# Integrating Artificial Intelligence into Research on Emotions and Behaviors in Science Education

Angel Ezquerra [1*] (iD), Federico Agen [1] (iD), Iñigo Rodríguez-Arteche [2] (iD), Ivan Ezquerra-Romano [3] (iD)

[1] Department of Science, Social Science and Mathematics Education, Faculty of Education, Complutense University of Madrid, Madrid, SPAIN
[2] Department of Physics and Mathematics, Faculty of Education, University of Alcala, Alcala de Henares, SPAIN
[3] Institute of Cognitive Neuroscience, University College London, London, UK

**Abstract**

Most research on emotions and behaviors in science education has used observational or declarative methods. These approaches present certain strengths, but they have important limitations for deepening our understanding of the affective domain. In this work, we develop a method for analyzing the dynamics of affective variables during an inquiry-based activity with an artificial intelligence system that recognizes facial expressions. Although the study was carried out on 12 students, here we analyze data from one person to describe the method in detail. The videos were processed with a software which outputs behavioral and emotional signals. To analyze them, we applied centered moving averages with different widths. This allowed us to align and interpret the dynamics of emotional, behavioral, and learning actions. We found spikes of Surprise when the student seemingly implemented their models, and their predictions were not met. Our analysis suggests the existence of four phases in the inquiry-based activity with specific dynamic profiles. This work lays the foundations for researchers and teachers to develop tools to monitor emotions and behaviors.

**Keywords:** artificial intelligence, gesture analysis, affective domain, emotions, inquiry-based learning

## INTRODUCTION

Emotional and behavioral components are involved in teaching and learning processes. Both components play an important role in academic performance from a cognitive and motivational perspective (Artino et al., 2012; Loderer et al., 2019; Pekrun et al., 2017). Specifically, it has been found that these affective parameters influence processes such as memory, attention, problem-solving, self-regulation, study strategies, and academic results (Barrett et al., 2019; Chevrier et al., 2019; Graesser, 2020). In addition, research has shown the reciprocal or bi-directional relation between emotions and teaching-learning processes in science. It is thought that previous emotions or the anticipation of future achievements condition cognitive processes (Marcos-Merino et al., 2021). Conversely, the results of learning may predict the emotions that students will experience (Putwain et al.,

2018). Thus, to optimize teaching and learning strategies, both teachers and students should learn to identify and regulate their emotions and behaviors (Borrachero et al., 2014; Pekrun & Linnenbrink-Garcia, 2014).

In recent years, the number of studies on the role of the affective domain in education and in science teaching and learning has increased (Sinatra et al., 2014). As such, research in science education has focused on the emotions of students (Bellocchi & Ritchie, 2015; Dávila et al., 2021), in-service teachers (Bellocchi, 2019), pre-service teachers (Borrachero et al., 2014; Jeong et al., 2016; Jiménez-Liso et al., 2021b), or a combination of the above (Lombardi & Sinatra, 2013). Moreover, there are a few studies which have focused on single subjects–such as physics or biology–(Laukenmann et al., 2003; Marcos-Merino, 2019), self-regulation of emotions (Fredricks, 2011), and self-regulation of conceptual change in science (Sinatra & Taasoobshirazi, 2018).

**Contribution to the literature**

- Most studies on emotions and behaviors have collected data using observational and declarative (self-reports) procedures. To overcome some of their limitations, here we focus on the use of artificial intelligence (AI) to collect and codify facial expressions during science learning.
- To access different time scales, we applied centered moving averages with different window widths. Each of these allows to explore distinctive emotional and behavioral features.
- We were able to connect the dynamics of surprise with the educational actions carried out during an inquiry-based activity. Furthermore, we identified four phases in the inquiry process, according to the emotional and behavioral parameters.

Most studies on the emotional and behavioral components in education have collected data using observational (based on expert analyses), declarative (based on self-reporting) or both procedures at the same time (Harley et al., 2019; Loderer et al., 2019). These observational or declarative strategies present certain strengths such as low cost and versatility of implementation in common teaching and learning scenarios (Todd et al., 2020). However, they also have some important limitations.

Regarding observational methods, describing the presence and intensity of emotions requires training in identifying the expressive action units (eAU) that characterize facial movements. For people to reliably label eAUs, they need to undergo intensive and prolonged training and practice (Barrett et al., 2019). Regarding declarative methods, it remains unclear whether participants are capable of perceiving and self-reporting their emotions correctly, whether they really want to show them or whether they are dragged into their responses by 'social desirability' (Izard, 2009; Pekrun, 2006). Also, self-reports may capture only the effect of some events, depending on the filters implicitly applied by the subjects (Jiménez-Liso et al., 2021a).

## INTRODUCTION TO AUTOMATED MEASUREMENT OF AFFECTIVE & BEHAVIORAL COMPONENTS

To overcome the limitations faced by observational and declarative methods, a variety of strategies for measuring neurobiological and behavioral parameters have been incorporated into educational research. Importantly, research has shown strong overlaps among the brain representations of experiencing and expressing emotions (de Gelder, 2006; Vaessen et al., 2019). Therefore, both neurobiological and behavioral methods arguably provide reliable and consistent access to emotional components.

The existing neurobiological methods can be divided into two groups according to the part of the nervous system that is being measured: the autonomic or the central nervous system. The heart rate, respiratory rate, and blood pressure are examples of parameters measured from the autonomic nervous system (Clark et al., 2020; Monkaresi et al., 2017). The electrical activity of the brain, its changes in blood flow and other metabolic processes are examples of parameters measured from the central nervous system (Azari et al., 2020; Ihme et al., 2018; Wu et al., 2021). These techniques are helping us understand the underlying mechanisms involved in teaching and learning processes. However, they are often not suitable for real-life scenarios and interfere with students' behaviors.

The existing behavioral methods are mostly driven by artificial intelligence (AI) algorithms. These algorithms extract behaviorally relevant patterns from videos, soundtracks and texts. In science education research, eye-tracking and voice recognition systems have opened the door to research reading strategies, selection of scaffolds, the use of mono-modal and multimodal representations, or the observation of experimental setups by students (Abbaschian et al., 2021; Jarodzka et al., 2021; Tóthová & Rusek, 2021). However, accessing emotional components with eye-tracking and voice recognition still remains challenging.

Here, we focus on the use of an automated method to collect and codify facial expressions in an inquiry-based science activity. In general, facial recognition software uses a series of algorithm chains for identifying shapes, detecting faces, assigning facial points, identifying expressive action units, and determining associated emotions. The iMotions® software is an example of the current capacities of artificial intelligence, which we have here integrated into educational research.

