Computers & Education 183 (2022) 104495

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology

Ilana Dubovi *

Nursing Department, The Stanley Steyer School of Health Professions, Sackler Faculty of Medicine, Tel-Aviv University, Israel

A R T I C L E I N F O

Keywords:
Eye-tracking
Facial expression
Electrodermal activity
Engagement
Virtual reality

A B S T R A C T

This study uses a multimodal data analysis approach to provide a more continuous and objective insight into how students' engagement unfolds and impacts learning achievements. In this study, 61 nursing students' learning processes with a virtual reality (VR)-based simulation were captured by psycho-physiological data streams of facial expression, eye-tracking, and electrodermal activity (EDA) sensors, as well as by subjective self-reports. Students' learning achievements were evaluated by a pre- and post-test content knowledge test. Overall, while both facial expression and self-report modalities revealed that students experienced significantly higher levels of positive than negative emotions, only the facial expression data channel was able to detect fluctuations in engagement during the different learning session phases. Findings point towards the VR procedural learning phase as a reengaging learning activity, which induces more facial expressions of joy and triggers a higher mental effort as measured by eye tracking and EDA metrics. Most importantly, a regression analysis demonstrated that the combination of modalities explained 51% of post-test knowledge achievements. Specifically, higher levels of prior knowledge and self-reported enthusiasm, and lower levels of angry facial expressions, blink rate, and devotion of visual fixations to irrelevant information, were associated with higher achievements. This study demonstrates that the methodology of using multimodal data channels encompassing different types of objective and subjective measures, can provide insights into a more holistic understanding of engagement in learning and learning achievements.

1. Introduction

Engagement is a multidimensional construct that refers to an individual's active involvement in a learning activity (Christenson, Reschly, & Wylie, 2012). Although educational research emphasizes the importance of engagement in the learning process and has demonstrated a robust link between engagement and achievement, more research is needed to make this complex construct more visible and measurable (D'Mello, Dieterle, & Duckworth, 2017; Ladd & Dinella, 2009). O'Brien, Roll, Kampen, and Davoudi (2021) suggest that to piece together the bigger picture of students' engagement, more holistic measures are needed to account for engagement as an ongoing cycle of engagement, disengagement, and re-engagement. Indeed, recent advances in the development of new data-capturing devices propose a multimodal approach of psycho-physiological signals and self-report data streams as a means of providing a person-oriented and continuous perspective on engagement. Such a perspective might provide a more comprehensive and objective insight into how engagement unfolds moment by moment and how it impacts learning (D'Mello et al., 2017). Combining

* Nursing Department, The Stanley Steyer School of Health Professions, Ramat Aviv, Tel Aviv University, Tel Aviv, 69978, Israel.
E-mail address: ilanadubovi@tauex.tau.ac.il.

https://doi.org/10.1016/j.compedu.2022.104495
Received 16 September 2021; Received in revised form 25 January 2022; Accepted 6 March 2022
Available online 12 March 2022

multiple measures of engagement will help us identify causes of disengagement, with the aim of further helping learners re-engage (Sharma & Giannakos, 2020).

Multimodal measures are now acknowledged as the basis for promising research in the field of multimodal data and learning analytics. For instance, Baceviciute and his colleagues (Baceviciute, Lucas, Terkildsen, & Makransky, 2021; Baceviciute, Terkildsen, & Makransky, 2021) incorporated eye-tracking and EEG in their investigation of cognitive engagement during learning in a virtual reality (VR) environment. Lee, Fischback, and Cain (2019) utilized an electrodermal activity (EDA) measure to examine the cognitive dimension of engagement during maker learning activities. Similarly, Taub, Sawyer, Lester, and Azevedo (2020) used video recordings of facial expressions to measure emotional engagement while students interacted with a game-based learning environment. Nevertheless, and despite the increased research in multimodal data for the learning field, a recent systematic literature review concluded that the "research in this direction has not reached its potential" (Sharma & Giannakos, 2020) and suggested additional research trajectories. Notably, Sharma and Giannakos (2020) claimed that engagement should be assessed not only as a unidimensional construct but rather as a multidimensional complex construct, and that multiple objective and subjective data measures could be used (Sinatra, Heddy, & Lombardi, 2015). In this context, the current study utilizes a multimodal approach to capture engagement as a multidimensional construct from a number of different standpoints in order to evaluate the impact on learning achievements.

1.1. *The construct of engagement*

Engagement captures behavioral, emotional, and cognitive manifestations of student learning (D'Mello et al., 2017; Fredricks, Blumenfeld, & Paris, 2004; Sinatra et al., 2015; Skinner, 2016). This paper explores only the emotional and cognitive aspects of engagement. The *emotional engagement* focuses on states related to students' emotional involvement during learning activities (Christenson et al., 2012). Positive emotions include enthusiasm, interest, and enjoyment while learning (Renninger & Bachrach, 2015) and negative emotional components include boredom, sadness, and frustration in the classroom (Skinner, 2016; Skinner, Furrer, Marchand, & Kindermann, 2008). Theories of motivation, including the self-determination theory and the control-value theory (CVT) of academic emotions (Deci & Ryan, 1985; Pekrun & Linnenbrink-Garcia, 2012), emphasize the role of both positive and negative emotions on students' involvement in learning activities, and underscore how affective dynamics can sustain or disrupt learners' engagement to impact learning performance (D'Mello & Graesser, 2012; Gupta, Elby, & Danielak, 2018; Pekrun & Perry, 2014).

According to the CVT, emotional experiences can be categorized by their valence (negative or positive), as well as by their activation (deactivating or activating). The CVT predicts that positive activating emotions (e.g., curiosity and enjoyment) promote learning outcomes, whereas the converse is true for negative-deactivating emotions (e.g., boredom) (Pekrun & Perry, 2014). Although the CVT represents an exemplary integrative approach to emotions and learning, concern has been raised regarding its generalizability to learning in technology-based learning environments. For example, contrary to the expectations, an evaluation of learning with Massive Open Online Courses (MOOCs) via machine learning techniques demonstrated that the expression of positive-activating emotions (e.g., enjoyment) and negative-deactivating emotions (e.g., anger) had no influence on student dropout (Xing, Tang, & Pei, 2019). In addition, a study with a virtual game-based environment suggested that negative emotions are more beneficial for long-term learning retention than positive emotions (Cheng, Huang, and Hsu (2020). However, a meta-analysis of emotions in technology-based learning studies published between 1965 and 2018 has provided evidence that face-to-face learning and learning with technology share similar emotional mechanisms (Loderer, Pekrun, & Lester, 2020). This controversy makes it clear that more research on emotional engagement with educational technology is needed.

The *cognitive dimension* of engagement is often defined as psychological investment. This encompasses students' mental orientation and cognitive efforts expended during learning activities, as well as the thoughts or focus aroused (Sinatra et al., 2015). Cognitive engagement measures are also considered to have self-regulation and motivation components (Ainley, 2012; Christenson et al., 2012; Fredricks et al., 2004) and have been found to affect a variety of positive outcomes, including motivation and learning achievements (Chi & Wylie, 2014; Greene, 2015; Guthrie et al., 2004).

