartphone-use explain their academic performance. *Computers in Human Behavior*, *117*, 106624. https://doi.org/10.1016/j.chb.2020.106624 van Kleef, G. A. (2016). *The interpersonal dynamics of emotion: Toward an integrative theory of emotions as social information*. Cambridge University Press. https://doi.org/10.1017/CBO9781107261396

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, 1*, Kauai, HI (pp. I-511–I-518). IEEE. https://doi.org/10.1109/CVPR.2001.990517

Werner, K. M., & Milyavskaya, M. (2019). Motivation and self-regulation: The role of want-to motivation in the processes underlying self-regulation and self-control. *Social and Personality Psychology Compass*, *13*(1), e12425. https://doi.org/10.1111/spc3.12425

Zafeiriou, S., Kollias, D., Nicolaou, M. A., Papaioannou, A., Zhao, G., & Kotsia, I. (2017). Aff-wild: Valence and arousal 'In-the-wild' challenge. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI (pp. 1980–1987). IEEE. https://doi.org/10.1109/CVPRW.2017.248

Zhoc, K. C. H., Chung, T. S. H., & King, R. B. (2018). Emotional intelligence (EI) and self-directed learning: Examining their relation and contribution to better student learning outcomes in higher education. *British Educational Research Journal*, *44*(6), 982–1004. https://doi.org/10.1002/berj.3472

How to cite this article: Fernández Herrero, J., Gómez Donoso, F., & Roig Vila, R. (2023). The first steps for adapting an artificial intelligence emotion expression recognition software for emotional management in the educational context. *British Journal of Educational Technology*, *54*, 1939–1963. https://doi.org/10.1111/bjet.13326