The detection of faces is carried out by algorithms such as the Viola Jones cascade classifier (Viola & Jones, 2004). Within the facial image, the assignment of reference points is performed through different algorithms. These algorithms assign facial markers (nose, eyebrows, lips, etc.) to the image. These markers are then compared with pre-existing reference patterns. The analysis of the differences between the facial markers and the pre-existing reference patterns defines each eAU (Ekman & Friesen, 1976). For example, the distance between the eyebrows defines the intensity of the expressive action of frowning. Each eAU identified by a number (eAU1, eAU2, etc.) corresponds to the contraction of a facial muscle or group of them. As such, the work of several authors has established the

Figure 1. The expressive AUs involved in the emotion of joy. The figure is reproduced from Barrett et al. (2019) with permission from the authors and PSPI (Psychological Science in the Public Interest)

relationship of different eAUs with specific emotions (Barrett et al., 2019; Ihme et al., 2018; Sayette et al., 2001). Importantly, these facial expressions are often visible and recognizable by all of us in our personal interaction.

The databases used by artificial intelligence software contain statistical distributions of facial expressions from multiple geographic locations, demographic profiles and recording conditions. This means that the categorization of emotions is carried out at a purely statistical level, considering movements of the facial reference points corresponding to each eAU. For example, the emotion of joy is identified by the 'Duchenne configuration' (Figure 1). In this example, the eAUs are essentially eAU6 and eAU12, which correspond to the activity of the muscle that lifts the cheeks, the orbicular muscle of the eye in its orbital portion, and the muscle that lifts the angle of the mouth, the zygomaticus major (Barrett et al., 2019).

The artificial intelligence software extracts the eAUs independently and then applies algorithms to assess the probability that a specific configuration of eAUs is produced by a given emotion (e.g., joy). Thus, the facial action code system (FACS) system performs a modular identification of the emotions based on the combination of the different eAUs (Ekman & Friesen, 1978). This provides standardization and objectivity in the identification of emotions as opposed to the limitations found when the identification is performed by human observers.

The iMotions® software detects human faces and assigns 34 facial reference points. Based on these data, this software uses the algorithm affectiva AFFDEX® to identify head movements, the interocular distance and 20 facial eAUs. Finally, it associates various facial expressions with the seven basic emotions as well as with some behavioral parameters.

To bring this technology into science education, the relationship of emotional and behavioral parameters with teaching-learning processes needs to be further studied. Hence, this work has the following objectives:

1. To implement an experimental design which allows the study of emotional and behavioral parameters from facial expressions in a science education activity.
2. To develop an analysis procedure to use the signals provided by the AI system for studying the dynamics of emotions and behaviors manifested by students.
3. To link the data provided by the AI system with educational actions.

## EXPERIMENTAL DESIGN

The experimental strategy involved the selection of an appropriate science education activity, the development of a procedure to analyze the emotional and behavioral parameters provided by the AI system and the educational actions carried out by the students.

### Science Education Activity: The Black Box

To integrate facial expression recognition technology into research in science education, we selected and adapted an activity that elicits changes in emotional and behavioral states. This activity was carried out in the context of the Spanish master's in secondary education (major in science, specifically in biology and geology) of the Faculty of Education at Complutense University of Madrid, Spain. Once the activity had been completed, there was a discussion with the pre-service teachers on the emotions that they had felt. The overarching aim of the session was to explore the different emotions they had felt and to consider what their future pupils would feel in equivalent activities. This qualitative data is not shown in this article.

The activity is coherent with the inquiry-based science education (IBSE) approach (Abd-El-Khalick et al., 2004; Minner et al., 2010). In this activity, students had to discover the content of a black box without opening or breaking it. This activity has been previously implemented by different authors such as Haber-Schaim et al. (1979), and Lederman and Abd-El-Khalick (1998).

The black box (9x6x20 cm) used in the activity contained coins of different sizes and materials: one of 1 euro, one of 20 cents, one of 10 cents, one of five cents and two of one cent. All the coins moved freely inside it. Students were given a couple of magnets and a set of all the existing euro coins to explore their interaction with the magnets, their friction with the box surface, the sounds they produced, etc. The students had to use their own skills to formulate hypotheses, carry out tests and draw conclusions regarding the content of the box. At

the end of the activity, the students had to report what they thought the box contained.

Videos were recorded for their subsequent analysis with the iMotions® software. Video cameras were placed on tripods one meter away from the students, who were sitting down on a chair by a table with the items and had to solve the task individually. This arrangement appeared to be the most appropriate to obtain the best view of the face, the upper body, and what the students did with the items.

Before starting the experimental section, students were given instructions on how to carry out this activity. They were told to remain within the image frame throughout the recording. Students were also told to avoid placing the box and their hands between the camera and their face. In addition, the necessary permissions were requested for the recording and the later use of the images for research purposes.

The experimental section lasted twenty minutes. The recordings made during the activity were stored and labelled according to a system generated ad hoc. The videos were edited for subsequent analysis (e.g., by cropping out the frames recorded outside of the duration of the activity).

### Procedure to Analyze the Parameters Provided by the AI System & the Educational Actions Carried Out by the Students

We recorded 12 students performing the inquiry-based task. In this article, we aim to explain in detail how to integrate automated facial expression recognition on educational research, and particularly in science education. Therefore, we will focus on data from one single individual. Specifically, we selected the video in which the face was visible and within frame for most of the time.

After edition, the selected video lasted 1,203,569 milliseconds (approximately 20 minutes) and it had 36,070 frames, so the time resolution was about 33 ms. This video was processed by the artificial intelligence system. This analysis provided a total of 141 entries per frame, i.e., 36,070 frames x 141 data per frame=5,085,870 total entries.

The data was exported as a CSV file to be analyzed by custom-written Python code. iMotions® also provides an output video that includes animated graphs of the recorded emotions alongside the original video. Thus, it was possible to observe in parallel the emotional and behavioral signals, and the student's actions to understand their relationship.

The system detects human faces and assigns 34 reference points. This set of points allows to associate facial expressions with emotional and behavioral parameters. In this study, the parameters analyzed were surprise, joy, disgust, fear, contempt, sadness, and anger (basic emotional parameters), plus attention and engagement (behavioral parameters).

The system assigns to each parameter a numeric score in intensity, from absent (0%) to fully present (100%). This percentage is a measure of the similarity between the graph of facial markers detected and the set of reference patterns stored by the artificial intelligence system. The software estimates the presence of each parameter at each frame.

To reduce the noise from the signals, we applied a moving average with custom-written Python code. This statistical technique replaces each value in a time-series by the average of the window around each original value. To smooth the signals and keep relevant patterns in the data, the appropriate mathematical factors must be selected, such as the width of the window around the original value and the position of the original value within the window.

By adjusting these factors, we could find an optimal way of associating the emotional and behavioral signals to the dynamics of the observed learning actions. Specifically, the position of the original value within the window was established to be in the center. Thus, we used a centered moving average, since the emotional and behavioral parameters measured are influenced by the recent past (e.g., the challenges faced in the inquiry-based process) and modulate the future behavior (e.g., they drive students' decision-making). The widths of the windows initially considered were of 0.5, 1, 2, 4, 12, 24, 30, 60, and 120 seconds.