Despite the significant advances in conceptualizing students' engagement, some important aspects of engagement still need to be tackled. In particular, until now most researchers have used self-reports or observations to measure engagement (D'Mello, 2017; Greene, 2015; Sinatra et al., 2015; Xie, Heddy, & Greene, 2019). Even studies employing automated continuous psycho-physiological measures are usually restricted to individual data streams because of the challenges related to multimodal data synchronization (Sharma & Giannakos, 2020). The current study was designed to address this issue and provide a multimodal approach to investigating the dynamic interplay of engagement during the learning process, by combining self-reports with automated psycho-physiological measurements.

1.2. *The multimodal data approach*

Multimodal data originate from different data channels which may be subjective and/or objective. Analyzing multiple signals and their mutual interdependence is expected to yield models that more accurately reflect the underlying nature of how engagement unfolds during a learning process (Sinatra et al., 2015). Incorporating traditional measures of engagement, such as self-reports, with fully automated measures of engagement, can capture the grain size of engagement dynamics and provide both microlevel and subjective macrolevel perspectives (D'Mello et al., 2017). Facial expression recognition, which is an example of a microlevel measure, is based on deep learning techniques that can match sets of action unit activities to different discrete emotions (Ekman, 1992; Ekman & Rosenberg, 1997). Moreover, it enables evaluation of changes in facial movements over very short time intervals and can therefore provide an unobtrusive but detailed view of the changes in students' emotions over the course of a learning session (Dindar, Jarvela, Ahola, Huang, & Zhao, 2020). Recent studies have demonstrated that prediction models of recognized emotional responses are highly accurate and show a strong agreement with self-reported data (Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015; Kulke, Feyerabend, & Schacht, 2020; Munshi et al., 2020). These models therefore represent a promising way of assessing affective engagement while learning (Ahn & Harley, 2020; Emerson, Cloude, Azevedo, & Lester, 2020). Importantly, while self-report measures predominantly reflect emotions of which the reporter is consciously aware, automatic facial expression analysis can capture fast changing emotions in the implicit dimension of the subconscious (He, Boesveldt, de Graaf, & de Wijk, 2016; Köster & Mojet, 2015; van Bommel, Stieger, Visalli, de Wijk, & Jager, 2020).

A visual attention data channel is another objective means of discerning cognitive engagement fluctuations while learning (D'Mello et al., 2017). Eye tracking technology enables monitoring of attention processes, namely, what, when, in what sequence, and for how long the learner responds to certain stimuli (Alemdag & Cagiltay, 2018; Miller, 2015). Eye tracking sensors can record three main types of visual attention measures: fixations, saccades, and blinks. Eye fixations, which describe the stable state of the eye at a single point, are theorized to indicate greater cognitive processing (Alemdag & Cagiltay, 2018). In particular, a longer mean fixation duration on a stimulus can indicate greater processing difficulty (Krejtz, Duchowski, Krejtz, Kopacz, & Chrząstowski-Wachtel, 2016; Kruger & Doherty, 2016; Liu & Chuang, 2011). In addition, eye-tracker metrics can reveal the decision making process. Since choice behavior and gaze allocation are apparently related, people tend to look longer at the item they will choose than at the item they will reject, thereby generating a gaze bias effect (Shimojo, Simion, Shimojo, & Scheier, 2003; Thomas, Molter, Krajbich, Heekeren, & Mohr, 2019). Another important eye-tracker metric is the blinking rate, which is thought to indicate mental load (Holland & Tarlow, 1972, 1975). Researchers have observed that blinking decreases during cognitive and memory processing, suggesting an inverse relationship between task difficulty and blinking; that is, increased difficulty lowers the blinking rate (Martins & Carvalho, 2015; Stern & Skelly, 1984).

An additional measurement of engagement employs an automated electrodermal activity (EDA) sensor (Andreassi, 2010). The mechanism for this is tied to the sympathetic nervous system, which in response to situations that comprise behavioral, cognitive, and affective phenomena immediately activates sweat glands near a person's hands and feet to prepare the body for action (Matsumoto, Walker, Walker, & Hughes, 1990). The measurable change in conductivity resulting from the influx of conductive liquid sweat has been interpreted as indicating engagement (Daily, James, Roy, & Darnell, 2015). Despite the use of EDA as an indicator of emotional engagement, recent studies suggest that electrodermal activity is not sensitive enough to changes in emotional states and may be more suitable for the measurement of cognitive engagement (Harley et al., 2015; Larmuseau, Vanneste, Cornelis, Desmet, & Depaepe, 2019; Lee et al., 2019; Parong & Mayer, 2021). Specifically, EDA measurements were used as an indicator of cognitive effort and high cognitive load during arithmetic and reading complex tasks (Armougum, Orriols, Gaston-Bellegarde, Joie-La Marle, & Piolino, 2019; Nourbakhsh, Wang, Chen, & Calvo, 2012).

The current study combines self-reports and psycho-physiological measures (facial expressions, eye-tracking metrics, and EDA) to assess the cognitive and emotional engagement flow while learners interact with a VR learning environment.

1.3. Learning with virtual reality

VR is a computer-generated environment that supports high-level synthetic and stimulating interactions in a 3D context. This provides an interactive experience of alternate reality that actively involves the learner in the learning process. While learning with VR, participants can move, sense, touch, and act upon simulated computer graphics, which in turn supports the perception that these objects truly exist (Ghanbarzadeh, Ghapanchi, Blumenstein, & Talaei-Khoei, 2014). VR simulations create a digital psychological sense of being immersed in a synthetic, computer-generated virtual world, as if it were a real-life place. This evokes a feeling of "being there", or as Witmer and Singer (1998) defined it, a 'sense of presence'. As such, these unique immersive characteristics can increase learners' affective arousal and boost their cognitive processing (Guan, Wang, Chen, Jin, & Hwang, 2021; Parong & Mayer, 2021). By building upon VR as a highly emotionally and cognitively stimulating learning environment, VR has been chosen for the present study to evaluate cognitive and emotional engagement processes.

VR environments can be classified as high or low-immersive. High immersive VR involves a high degree of interactivity obtained via peripheral devices, i.e., a head-mounted display in which a high graphical fidelity screen is mounted in front of the eyes with separate lenses for each eye, and sound is delivered through earphones. By replacing visual elements of the real-world environment with sensory stimuli that correspond with the virtual environment, this type of VR allows the user to immerse themselves in the virtual environment. In contrast, low-immersive VR, often called desktop VR, takes the form of a window into a virtual world displayed on a computer monitor, with interactions via a mouse, keyboard, or joystick (Choi et al., 2016). Importantly, recent studies report that immersive technology is inversely correlated with the level of learning, suggesting that high immersive VR is associated with more presence but less learning due to the increased cognitive load (Jensen & Konradsen, 2018; Makransky, Terkildsen, & Mayer, 2019; Ning Woon et al., 2020, p. 104655; Parong & Mayer, 2021). In consequence, the current study focused on learning with desktop low-immersive VR.

The benefits of VR learning have given rise to a number of educational theories and perspectives, and specifically the constructivist perspective, as well as situated and embodied learning theories and experiential learning approaches (Baceviciute, Terkildsen, & Makransky, 2021; Brown & Ford, 2002; Dede, Jacobson, & Richards, 2017; Fromm et al., 2021; Johnson-Glenberg, 2018; Lindgren, Tscholl, Wang, & Johnson, 2016). In this context, a recent meta-analysis revealed how widely VR technology has been adopted in the K-12 educational field (Yu, 2021). Interestingly, such technological advances have been even more widely adopted by medical and nursing schools and have shown clear learning benefits in teaching and learning anatomy (Huang, Liaw, & Lai, 2016), for surgical training (Li et al., 2017), for administration of medication and aseptic technique procedures, and in promoting interpersonal communications skills and decision making in a variety of clinical conditions (Chang & Lai, 2021; Dubovi, 2022; Harmon et al., 2021; Shorey & Ng, 2021; Woon et al., 2021). Cochrane systematic reviews concluded that VR training is at least as good as conventional training and should be implemented as an additional learning tool for medical students (Khan et al., 2019; Piromchai et al., 2015). However, to the best of the author's knowledge, this study is the first attempt to apply the multimodal data channel approach toward VR learning experiences in the nursing field.

1.4. *Research aims and questions*

Objective psycho-physiological measures for tracking engagement during learning, in addition to subjective self-reports, can inform us about what initiates and diminishes engagement, and how emotional and cognitive aspects are interrelated. Since research on multimodal measures and engagement is still in its infancy, the current study was designed to explore the interplay of nursing students' emotional and cognitive engagement while learning with a VR simulation. Specifically, the research questions relate to how students' cognitive and emotional engagement unfolds while they learn, as revealed by their self-reports and by psycho-physiological real-time measurements from facial expression analysis, eye tracking, and EDA sensors; and the nature of the synergistic effect of this cognitive and emotional engagement on students' learning achievements, as measured by multimodal data streams.

2. Methods

2.1. *Participants*

A total of 65 freshmen nursing students at an Israeli university volunteered to participate in the study. Due to technological issues of data collection (e.g., calibration errors), the data from 4 participants were excluded, resulting in a final sample size of 61. Most participants were female (n = 44) and their mean age was 23 ± 5.1 years.

The study was conducted following the approval of the university's ethics committee (#0001776–2).

2.2. *Research design and procedure*

This was a prospective study with a pre-test – post-test design. The students' demographics and prior content knowledge were obtained by a pre-test paper-and-pencil survey. Each participant was then asked to sit in front of a computer and their eye tracking and facial expressions were calibrated, followed by a learning session with a VR-based simulation lasting for a mean time of 41.18 (±13.7) minutes. Psycho-physiological cognitive and emotional engagement characteristics were captured during the entire learning experience using the eye tracker, EDA wearable wristband sensor, and facial expressions, which were extracted from webcam video recordings. In addition, students were asked by the VR simulation to complete three Positive and Negative Affect Scale (PANAS) questionnaires (Watson, Clark, & Tellegen, 1988) at Time 1, before the intro phase; Time 2, after the intro phase; and Time 3, after the last phase of the summary video.

After the learning session, participants completed a paper-and-pencil content knowledge post-test and the Presence questionnaire. The iMotions 9.0 Biometric Research Platform (https://imotions.com) was used to perform the experimental design and to collect the data from the multiple physiological measurements.

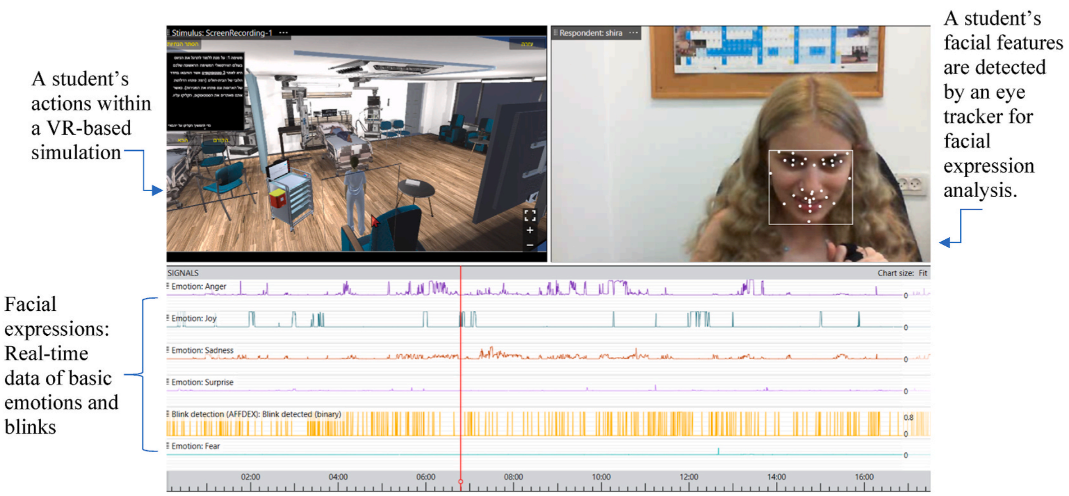


Fig. 1. A screenshot of the multimodel data collection while a student is experiencing the VR based simulation.

2.3. VR-based simulation

The VR simulation was an extension of a previous version, developed for this study using the Unity platform (https://unity.com) ((Dubovi et al., 2017) and presented as a 3D hospital. The simulation sequence followed six phases: intro phase, patient assessment; tutorial about the procedure; application of the procedure; follow-up tutorial about the procedure; and summary video about the procedure. During the introduction, students playing the role of nurse avatar were asked to explore the virtual hospital by moving around, opening cupboards, interacting with medical equipment, and reading charts. After this initial phase of acclimation, the students were asked to collect clinical data about the patient and to discern the rationale for the patient's hospitalization and treatment plan. The participants were then introduced to an embedded tutorial about medication administration procedures, followed by the application phase during which the participants were asked to apply their knowledge by conducting basic medication procedures. This involved choosing the correct medication from a selection of look-alike and sound-alike medications that are commonly confused. This could be achieved by correctly following the medication administration procedure learned during the previous phase. The two final phases involved a follow-up tutorial about the procedure, which emphasized dosage calculations, and a summary video about the procedure. The learning process was supported by textual prompts in order to provide instructional support, as well as performance feedback (Fig. 1).

2.4. Data collection instruments

A summary of the instruments and metrics used in this study are presented in Table 1.