The windows with widths between 0.5 and 4 seconds were considered because some facial macro-expressions have durations that fall within these time periods (Adegun & Vadapalli, 2020; Ekman, 2003). This kind of emotional macro-expressions occur in daily interactions, and they are evident at first sight. The windows between 12 and 120 seconds were included because the duration of actions and events in inquiry-based science activities is within this range (Lämsä et al., 2018).

To interpret the emotional and behavioral signals from an educational perspective, the actions carried out by the student over time were manually and independently tabulated by two of the researchers. Later, we worked together to reach an agreement on the final tables that best capture the behavior of the student. By combining the tabulated actions with the signals at the appropriate time scales, we could associate educational, emotional, and behavioral events. The aim of this approach was therefore to bridge the affective, behavioral, and educational spheres.

## RESULTS AND ANALYSIS

The analysis of the entire process implied carrying out different studies in parallel and linking the successive results. For clarity, we show (see Figure 2) the
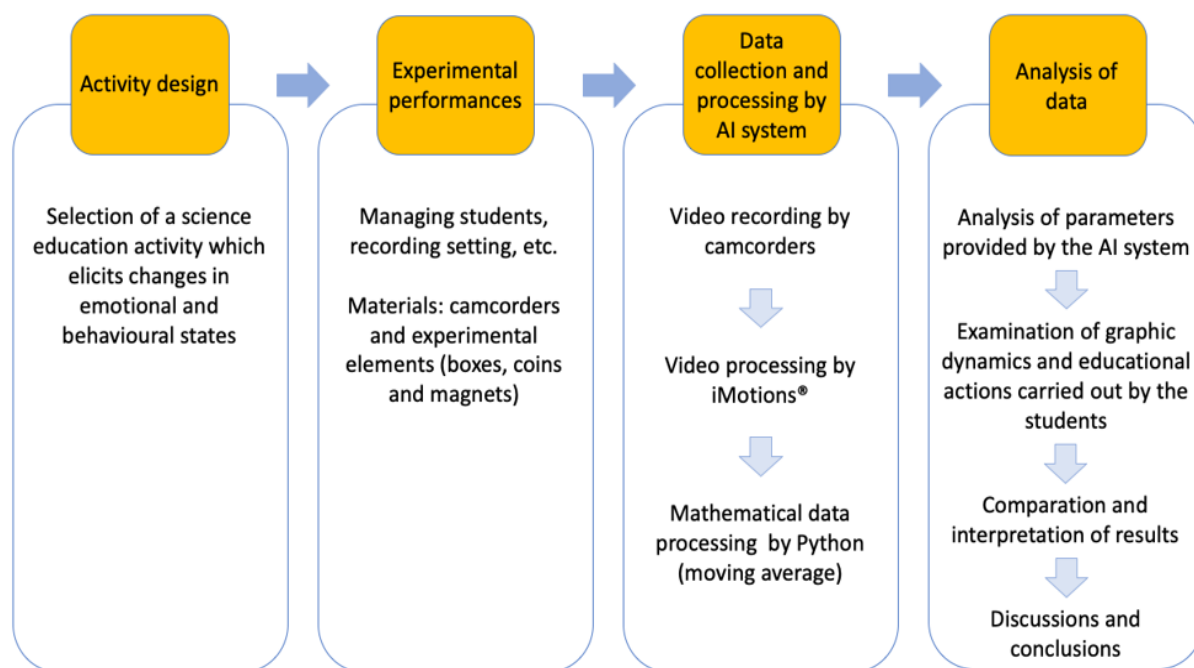
**Figure 2.** Flow chart to integrate AI into research on emotions and behaviors in science education

studies sequentially, although sometimes we will refer to elements that are shown later or further back.

**Tabulating Educational Actions**

The actions of the student were tabulated manually by the researchers. In Table 1, the second column contains a summary of the actions observed during the activity. This analysis is consistent with the content analysis applied in research in science education, in particular with the analysis of inquiry-based activities (Crujeiras & Jiménez-Aleixandre, 2019; Lämsä et al., 2018).

Based on the dynamics of the educational (Table 1), behavioral, and emotional events, we were able to define four phases which the student followed during the inquiry-based activity, which are coherent with the main features and practices involved in this educational approach (Jiménez-Liso et al., 2021a, 2021b). For

Table 1. Summary of observed relevant actions and proposed inquiry phases

| Time (min.) | Actions | Inquiry phases | |
|---|---|---|---|
| 0:00-3:00 | The student observes and lifts the box. They open the envelope and take the coins and magnets out. They shake the box and the envelope (with some coins inside) a few times. They laugh nervously. They bring the magnets close to the box and slide them over it. They observe and shake the envelope and the box a few times, while they listen to the resulting sounds. They move the magnets close to different coins. They slide a magnet over the box. They shake the box and the envelope (with some previously selected coins). They slowly turn the box upside-down. | Phase 1a-Non-systematic observation | Phase 1-Accommodation to the problem |
| 3:00-4:15 | They pick up and look at the envelope. They shake it. They look at the coins and magnets, and seem to ponder. They slowly turn the box upside-down. They bring a magnet towards different points around the box. While holding the magnet in one position, they shake the box. They bring a second magnet to the opposite side of the box, turn it over and seem to ponder. | Phase 1b-Pre-systematic observation & reflection | |
| 4:15-8:15 | They move a magnet towards and away from the box a number of times. They get the envelope and take some coins out. They shake the envelope and the box a few times. They get the box and observe it. They turn it over. They take a magnet and hold it against one side of the box. Then, they turn the box over a few times. They laugh. They explore the coins on the table with a magnet. They move a magnet towards and away from the box several times. They shake the box (without any magnets). While holding the box in their hands, they think and look at the items on the table. They simultaneously move close or away one, two or three magnets to different parts of the box. They take a coin out of the envelope and check whether it is attracted by the magnets. They move a magnet close to one side of the box. While keeping one magnet in a fixed position, they change the angle of the box. | Phase 2-Exploration of problem-solving strategies | |

**Table 1 (Continued).** Summary of observed relevant actions and proposed inquiry phases

| Time (min.) | Actions | Inquiry phases |
|---|---|---|
| 8:15-19:00 | They gradually change the tilt of the box while keeping one magnet in a fixed position. At a given angle, they move the magnet away. They change the tilt of the box several times without magnets. They check at what angles the coins start to slide. They carry out several tests in which they gradually modify the angle of the box, while keeping a magnet fixed. They stop and seem to ponder. They place a magnet on the lower part of the box and carry out some tests: they shake the box, turn it around, bring a magnet towards and away of the box, etc. They gradually modify the angle of the box, while keeping a magnet fixed. They design the next test. They change the tilt of the box until they set a given position. They then move the magnet away. They repeat the previous strategy several times. Sometimes, after setting an angle and moving the magnet away, they continue to increase the inclination of the box. They try a test in which they slide a few magnets over the bottom of the box, while the box is tilted. They perform several similar tests with two magnets arranged in an L-shape. As they reach a certain angle, they move one of the magnets away. | Phase 3-Implementation of problem-solving strategies |
| 19:00-the end | The student is warned that there is 1-minute left. They perform their actions at a faster pace. Some are repeated actions (2 magnets arranged in an L-shape) & there are new ones (3 magnets arranged in a U-shape). They nervously think of & write their conclusions. | Phase 4-Conclusion & communication |

convenience, this is displayed in Table 1 (third column), although it was definitely determined in subsequent analyses.