2.4.1. Emotional engagement

2.4.1.1. Facial expression recognition. Facial expressions were collected throughout the learning experience. For this purpose, the Affectiva Affdex algorithm provided by iMotions 9.0 (https://imotions.com) was used to obtain real-time data for the 7 basic emotion likelihoods for joy, anger, surprise, contempt, fear, sadness, and disgust, at a 30 Hz frequency. The algorithm uses the Facial Action Coding System to identify and categorize facial expressions based on specific facial 'action units' and was shown to be reliable (Ekman & Rosenberg, 1997; Kulke et al., 2020). The metrics can be thought of in terms of probability; as the emotion or facial expression occurs

Table 1
Detailed description and definition of study variables by the type of modality.

	Modality & Variable	Type of data	Metrics
Emotional engagement	*Facial expressions* Emotion classification	Continuous, objective	Seven emotional basic states were automatically extracted using the Affectiva Affdex algorithm: anger, sadness, disgust, joy, surprise, fear, and contempt. Following data extraction and thresholding the following metric was used: • *Time percentage* – A time percentage metric was calculated for each emotion, namely, the amount of time that each emotion was displayed out of the total time participants were learning with VR.
	Self-report	Taken three times during learning, subjective	Self-reported PANAS questionnaire of 10 positive emotions and 10 negative emotions. The mean score was calculated for each emotion as well as for the overall positive and negative domains.
Cognitive engagement	*Eye tracking* Gaze metrics	Continuous, objective	*Gaze points are the raw* data-points collected by an eye tracker, and provide a continuous estimation of where a respondent's gaze is currently located. • *Time spent* - The amount of time that participants spent gazing at a particular AOI in relation to the time during which the AOI was active. The unit of measure is percentage.
	Fixations		Fixation is a period during which the eyes are locked on a specific object, when the eyeball remains relatively still. • *Time to first fixation* – Average time from the AOI appearance until the first fixation on the AOI is detected. Measured in milliseconds. • *Fixation counts*- Average number of fixations on a certain AOI • *Fixation dwells* - The mean duration of fixations on the AOI in relation to the time during which the AOI is active. The unit of measure is percentage.
	Saccades		Saccades are quick eye movements between fixations. • Saccade counts- Average number of saccades detected inside the AOI.
	Blinks		• Blink rate - Blinks per minute.
	EDA	Continuous, objective	EDA picks – number of phasic fast alterations per minute.
Presence	*Self-report*	Post-test, subjective	Descriptive statistics (Mean, SD) were calculated for the three presence questionnaire subscales.
Knowledge	*Self-report*	Pre and post learning, subjective	MAT was calculated as the Mean (SD) percentage of correct answers.

and intensifies, the score rises from 0 (no expression) to 100 (expression fully present). A thresholding for time and for absolute amplitude probability, based on the iMotions guidelines (https://imotions.com/blog/facial-expression-analysis/), was used to extract emotion percentage metrics.

2.4.1.2. Self-report on the Positive and Negative Affect Scale (PANAS). The PANAS provides a list of 20 adjectives used to describe 10 positive emotions and 10 negative emotions (Watson et al., 1988). The positive aspects relate to feelings of being attentive, active, alert, excited, enthusiastic, determined, inspired, proud, interested, and strong, while the negative aspects involve being hostile, irritable, ashamed, guilty, distressed, upset, scared, afraid, jittery, and nervous. The order of questions in the questionnaire mixed positive and negative descriptions. Respondents are required to indicate the extent to which they feel the emotion “at this moment” on a 5-point scale. Cronbach’s alpha indicated a good internal consistency score of 0.86 for the global positive affect and 0.82 for the global negative affect.

2.4.2. Cognitive engagement

2.4.2.1. Eye tracking. Smart Eye Aurora eye tracking hardware collected real-time gaze data at 60 Hz. Fixations and saccades for three dynamic areas of interest (AOIs; Fig. 2b) were collected and classified using the Duration Dispersion Filter (Holmqvist et al., 2011) and were then post processed using the Moving tool provided by iMotions to animate the AOIs in relation to the participant reference (Friedrich, Rußwinkel, & Möhlenbrink, 2017; Orquin, Ashby, & Clarke, 2016). The three pre-defined AOIs captured the students’ visual attention while interacting with the VR simulation procedure application phase (when they were required to select the correct medication). The first AOI covered the correct Aspart Insulin medication; the second Aspart Insulin that had passed the expiration date (i.e., the wrong medication); and the third Humalog Mix Insulin, a medication with the wrong trade and generic names.

The metrics extracted were time to first fixation (TTFF), percentage of time spent, fixation counts, percentage of fixation dwells, and saccade counts (for detailed definitions see Table 1). Furthermore, the blink rates during learning with the VR were captured to provide information about the overall learning experience.

2.4.2.2. EDA. The EDA signal was continuously recorded by the Shimmer 3 wristband to calculate the skin conductance level (also known as tonic level), as well as a fast-changing component, often referred to as the phasic response or skin conductance response; SCR. A smoothing filter was applied in order to extract the faster emotional related changes in the SCR data from the underlying tonic signal (Braithwaite, Watson, Jones, & Rowe, 2013). This was followed by the application of a peak detection algorithm and artifact rejection filter in order to calculate the number of EDA peaks per minute (Benedek & Kaernbach, 2010).

2.4.3.3. Presence questionnaire. The Presence Questionnaire PQ was developed by Witmer and Singer (1998) to measure the degree to which a direct interaction with a VR environment triggers participants to feel that the simulation sensations are real. The instrument addresses three subscales: (1) Involved/Comparison (11 items), which refers to the degree of VR responsiveness to the initiated actions; (2) Natural (3 items), which refers to the degree to which the interactions with the environment feels natural; and (3) Interface Quality (3 items), which refers to the degree of distraction caused by the quality of the interface. The PQ was completed immediately after learning with the VR simulation. The overall internal consistency in the current study was $\alpha = 0.88$, which is similar to the findings of previous reports (Witmer & Singer, 1998).

2.4.4. Content knowledge test

The Medication Administration Test (MAT) assessed the nursing students’ understanding of the practical applications of the

a. b.

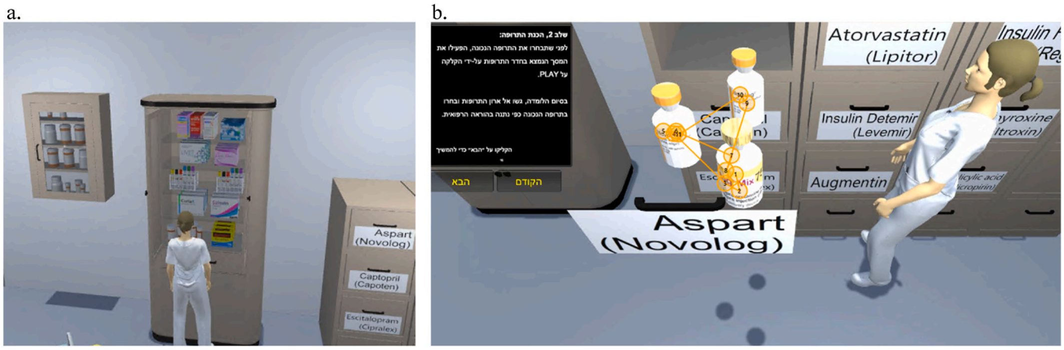


Fig. 2. A screenshot of a student’s avatar during the application phase of the medication administration procedure.
a. A student looking for the correct medication; b. The medications represent the three areas of interest designated for visual attention analysis. The numbered yellow points represent the order of visual attention as depicted by the eye-tracker. Namely, when and in what order the student directed their gaze while looking for the correct medication. . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

medication administration guidelines. The test, which was validated in a previous study (Dubovi et al., 2017), comprises 13 multiple-choice items. For example, one of the questions asks students whether two different medications in similar packages but with different generic names are the same medication. Another example is a question regarding what guideline should be applied if a patient accidently drops the cup with their pills on the floor while taking their medication. Pre- and post-test MAT evaluations were identical tests with shuffled questions, undertaken before learning with VR and immediately after. Analysis of the MAT results indicated a good internal consistency score, with a Cronbach's alpha of .74.

2.5. Data analysis

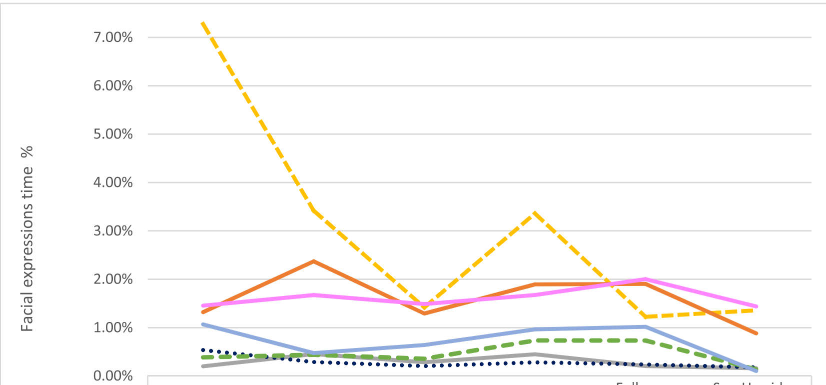
2.5.1. Emotional engagement

Because the facial expressions data did not have a normal distribution, a Kruskal-Wallis H test was performed for each of the seven emotions (joy, surprise, fear, contempt, disgust, sadness, and anger) to calculate the emotion distribution frequencies during the overall learning experience with VR. This was followed by a Linear Mixed Effects Model (LMM) for repeated measures in order to follow the emotional fluctuations of each emotion during the six different VR phases (an ordinal variable). The dependent variable was the time percentage of each of the seven emotions as extracted by facial expression analysis, and the independent variable was the VR phase (six phases), with the random effect being the individual and the VR phase. Slope and intercept were defined as random factors. The covariance structure was unstructured (Cobb, Jashami, & Hurwitz, 2021; Mancini, Biolcati, Agnoli, Andrei, & Trombini, 2018; McDuff, Kodra, Kaliouby, & LaFrance, 2017; Peng-Li, Mathiesen, Chan, Byrne, & Wang, 2021).

A LMM was established to estimate the self-reported changes in the PANAS self-report, administered three times during the learning process. The dependent variable was the self-reported emotion, and the independent variable was the time point at which the questionnaire was administered (an ordinal variable), with the random effect being the individual and the time point. Slope and intercept were defined as random factors. The covariance structure was unstructured.

2.5.2. Cognitive engagement

A one-way ANOVA analysis was performed to assess the eye-tracking metrics differences between the three AOIs (Aspart Insulin, the correct medication; Aspart Insulin, the wrong expired medication, and Humalog Mix Insulin, the wrong medication) during the procedural application VR phase. This was followed by LMM analysis to assess changes in blink rate during the learning process. The dependent variable was blink rate, and the independent variable was VR phase (six phases; an ordinal variable), with the random effect being the individual and the VR phase. Slope and intercept were defined as random factors. The covariance structure was unstructured. LMM analysis was also used to estimate the fluctuation of EDA picks. The dependent variable was EDA pick, and the independent variable was VR phase (six phases; an ordinal variable), with the random effect being the individual and the VT phase. Slope and intercept were defined as random factors. The covariance structure was unstructured.



	Intro phase	Patient assessment	Tutorial about the procedure	Application of the procedure	Followup tutorial about the procedure	SumUp video of the procedure
Joy	7.28%	3.42%	1.42%	3.36%	1.22%	1.36%
Surprise	1.32%	2.37%	1.29%	1.89%	1.90%	0.88%
Fear	0.20%	0.45%	0.28%	0.45%	0.20%	0.16%
Contempt	1.45%	1.67%	1.49%	1.67%	2.00%	1.44%
Disgust	0.54%	0.29%	0.21%	0.28%	0.24%	0.18%
Sadness	0.39%	0.44%	0.36%	0.73%	0.73%	0.13%
Anger	1.07%	0.47%	0.64%	0.96%	1.02%	0.10%

Fig. 3. Changes in students' emotions according to the VR simulation sequence phase.
The figure shows the percentage amount of time the students displayed each emotion out of the total time recorded for each VR simulation phase.

2.5.3. The impact of emotional and cognitive engagement on learning achievements

The MAT content knowledge responses were coded as correct or incorrect, and the total test score was calculated as the percentage of correct answers. The pre- and post-test results were analyzed with descriptive statistics (mean, SD). A two-phase approach was utilized to estimate the contribution of cognitive and emotional engagement characteristics to post-test MAT content knowledge. First, bivariate correlation analysis was used to examine the associations between the independent variables (facial expressions, PANAS self-report, eye-tracking metrics, EDA, and presence questionnaire) and the dependent post-test knowledge variable. This was followed by a second phase, based on the results of the correlation analyses, with only variables found to be significantly associated with the dependent post-test knowledge variable selected for inclusion in the subsequent regression analysis (for a similar approach to regression see (Duffy, Lajoie, Pekrun, & Lachapelle, 2018, p. 101150; Greene, Seung, & Copeland, 2014). The first regression analysis model included only the pre-test knowledge scores in order to relate to the baseline knowledge effect and the emotional engagement variables. Subsequently, the cognitive engagement variables were also included in order to assess the integrative contribution of emotional and cognitive engagement characteristics.

Data were analyzed using SPSS (version 25, IBM Corporation, Armonk, NY).