**Global Average Presence of Emotional & Behavioral Parameters**

To begin with the exploratory analysis, the global average presence of each emotional and behavioral signal throughout the whole activity was calculated (Table 2). The parameter attention had the highest average presence, 79.30%. The parameter engagement had an average presence of 45.45%, while the data for the signal of surprise was 16.60%. In contrast, the other parameters were considerably less frequent. Disgust, fear, joy, and contempt had a mean presence in the order of tenths. Meanwhile, anger and sadness showed values in the order of hundredths.

These results suggest that the student was very attentive and engaged, and they were also surprised during the activity. In the future, the application of this analysis will allow us to compare the average presence of each parameter in different individuals for a given activity, and between different types of activities and contents. Furthermore, the results of the global average presence (Table 2) informed the choice of window widths for smoothing the signals with the centered moving average. Thus, longer windows between 12 and 120 seconds were applied to the parameters with high presence (e.g., attention and engagement); in contrast, the only window applied to the parameters with low presence (joy, disgust, fear, contempt, sadness, and anger) was the shortest one (0.5 s). The parameters with medium presence (in this case, surprise) could be analyzed by both procedures.

**Table 2.** Global average presence of each parameter throughout the activity

| Parameter | Global average presence (%) |
|---|---|
| Attention | 79.30 |
| Engagement | 45.45 |
| Surprise | 16.60 |
| Joy | 0.61 |
| Disgust | 0.58 |
| Fear | 0.36 |
| Contempt | 0.29 |
| Sadness | 0.03 |
| Anger | 0.02 |

The parameters with medium and high presence offer information about the global dynamics of the activity. However, the high-presence parameters do not seem to be so clearly associated with specific actions carried out by the student. For this reason, here we first analyze surprise (medium presence) which seems to be associated both with global dynamics of the activity and specific actions carried out by the student.

**Study of Parameters with Medium Presence**

This analysis was carried out by applying a centered moving average of a width of four seconds. This time scale is close to the facial expressions of basic emotions (Adegun & Vadapalli, 2020; Ekman, 2003). In Figure 3, surprise shows several low intensity spikes of presence at the beginning of the activity. The analysis of the student's actions in this period (Table 1) indicated that the student appeared to play with the box. They made unsystematic observations through which they seemed to familiarize with the problem and the items. During this phase, the behavior of the box did not substantially surprise the student. Therefore, this initial period, which
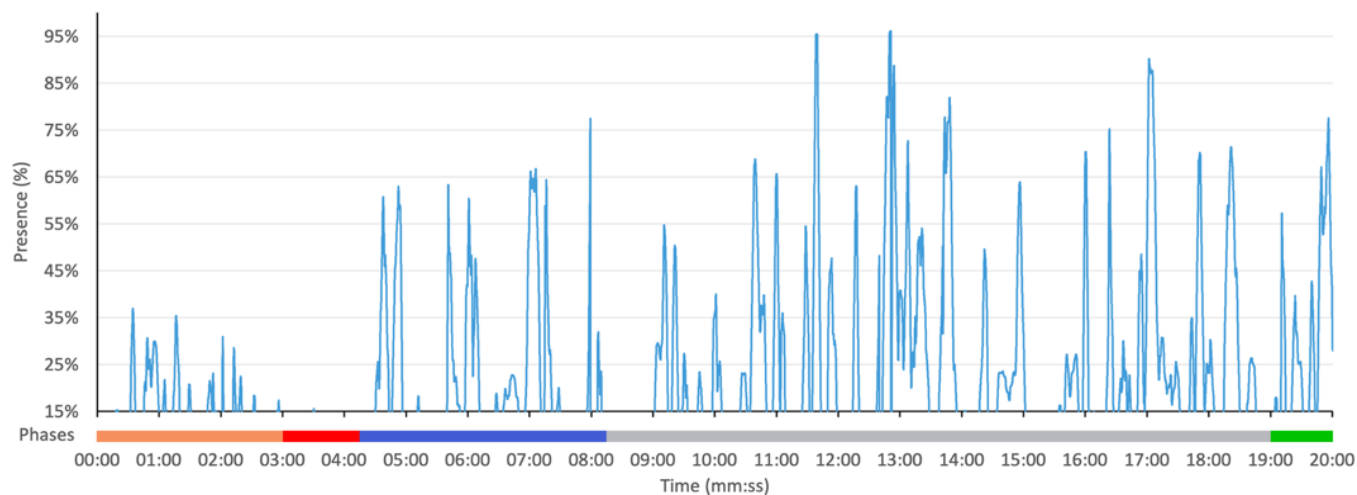
Figure 3. Surprise signal smoothed with a centered moving average with a width of four seconds. The Y-axis represents the percentage of presence (>15%), and the X-axis represents time. The colored band below the X-axis represents the phases proposed in Table 1: 1a/b (light/dark orange), 2 (blue), 3 (grey), and 4 (green)

was considered to be sub-phase 1a (from t≃0 min to t≃3:00 min), was labelled as *non-systematic observation*.

Then, we observed that the presence of Surprise became very low (<15%) (Figure 3). In this period, from t≃3:00 min to t≃4:15 min, the student performed small tests on the behavior of the supporting items and went through moments of reflection (Table 1). The tests on the behavior of the tools apparently confirmed that everything worked as expected, which did not surprise the student. This period, sub-phase 1b (from t≃3:00 min to t≃4:15 min), was called *pre-systematic observation and reflection*. Together with the previous sub-phase, both periods were considered as phase 1 (from t≃0 min to t≃4:15 min), and labelled as *accommodation to the problem.*

From t≃4:15 min to t≃8:15 min, the student presented sets of spikes of surprise separated by epochs of between 30 and 45 seconds during which this parameter had a low presence (Figure 3). In this period, the student performed actions to find possible problem-solving strategies. They carried out tests with the items available and they stopped every so often to observe the items and ponder (Table 1). When watching the video alongside Figure 3, we observed a potential correlation between the sets of surprise spikes and the successive tests carried out for solving the problem. This period, phase 2 (from t≃4:15 min to t≃8:15 min), was called *exploration of problem-solving strategies*.

The period from t≃8:15 min to t≃19:00 min had a continuous fluctuation of surprise. The student carried out systematic tests in which they aimed to find changes in the behavior of the box (Table 1). For example, they varied the angle until they perceived that the coins moved down. This process of systematic testing was repeated several times, but with different strategies. This repetition is reflected in the dynamics of surprise (Figure 3). The spikes seemed to occur after their predictions were not confirmed. This period, phase 3 (from t≃8:15

min to t≃19:00 min), was labelled as *implementation of problem-solving strategies*.