3. Results

3.1. Emotional engagement

Emotional engagement throughout the overall learning experience was captured via facial expressions extraction and PANAS self-reports. The results of a Kruskal-Wallis H analysis of *facial expressions* over the overall learning process revealed that participants experienced significantly ($\chi2(6) = 314.710$, $p < 0.001$) more joy (4.0 ± 3.6%) than other emotions, such as surprise (1.7 ± 2.6%), fear (0.3 ± 0.9%), contempt (1.5 ± 2.5%), disgust (0.3 ± 0.6%), sadness (0.4 ± 1.2%), and anger (0.8 ± 1.6%). The LMM for repeated measures analysis was conducted to examine the fluctuation of facial expressions throughout the six different learning phases with VR. The results revealed a significant interaction between the type of emotions derived from the facial expressions analysis and the six VR simulation phases ($F (30, 420.699) = 7.545$, $p < 0.001$, goodness of fit: 2 RLL: 11440.794; AIC: 11482.794; BIC; 11605.149). Post-hoc analysis indicated significantly more joyful expressions than other emotions in the introduction ($p < 0.001$) and procedure application phases ($p < 0.001$; Fig. 3). Facial expressions of anger and surprise were significantly lower during the final summary video phase than in other phases ($p < 0.05$, Fig. 3).

While facial expression analysis was designed to capture the fluctuations in the students' implicit and fast changing emotions, the PANAS self-report required the students to rate the intensity of emotions experienced at three different time-points. In accordance with the facial expression analysis, the PANAS paired *t* analysis revealed that students experienced significantly higher levels of positive than negative emotions ($t = 27.610$, $p < 0.001$; 3.8 ± 1.1 vs. 1.3 ± 0.7, Table 2). As shown in Table 2, the most predominant emotions while interacting with VR were interest (4.5 ± 0.4), attentiveness (4.5 ± 0.5), and alertness (4.3 ± 0.5). The results of the LMM analysis indicated that the emotions did not change significantly over the three time points of the PANAS test ($F (38, 2239.068) = 0.964$, $p = 0.533$, Goodness of fit: 2 RLL: 9336.596; AIC: 9342.596; BIC; 9361.162).

Table 2
Descriptive analysis of the PANAS self-report (N = 61).

Valence	Emotions	Time 1	Time 2	Time 3
		Mean (±SD)	Mean (±SD)	Mean (±SD)
Positive affect	Interested	4.5 (±0.6)	4.5 (±0.6)	4.6 (±0.5)
	Strong	3.6 (±1.00)	3.4 (±1.2)	3.6 (±1.3)
	Excited	3.5 (±0.9)	3.6 (±1.1)	3.6 (±1.1)
	Enthusiastic	3.5 (±1.1)	3.8 (±1.0)	3.7 (±1.2)
	Alert	4.4 (±0.6)	4.4 (±0.6)	4.2 (±0.7)
	Proud	3.3 (±1.1)	3.4 (±1.2)	3.5 (±1.3)
	Inspired	3.3 (±1.1)	3.4 (±1.1)	3.6 (±1.3)
	Attentive	4.6 (±0.5)	4.6 (±0.6)	4.4 (±0.7)
	Active	3.3 (±1.1)	3.6 (±1.1)	3.6 (±1.2)
	Determined	3.6 (±1.0)	3.7 (±1.2)	3.6 (±1.3)
	Total Positive	3.8 (±0.5)	3.8 (±0.6)	3.8 (±0.7)
Negative affect	Nervous	1.2 (±0.4)	1.1 (±0.4)	1.1 (±0.4)
	Ashamed	1.9 (±1.0)	1.5 (±0.8)	1.5 (±0.8)
	Scared	1.3 (±0.7)	1.2 (±0.5)	1.3 (±0.7)
	Hostile	1.1 (±0.4)	1.2 (±0.6)	1.2 (±0.6)
	Afraid	1.5 (±0.7)	1.3 (±0.6)	1.6 (±1.0)
	Upset	1.1 (±0.5)	1.0 (±0.1)	1.1 (±0.4)
	Distressed	1.4 (±0.8)	1.3 (±0.6)	1.2 (±0.7)
	Irritable	1.4 (±0.8)	1.2 (±0.6)	1.4 (±0.9)
	Guilty	1.1 (±0.4)	1.1 (±0.4)	1.2 (±0.7)
	Jittery	1.6 (±0.7)	1.3 (±0.7)	1.5 (±0.8)
	Total Negative	1.3 (±0.4)	1.2 (±0.3)	1.3 (±0.5)

3.2. Cognitive engagement

Students' cognitive engagement while learning with the VR simulation was captured by eye-tracking metrics of fixations and saccades and by the blink rate. The LMM analysis results of the blink rate revealed significant changes in the number of blinks per minute in the different VR simulation phases (Table 3; F (5, 60.174) = 19.841, $p < 0.001$, Goodness of fit: 2 RLL: 2261.777; AIC: 2303.777; BIC; 2385.268). Post-hoc analysis demonstrated significantly fewer blinks during the application of the procedure phase compared to the other phases (14.09 ± 9.19; Table 3).

The fixations and saccades in the three AOIs were used to investigate the cognitive engagement characteristics for the procedure application phase (see section 2.4.2.1 and Table 1), when the students were asked to choose Aspart Insulin with the correct expiry date from among a selection of other medications (Fig. 2b). Table 4 demonstrates that on average students first fixated on the wrong Humalog Mix medication. They also spent significantly more time looking and fixating on medications with similar packaging, presumably looking for differences (Table 4).

EDA metrics were an additional measure used to account for both cognitive and emotional engagement. As shown in Table 3, significant changes in the EDA peaks per minute were detected by an LMM for repeated measures analysis across the different VR simulation phases (F (5, 53.932) = 2.517, $p < 0.05$, Goodness of fit: 2 RLL: 1383.825; AIC: 1425.825; BIC; 1505.670). Post-hoc analysis demonstrated significantly more EDA peaks per minute during the application of the procedure and patient assessment phases compared to the other phases (Table 3).

Significant correlations were demonstrated between cognitive and emotional engagement measures. Particularly, Table 5 shows that blink rate per minute was significantly associated with facial expressions of joy ($r = 0.465$, $p < 0.01$) and surprise ($r = 0.256$, $p < 0.05$). In addition, while no significant association was identified between EDA arousal and emotional engagement, there was a significant negative association of EDA peaks with blink rate ($r = -0.345$, $p < 0.05$; Table 5).

3.3. Presence

Sense of presence within the VR environment is an important parameter of the learning experience. The results indicated that the values for the three PQ subscales: Involved/Comparison (5.2 ± 0.6), Natural (4.7 ± 1.1), and Interface Quality (5.7 ± 0.9), were either higher or similar to those reported by Witmer and Singer (1998) in their work with immersive VR. An intercorrelation analysis revealed a significant correlation between sense of presence and self-reported PANAS (Supplementary data A.3). In particular, there were positive associations between presence and interest level ($r = 410$, $p < 0.01$) and attentiveness ($r = 377$, $p < 0.01$); and a negative association between presence and jitteriness ($r = -321$, $p < 0.05$). In addition, while no significant association was identified between sense of presence and emotional engagement, there was a significant negative association of presence with fixation on the correct AOI ($r = 0.408$, $p < 0.05$; Table 5).