Finally, from the announcement of the end of the activity until its conclusion (from t≃19:00 min to t≃20 min), the level of surprise fell slightly. This period, phase 4 (from t≃19:00 min to t≃20 min), was labelled as *conclusion and communication*.

To sum up, it seems that the dynamics of surprise reflected the student's different actions. Interestingly, the student showed high levels of surprise once they had familiarized with the problem. Thus, the student appeared to firstly develop a framework based on the behavior of the items. After this, they implemented their models, observed the outcomes and compared these with their predictions. Then, surprise began to decline (less pronounced spikes) as the observed behaviors appeared to be in line with their conjectures. Thus, surprise seemed to be linked to the implementation of their models, rather than to the novelty or amusement of the activity. This result is in line with an up-to-date view of inquiry as an approach which conceives surprise (connected with data analysis) as a learning enhancer (Jiménez-Liso et al., 2021a; Rodríguez-Arteche & Martínez-Aznar, 2016).

**Study of Parameters with High Presence**

To analyze the parameters with high presence, we smoothed the signals with the moving average with 12-, 24-, 30-, 60-, and 120-second windows. These times scales suit better the duration of educational actions (Lämsä et al., 2018). Here, the signals after smoothing with the 24-second window are shown (Figure 4).

During phase 1, we observed that attention began at high levels of presence and decreased while fluctuating (from t≃0 min to t≃4:15 min). This pattern is consistent with the dynamics observed for surprise and with the
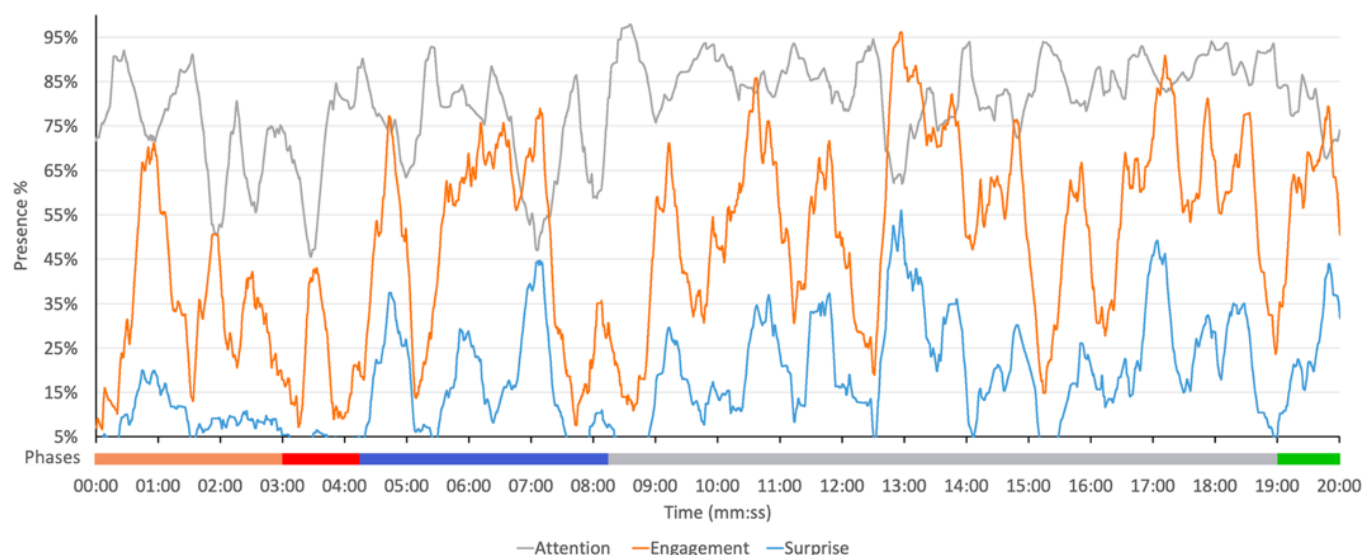
**Figure 4.** Attention (grey), engagement (orange), & surprise (blue) signals smoothed with a centered moving average with a width of 24 seconds. The Y-axis represents the percentage of presence (>5%), and the X-axis represents time. The colored band below the X-axis represents the phases proposed in Table 1: 1a/b (light/dark orange), 2 (blue), 3 (grey), & 4 (green)

actions performed by the student in sub-phases 1a and 1b.

During phase 2, attention decreased below 50% and increased above 90% while fluctuating (from t≃4:15 min to t≃8:15 min). In this phase, we could not find meaningful matches with other behaviors. From minute 8 onwards, right before the start of phase 3, attention began to stabilize at high levels.

It seems that during the first two phases the student was trying different strategies but could not find one that focused their attention. When they decided how to approach the problem (phase 3, from t≃8:15 min to t≃19:00 min), they began to carry out systematic tests that fixed their attention (over 80%, see Figure 4).

Moreover, the presence of engagement had large fluctuations during phase 1 (from t≃0 min to t≃4:15 min) and seemed out of synchrony with the other parameters. This pattern is consistent with the student's actions in the initial minutes (phase 1, Table 1). During this phase, the student appeared to get distracted and take breaks to ponder. Then, they started focusing on the box and the other items and tested problem-solving strategies (phase 2). From minute 8 onwards, right before the start of phase 3 (from t≃8:15 min to t≃19:00 min), the signals of engagement and surprise began to synchronize, and attention reached its highest levels of presence.

It seems that when the student familiarized with the problem and began applying appropriate strategies (phase 3), they remained attentive and engaged to check whether their predictions were met or not. Graphically, the signals of surprise and engagement increased as the student accommodated to the problem and showed synchrony with the tests and results obtained. This fact is consistent with the results obtained by Inkinen et al. (2020) and Jiménez-Liso et al. (2021a) on the greater engagement of students in the practices of using models and constructing explanations. These practices are preceded by an adequate accommodation to the problem.

Interestingly, there appeared to be lags of the order of half a minute between the behavior of some of the parameters provided by the AI system and the inquiry actions of the student. This means that the transition from one phase to another possibly had the time scale of problem-solving actions, which is about 30 seconds (Lämsä et al., 2018). Although this work is still preliminary, we could speculate that either the behavior of the student, the behavior of the items, or some of the parameters could act as triggers for the phase changes.

**Study of Parameters with Low Presence**

To study the parameters with low presence, the moving average with the 0.5-second window was calculated. This criterion is consistent with the duration of facial expressions associated with basic emotions (Adegun & Vadapalli, 2020; Ekman, 2003). Thus, the smoothed signals of the low presence parameters were expected to show short, high levels of presence (spikes) throughout the session.

In this analysis, an upward increase in the signal was considered a spike when the signal was above 50.0% of presence. Additionally, spikes were considered independent of each other when they were separated by Δt>4 s. Following this criterion, 69 peaks were obtained for the seven basic emotional parameters. After this, the videos were reviewed to identify the events associated with the spikes. We discarded those spikes which were triggered by external events (interaction with students and trainers, four peaks) or induced by partial occlusion of the face (two peaks). So, we considered 63 spikes: 53 (84.1%) of surprise, four of joy (6.3%), three of fear
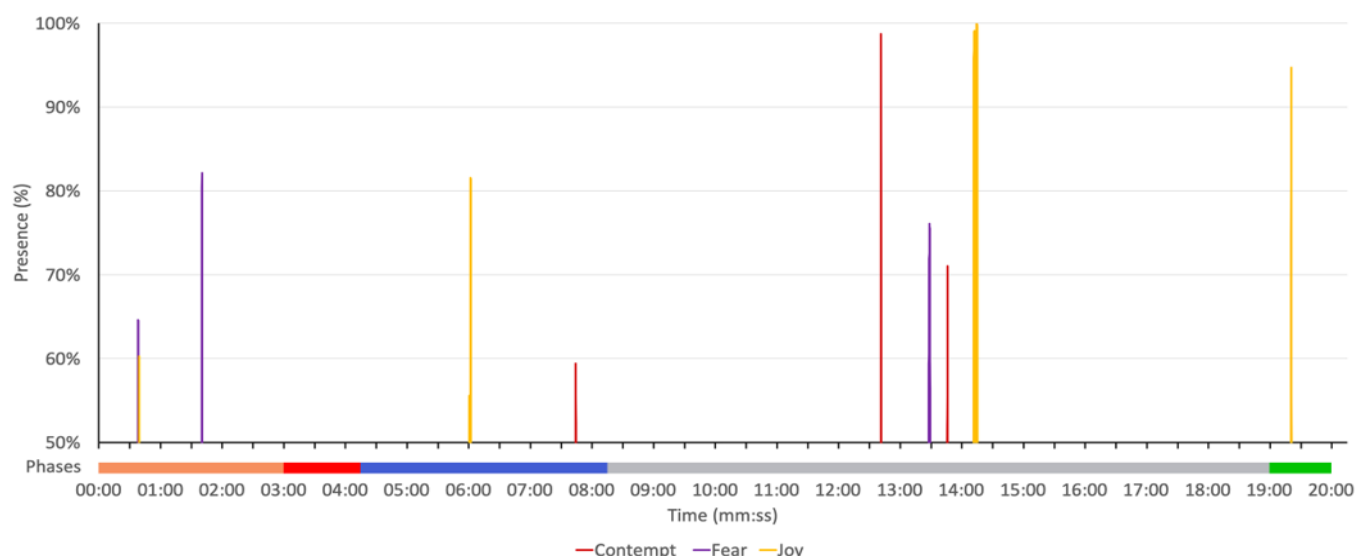
Figure 5. Spikes of the parameters with low presence: contempt (red), fear (purple), & joy (yellow). The signals were smoothed with a centered moving average with a width of 0.5 seconds. An upward increase was considered a spike when the signal was above 50% and separated by $\Delta t > 4$ s. The Y-axis represents the percentage of presence, and the X-axis represents time. The colored band below the X-axis represents the phases proposed in Table 1: 1a/b (light/dark orange), 2 (blue), 3 (grey), & 4 (green)

Table 3. Educational actions associated to the spikes of the parameters with low presence

| Emotion | Educational action | Time |
|---|---|---|
| Contempt | After some time changing the position of a magnet with respect to the box without apparent results, they change the strategy & start modifying the angle of inclination. | 07:44 |
| | They spend some time reflecting & looking at the box. After this, they seem uncertain & start a new test of changing the angle. | 12:41 |
| | The student faces a problem when increasing the angle of inclination of the box & has to stop & restart the test. | 13:46 |
| Fear | They shake the envelope with the example coins inside & then shake the box. They seem nervous, probably because they realize that some of the example coins might also be inside the box. | 00:38 |
| | They shake the envelope with example coins inside and then the black box. They seem to give meaning to these experiences and then feel nervous. | 01:40 |
| | The student carefully & nervously modifies the angle of inclination of the box while holding a magnet on one side. | 13:29 |
| Joy | They laugh after shaking the box & comparing this behavior to the envelope with coins. | 00:39 |
| | They happily look at the box after a small break of reflection. | 06:02 |
| | They laugh after finishing a test in which they shake the box with a magnet on one side & test the behavior of the coins inside. | 14:15 |
| | The student starts a new trial & laughs proudly. | 19:21 |

(4.8%), and three of contempt (4.8%). Figure 5 shows the spikes of the parameters with low presence (without surprise) *vs.* time.

To facilitate the interpretation of the basic emotions other than surprise, Table 3 shows the educational actions that were carried out at a given time. Contempt seemed to be related to the annoyance elicited when a test did not provide apparent results, or either to the uncertainty that followed periods of reflection. The actions associated to fear suggest that this parameter had a high presence during periods of excitement and agitation (e.g., at the beginning of the activity) and of careful handling of the materials. Finally, joy was associated to periods of appreciation of the activity (e.g.,

initial excitement) and of pride related to understanding or achievement of results (Bellocchi & Ritchie, 2015).

All in all, it seems that the study of parameters with low presence (spikes) does not contribute to the interpretation of the inquiry phases. However, it provides insights on the association between specific educational actions and the corresponding emotional responses.

## CONCLUSIONS

This paper focuses on showing a novel procedure to integrate artificial intelligence into research on emotions and behaviors during an inquiry-based science activity. In this regard, the study carried out presents some

relevant implications to improve the limitations of observational and declarative methods.

In particular, the use of automated facial expression recognition minimizes limitations such as the need to train learners and observers in the recognition and evaluation of emotions, or the difficulties in systematizing these processes (Barrett et al., 2019; Pekrun, 2006).

Importantly, the students' faces and upper bodies should always remain within the camera frame, since this allows collection of face expressions together with educational actions. However, the use of static cameras is a limitation in this study because students inevitably move out of frame at certain points.

The first analysis performed on the behavioral and emotional signals was a global average. The top-3 most present parameters were attention, engagement, and surprise for the dataset analyzed here. This is also observed when this preliminary analysis is conducted on all the datasets collected. We also found presence of the parameter joy in some of the datasets not included here. The results from the global average suggest that the student was very attentive and engaged, and they were also surprised during this activity. This kind of analysis provides a first-level assessment to compare individuals' behaviors for a given activity, and the effect of each type of activities and science contents.

To smooth the signals and extract relevant patterns, we developed an analysis pipeline. By using a centered moving average and different window widths, we were able to study the emotional and behavioral parameters at time scales relevant to educational actions. The results from the analyses with time windows between 0.5 and 4 seconds emphasized the connection between emotional or behavioral responses and specific educational actions. Thus, we identified the association between the spikes of surprise and the implementation of predictive models. In this sense, we could observe how the student firstly developed a framework about the behavior of the items, later implemented their models, observed the outcomes, and compared these with their predictions. In this process, the dynamics of surprise seemed to reflect their tests of their conjectures. Thus, surprise seemed to be linked to the validity of their models rather than to the novelty or amusement of the activity.