3.4. MAT content knowledge, emotional and cognitive engagement

The post-test MAT content knowledge scores were significantly higher than the pre-test scores (paired $t = -6.422$, $p < 0.001$; 90 ± 9 vs.78 ± 16). A bivariate intercorrelation analysis was performed to examine the impact of emotional and cognitive engagement on knowledge gain (Table 5 and Supplementary data A.1, A.2). Only significant intercorrelations were selected for inclusion in the regression analysis. As Table 6 demonstrates, Model 1 incorporated pre-test knowledge scores and emotional engagement variables to reveal that prior knowledge, self-reported enthusiasm, and facial expression of anger could explain 39% of post-test knowledge, ($F(2, 55) = 6.44$, $p < 0.001$). While enthusiasm had a positive impact on knowledge, anger had a negative impact. In Model 2, the addition of cognitive engagement in terms of blink rate (per minute) and visual fixations on the wrong medication (%), explained an additional significant 12% of the variance in post-test content knowledge scores; and the entire model explained 51% of the variance in post-test scores ($F(2, 53) = 3.72$, $p < 0.05$). Both blink rate and fixation on the wrong AOIs had a negative impact on post-test knowledge. In addition, a bootstrap method with 5000 resamples was used to investigate the parallel mediation effect of self-reported enthusiasm on the association between post-test interest, attentiveness, determination, and knowledge scores (Hayes, 2012). The results indicated that the indirect effect of enthusiasm was significant for the association between post-test interest and knowledge ($b = 0.15$, $SE = 0.07$, 95% CI = 0.360 to 0.3021); post-test attentiveness and knowledge ($b = 0.14$, $SE = 0.08$, 95% CI = 0.0191 to 0.3360), and post-test determination and knowledge ($b = 0.17$, $SE = 0.09$, 95% CI = 0.0328 to 0.3969).

Table 3
Changes in the number of blinks per minute in the VR simulation sequence phases (N = 61).

VR simulation sequence phase	Blinks per minute	EDA per minute
	Mean (±SD)	Mean (±SD)
Introduction phase	18.55 (±10.38)	4.25 (±3.03)
Patient assessment	16.14 (±11.07)	4.99 (±2.68)
Tutorial about the procedure	18.88 (±10.38)	4.45 (±2.58)
Application of the procedure	14.09 (±9.19)	5.10 (±2.91)
Follow-up tutorial about the procedure	17.04 (±10.76)	4.47 (±3.07)
Summary video of the procedure	19.45 (±12.94)	4.11 (±2.57)

Table 4
One-way ANOVA for visual attention measures by three different Areas of interest (N = 61).

Variable name	Areas of interest			F (2,120)	Post hoc
	Aspart Insulin (the correct medication)	Expired Aspart Insulin (the wrong medication)	Humalog Mix Insulin (the wrong medication)		
Time to First Fixation (ms)	3327 (±5494)	4734 (±7942)	1076 (±1597)	5.87*	Humalog (wrong) < Aspart (correct), Aspart (wrong)
Time spent (%)	22.5 (±14.2)	22.5 (±10.6)	14.5 (±9.6)	4.12*	Aspart (correct), Aspart (wrong) > Humalog (wrong)
Fixation counts	76.5(±64.2)	66.9 (±41.5)	55.9 (±46.3)	1.07	n.s.
Fixation dwells (%)	15.8 (±10.3)	16.8 (±10.2)	11.2 (±8.4)	4.40*	Aspart (correct), Aspart (wrong) > Humalog (wrong)
Saccade counts	150.4 (±184.8)	117.4 (±78.6)	78.7 (±59.0)	3.99*	Aspart (correct), Aspart (wrong) > Humalog (wrong)

4. Discussion

The main goal of this study was to provide a more holistic perspective on how emotional and cognitive engagement unfolds and synergistically impacts achievements. Analysis of emotional and cognitive engagement via multimodal metrics revealed that self-reported enthusiasm levels, facial expressions of anger, and visual attention in terms of blink rate and fixations explained 51% of post-test learning achievement. By integrating data from different modalities, automatic sensors, and self-reported components of engagement, the current study indicates methodological and instructional implications (Emerson et al., 2020).

4.1. *Emotional engagement*

Overall, the integration of a granular, continuous data-source of emotional engagement using automatic recognition of students' facial expressions together with self-reports, emphasizes the important impact of positively and negatively activating emotions on learning. More specifically, the facial expression continuum revealed that joy (i.e., enjoyment) was the most predominant emotion expressed by students. Interestingly, the frequency of the joy expression changed across the different VR learning phases. Specifically, the facial expression of joy was higher at the beginning of the VR learning session, then decreased, and increased again during the activity that required active application of a procedure. This finding that emotions change according to the stage of the learning session is consistent with the most recent study by Tonguç and Ozkara (2020), who evaluated students' facial expressions during a lecture and reported that the instructor's activities impacted students' expressions of joy. Interestingly, while the increase in the joy expression during the initial learning phase in the current study can be explained by the novelty of the VR simulation, the significant increase in joy during the application procedure phase can be attributed to the additive value of learning with the VR medium. In addition to declarative learning, VR also facilitates step-by-step skill procedural learning, which if experienced as successful can trigger the outcome related emotion of joy (Dubovi, 2022; Dubovi, Levy, & Dagan, 2017; Pekrun, 2006).

Despite the observation that facial expressions of joy were the most frequent emotions expressed by students, there was no significant impact of joy on post-test learning achievements. In contrast, although facial expressions of anger were expressed at a relatively low frequency with no significant changes across the different phases of the learning session, there was a significant association with post-test learning achievements. Specifically, more frequent anger expressions were associated with lower post-test scores. These findings are in line with previous research suggesting that facial expressions are more suitable for detecting and characterizing negative valence, because negative facial expressions are quicker to appear (van Bommel et al., 2020). This finding is also supported by the CVT framework, which categorizes anger as a negative activating emotion that is expected to undermine interest and intrinsic motivation and thereby hinder learning (Camacho-Morles et al., 2021; Pekrun & Perry, 2014). It is important to note that there is an ongoing debate regarding the effect of anger on achievement (Ahn & Harley, 2020; Pekrun, Elliot, & Maier, 2009).