Moreover, the results from the analyses with wider time windows (e.g., 24 seconds) appeared to be more appropriate to study the inquiry phases. Thus, we identified four phases in the inquiry process: *accommodation to the problem*; *exploration of problem-solving strategies*; *implementation of problem-solving strategies*; and *conclusion and communication*. We also observed that the transition between phases was potentially triggered by external stimuli (e.g., box behavior), by the actions carried out by the student (e.g., completion of a test), or by the emotional state. The preliminary analysis

performed on all the collected datasets shows the existence of the phase 1 in the 12 students, but with different duration from person to person. Actually, two of them remained in phase 1 during all activity, and only four students were able to reach the level of conclusion and communication (phase 4).

Science education needs to further understand the role of the affective domain and provide teachers with evidence of how emotions may affect the cognitive development of students (Mellado et al., 2014). By developing and implementing methods such as the one described here, we can explore the flow of emotions in science education. The understanding of affective and neuroscientific factors should be integrated into teacher training (Ezquerra & Ezquerra-Romano, 2019). This would allow teachers to understand the emotions that students feel when they are learning science and develop more effective teaching activities.

# REFERENCES

Abbaschian, B. J., Sierra-Sosa, J., & Elmaghraby, A. (2021). Deep learning techniques for speech emotion recognition, from databases to models. *Sensors*, *21*(4), 1249. https://doi.org/10.3390/s21041249

Abd-El-Khalick, F., Boujaoude, S. Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., & Tuan, H-L. (2004). Inquiry in science education: International perspectives. *Science Education*, *88*(3), 397-419. https://doi.org/10.1002/sce.10118

Adegun, I. P., & Vadapalli, H. B. (2020). Facial micro-expression recognition: A machine learning approach. *Scientific African*, *8*, e00465. https://doi.org/10.1016/j.sciaf.2020.e00465

Artino, A. R., Holmboe, E. S., & Durning, S. J. (2012). Control-value theory: Using achievement emotions to improve understanding of motivation, learning, and performance in medical education: AMEE guide no. 64. *Medical Teacher*, *34*(3), e148-e160. https://doi.org/10.3109/0142159X.2012.651515

Azari, B., Westlin, C., Satpute, A., Hutchinson, J. B., Kragel, P. A., Hoemann, K., Khan, Z., Wormwood, J. B., Quigley, K. S., Erdogmus, D., Dy, J., Brooks, D.

H., & Barrett, L. F. (2020). Comparing supervised and unsupervised approaches to emotion categorization in the human brain, body, and subjective experience. *Scientific Reports*, *10*, 20284. https://doi.org/10.1038/s41598-020-77117-8

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1-68. https://doi.org/10.1177/1529100619832930

Bellocchi, A. (2019). Early career science teacher experiences of social bonds and emotion management. *Journal of Research in Science Teaching, 56*(3), 322-347. https://doi.org/10.1002/tea.21520

Bellocchi, A., & Ritchie, S. (2015). "I was proud of myself that I didn't give up and I did it": Experiences of pride and triumph in learning science. *Science Education*, *99*(4), 638-668. https://doi.org/10.1002/sce.21159

Borrachero, A. B., Brígido, M., Mellado, L., Costillo, E., & Mellado, V. (2014). Emotions in prospective secondary teachers when teaching science content, distinguishing by gender. *Research in Science & Technological Education*, *32*(2), 182-215. https://doi.org/10.1080/02635143.2014.909800

Chevrier, M., Muis, K. R., Trevors, G. J., Pekrun, R., & Sinatra, G. M. (2019). Exploring the antecedents and consequences of epistemic emotions. *Learning and Instruction, 63*, 101209. https://doi.org/10.1016/j.learninstruc.2019.05.006

Clark, E. A., Kessinger, J., Duncan, S. E., Bell, M. A., Lahne, J., Gallagher, D. L., & O'Keefe, S. F. (2020). The facial action coding system for characterization of human affective response to consumer product-based stimuli: A systematic review. *Frontiers in Psychology*, *11*, 920. https://doi.org/10.3389/fpsyg.2020.00920

Crujeiras, B., & Jiménez-Aleixandre, M. P. (2019). Students' progression in monitoring anomalous results obtained in inquiry-based laboratory tasks. *Research in Science Education, 49*(2), 243-264. https://doi.org/10.1007/s11165-017-9641-3

Dávila, M. A., Cañada, F., Sánchez-Martín, J., Airado, D., & Mellado, V. (2021). Emotional performance on physics and chemistry learning: The case of Spanish K-9 and K-10 students. *International Journal of Science Education, 43*(6), 823-846. https://doi.org/10.1080/09500693.2021.1889069

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*(3), 242-249. https://doi.org/10.1038/nrn1872

Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. Henry Holt and Company.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, *1*(1), 56-75. https://doi.org/10.1007/BF01115465

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press. https://doi.org/10.1037/t27734-000

Ezquerra, A., & Ezquerra-Romano, I. (2019). Using neuroscience evidence to train pre-service physics teachers on the concepts of heat and cold. *Journal of Physics: Conference Series*, *1287*, 012038. https://doi.org/10.1088/1742-6596/1287/1/012038

Fredricks, J. A. (2011). Engagement in school and out-of-school contexts: A multidimensional view of engagement. *Theory into Practice*, *50*(4), 327-335. https://doi.org/10.1080/00405841.2011.607401

Graesser, A. C. (2020). Emotions are the experiential glue of learning environments in the 21st century. *Learning and Instruction, 70*, 101212. https://doi.org/10.1016/j.learninstruc.2019.05.009

Haber-Schaim, U. et al. (1979). *Curso de introducción a las ciencias físicas* [*Introductory course in physical sciences*]. Reverté [Reverted].

Harley, J. M., Jarrell, A., & Lajoie, S. P. (2019). Emotion regulation tendencies, achievement emotions, and physiological arousal in a medical diagnostic reasoning simulation. *Instructional Science*, *47*(2), 151-180. https://doi.org/10.1007/s11251-018-09480-z

Ihme, K., Unni, A., Zhang, M., Rieger, J. W., & Jipp, M. (2018). Recognizing frustration of drivers from face video recordings and brain activation measurements with functional near-infrared spectroscopy. *Frontiers in Human Neuroscience*, *12*, 327. https://doi.org/10.3389/fnhum.2018.00327

Inkinen, J., Klager, C., Juuti, K., Schneider, B., Salmela-Aro, K., Krajcik, J., & Lavonen, J. (2020). High school students' situational engagement associated with scientific practices in designed science learning situations. *Science Education*, *104*(4), 667-692. https://doi.org/10.1002/sce.21570

Izard, C. E. (2009). Emotion theory and research: Highlights, unanswered questions, and emerging issues. *Annual Review of Psychology, 60*, 1-25. https://doi.org/10.1146/annurev.psych.60.110707.163539

Jarodzka, H., Skuballa, I., & Gruber, H. (2021). Eye-tracking in educational practice: Investigating visual perception underlying teaching and learning in the classroom. *Educational Psychology Review*,

*33*(1), 1-10. https://doi.org/10.1007/s10648-020-09565-7

Jeong, J. S., González-Gómez, D., & Cañada, F. (2016). Students' perceptions and emotions toward learning in a flipped general science classroom. *Journal of Science Education and Technology*, *25*(5), 747-758. https://doi.org/10.1007/s10956-016-9630-8

Jiménez-Liso, M. R., Bellocchi, A., Martínez-Chico, M., & López-Gay, R. (2021a). A model-based inquiry sequence as a heuristic to evaluate students' emotional, behavioural, and cognitive engagement. *Research in Science Education*. https://doi.org/10.1007/s11165-021-10010-0

Jiménez-Liso, M. R., Martínez-Chico, M., Avraamidou, L., & López-Gay, R. (2021b). Scientific practices in teacher education: The interplay of sense, sensors, and emotions. *Research in Science & Technological Education*, *39*(1), 44-67. https://doi.org/10.1080/02635143.2019.1647158

Lämsä, J., Hämäläinen, R., Koskinen, P., & Viiri, J. (2018). Visualising the temporal aspects of collaborative inquiry-based learning processes in technology enhanced physics learning. *International Journal of Science Education*, *40*(14), 1697-1717. https://doi.org/10.1080/09500693.2018.1506594

Laukenmann, M., Bleicher, M., Fuß, S., Gläser-Zikuda, M., Mayring, P., & von Rhöneck, C. (2003). An investigation of the influence of emotional factors on learning in physics instruction. *International Journal of Science Education*, *25*(4), 489-507. https://doi.org/10.1080/09500690210163233

Lederman, N., & Abd-El-Khalick, F. (1998). Avoiding de-natured science: Activities that promote understandings of the nature of science. In W. F. McComas (Ed.), *The nature of science in science education* (pp. 83-126). Springer. https://doi.org/10.1007/0-306-47215-5_5

Loderer, K., Pekrun, R., & Plass J. L. (2019). Affective foundations of game-based learning. In J. L. Plass, E. Richard, R. E. Mayer, & B. D. Homer (Eds.), *The handbook of game-based learning* (pp. 111-151). MIT Press.

Lombardi, D., & Sinatra, G. M. (2013). Emotions about teaching about human-induced climate change. *International Journal of Science Education*, *35*(1), 167-191. https://doi.org/10.1080/09500693.2012.738372

Marcos-Merino, J. M. (2019). Análisis de las relaciones emociones-aprendizaje de maestros en formación inicial con una práctica activa de biología [Analysis of the emotions-learning relationships of teachers in initial training with an active practice of biology]. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias* [*Eureka Magazine on Teaching and Dissemination of Sciences*], *16*(1), 1603. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2019.v16.i1.1603

Marcos-Merino, J. M., Esteban M. R., & Ochoa de Alda, J. A. G. (2021). Conocimiento previo, emociones y aprendizaje en una actividad experimental de ciencias [Prior knowledge, emotions and learning in an experimental science activity]. *Enseñanza de las Ciencias* [*Science Education*], 1-18. https://doi.org/10.5565/rev/ensciencias.3361

Mellado, V., Borrachero, A. B., Brígido, M., Melo, L. V., Dávila, M. A., Cañada, F., Conde, M. C., Costillo, E., Cubero, J., Esteban, R., Martínez-Borreguero, G., Ruiz, C., Sánchez-Martín, J., Garritz, A., Mellado, L., Vázquez-Bernal, B., Jiménez, R., & Bermejo, M. L. (2014). Las emociones en la enseñanza de las ciencias [Emotions in science teaching]. *Enseñanza de las Ciencias* [*Science Education*], *32*(3), 11-36. https://doi.org/10.5565/rev/ensciencias.1478

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction–What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474-496. https://doi.org/10.1002/tea.20347

Monkaresi, H., Bosch, N., Calvo, R. A., & D'Mello, S. K. (2017). Automated detection of engagement using video-based estimation of facial expressions and heart rate. *IEEE Transactions on Affective Computing*, *8*(1), 15-28. https://doi.org/10.1109/TAFFC.2016.2515084

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., & Linnenbrink-Garcia, L. (Eds.). (2014). *International handbook of emotions in education*. Routledge. https://doi.org/10.4324/9780203148211

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*(5), 1653-1670. https://doi.org/10.1111/cdev.12704

Putwain, D. W., Becker, S., Symes, W., & Pekrun, R. (2018). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learning and Instruction*, *54*, 73-81. https://doi.org/10.1016/j.learninstruc.2017.08.004

Rodríguez-Arteche, I., & Martínez-Aznar, M. M. (2016). Introducing inquiry-based methodologies during initial secondary education teacher training using an open-ended problem about chemical change. *Journal of Chemical Education*, *93*(9), 1528-1535. https://doi.org/10.1021/acs.jchemed.5b01037

Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., & Parrott, D. J. (2001). A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Nonverbal Behavior, 25*(3), 167-185. https://doi.org/10.1023/A:1010671109788

Sinatra, G. M., & Taasoobshirazi, G. (2018). The self-regulation of learning and conceptual change in science: Research, theory, and educational applications. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 153-165). Routledge. https://doi.org/10.4324/9781315697048-10

Sinatra, G. M., Broughton, S. H., & Lombardi, D. (2014). Emotions in science education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 415-436). Routledge.

Todd, R. M., Miskovic, V., Chikazoe, J., & Anderson, A. K. (2020). Emotional objectivity: Neural representations of emotions and their interaction with cognition. *Annual Review of Psychology*, *71*, 25-48. https://doi.org/10.1146/annurev-psych-010419-051044

Tóthová, M., & Rusek, M. (2021). The use of eye-tracking in science textbook analysis: A literature review. *Scientia in Educatione* [*Science in Education*], *12*(1), 1-12. https://doi.org/10.14712/18047106.1932

Vaessen, M., Van der Heijden, K., & de Gelder, B. (2019). Decoding of emotion expression in the face, body and voice reveals sensory modality specific representations. *bioRxiv*, 869578. https://doi.org/10.1101/869578

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, *57*(2), 137-154. https://doi.org/10.1023/B:VISI.0000013087.49260.fb

Wu, C.-L., Lin, T.-J., Chiou, G.-L., Lee, C.-Y., Luan, H., Tsai, M.-J., Potvin, P., & Tsai, C.-C. (2021). A systematic review of MRI neuroimaging for education research. *Frontiers in Psychology*, *12*, 617599. https://doi.org/10.3389/fpsyg.2021.617599

**https://www.ejmste.com**