The PANAS self-report was incorporated in order to capture the subjective perception of the emotional experience while learning. The results indicated that despite the great fluctuations in facial expressions during the learning experience, there was no significant change in the three PANAS self-reported questionnaires completed at different stages of the learning process. This finding highlights the additive value of a continuous automated measure of facial expressions, which appears more sensitive to implicit emotional fluctuations (He et al., 2016; Köster & Mojet, 2015). Indeed, previous studies have documented the fact that self-reports predominantly reveal only the emotions of which the responder is aware, while complex emotion processes in other subsystems remain hidden (Köster & Mojet, 2015; Scherer, 2009). Another study concluded that the results of self-reports are not directly comparable to facial expression analysis and indeed show little overlap because of differences in intensities and durations (van Bommel et al. (2020). The current study findings illustrate the advantages of a multimodal approach that is able to depict a more holistic perspective on engagement and learning. Interestingly, the findings also demonstrated that the impact of self-reported interest, attentiveness, and determination on learning achievements was mediated by enthusiasm. Enthusiasm has been defined as "energetic interest" that, once stimulated by a particular activity, provides excitement to perform (Setianingsih & Nafisah, 2021). Moreover, Renninger and Hidi (2016) defined interest as a dual state that comprises both cognitive and affective motivational predispositions of engagement. The findings are further supported by the CVT, which suggests that personal interest in the activity domain can give rise to appraisals of controllability and value, which lead to the promotion of achievement emotions and have a positive impact on performance.

Table 5
Bivariate intercorrelations between MAT content knowledge post-test and emotional engagement, cognitive engagement, and presence (N= 61).

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Content post-test knowledge	–																
Emotional Engagement:																	
Facial expressions																	
2. Joy	-.116	–															
3. Surprise	-.048	.008	–														
4. Fear	.135	-.031	**.294***	–													
5. Contempt	.017	**.265***	.090	-.048	–												
6. Disgust	.088	.090	.080	-.027	**.545****	–											
7. Sadness	-.122	-.025	.084	-.075	-.057	.048	–										
8. Anger	**-.323***	-.119	-.095	-.033	-.018	.119	.114	–									
PANAS self-report (Positive/Negative)																	
9. Positive	**.294***	-.003	-.106	.106	0.190	**-.267***	-.06	-.104	–								
10. Negative	-.111	.106	-.055	-.054	.026	-.111	.081	-.085	-.044	–							
11. Presence	.068	.081	-.229	.192	.145	.035	.006	.021	.229	-.13	–						
Cognitive engagement:																	
12. Blink rate per minute	**-.264***	**.465****	**.256***	-.010	-.003	.028	-.005	.208	-.073	-.141	-.099	–					
13. Fixation dwells on the wrong choice (Humolog Mix)	-.026	-.199	-.123	.188	-.007	-.321	-.077	.125	.061	-.017	.308	-.155	–				
14. Fixation dwells on the correct choice (Novolog)	.123	.041	-.309	-.118	.204	-.096	-.112	-.031	.239	-.029	**.408***	.095	.018	–			
15. Fixation dwells on the wrong choice (Novolog)	**-.417****	-.316	-.226	-.089	-.284	-.315	-.228	.142	-.244	-.014	.182	-.144	.093	.212	–		
16. EDA per minute	.121	-.019	.051	-.059	.033	-.01	-.06	.154	.082	-.096	-.002	**-.345***	-.203	.29	.126	–	
17. Content pre-test knowledge	**.489*****	-.102	.072	-.051	.012	.156	-.020	-.151	.104	.082	**-.259***	**-.279***	.026	-.007	-.033	.114	–

Data represents the Pearson r or Spearman r_s.
For a more detailed bivariate intercorrelation analysis please refer to Supplementary data A.1 – A.2* $p < 0.05$; **$p < 0.01$.

Table 6
Regression analysis for variables explaining MAT content knowledge post-test ($N = 61$).

Engagement type	Measures	Variable	Model 1		Model 2	
			β	t	B	t
	Self-report	Knowledge pre-test	0.39	3.722***	0.35	3.301**
Emotional engagement	Self-report	Enthusiastic	0.34	2.661*	0.31*	2.493*
		Interested	0.06	0.388	0.07	0.554
		Attentive	0.15	1.22	0.16	1.266
		Determined	0.05	0.391	0.03	0.256
	Facial expressions	Anger	−0.28*	−2.556*	−0.29*	−2.897*
Cognitive engagement	Eye tracking	Blinks rate			−0.234*	−2.054*
		Fixation dwells on the irrelevant choice			−0.27*	−2.602*
		R^2	0.39		0.51	
		F for change in R^2	6.44***		3.72*	
		ΔR^2	0.39		0.12	

*$p < 0.05$; ***$p < 0.001$.

This study proposes that sense of presence, provided by the illusion of being in a virtual place, is part of an emotional engagement construct. This is supported by a significant interaction between perceived presence and the interest and attentiveness levels in the PANAS self-reports and adds to the existing affective computing research, which shows that sense of presence is a precondition for emotions (Felnhofer et al., 2015; Marín-Morales et al., 2018).

4.2. Cognitive engagement

Another major contribution of this study stems from the cognitive engagement fluctuation obtained by eye movement metrics. First, the results demonstrated that blink rate is negatively associated with post-test performance. Previous studies have linked a lower blink rate to higher on-task mental effort (Chen, Epps, Ruiz, & Chen, 2011; Holland & Tarlow, 1972; Martins & Carvalho, 2015). This inverse relationship between cognitive effort and blink rate suggests that students who experienced the VR simulation as mentally difficult were more likely to achieve low post-test scores. Furthermore, a more precise analysis revealed that the blink rate changed over the course of the learning session, with a significantly lower blink rate recorded for the application procedure phase. This implies that procedural knowledge application to solve a complex problem of identifying the correct medication was experienced by the students as cognitively effortful. Indeed, the literature shows that knowing how to practice a skill requires more cognitive capacity than acquiring declarative knowledge (Anderson, 2015; Hong, Pi, & Yang, 2018).

Cognitive processes were also detected by monitoring the visual attention characteristics in terms of fixation counts and dwells. These eye movement parameters revealed that students devoted more visual attention to medications with similar packaging than to medications with a different package, which indicates a processing difficulty associated with distinguishing between two similar looking packages while applying the medication administration procedure regulations (Alemdag & Cagiltay, 2018). According to the gaze bias theory, when reaching decisions we often spend more time examining the options that we ultimately choose than those we do not choose (Schotter, Berry, McKenzie, & Rayner, 2010). Accordingly, the results indicate that students who spent more time on the irrelevant choice tended to score lower in the post-test. This finding is an important contribution since Alemdag and Cagiltay (2018), in their systematic review, concluded that there was a “lack of evidence” for any association between eye tracking metrics and learning performance.

Another autonomic data channel that exhibited an association with cognitive engagement was the EDA wristband sensor. Like the facial expression and eye tracking data streams, the EDA response was significantly increased during the application procedure phase, indicating maximal engagement. The observation of a significant correlation between EDA peak responses and blink rate but not emotional engagement accords with the suggestion by Lee et al. (2019) that the EDA response is less suited for capturing arousal due to affect than for detecting cognitive processing.

4.3. Interrelated relationship between emotional and cognitive engagement

Findings demonstrate some essential interrelated associations between the cognitive and emotional engagement characteristics. Positive emotions of joy and surprise were shown to be related to induced blink rate. As explained earlier, a higher blink rate is associated with higher on-task mental effort. This might suggest that positive emotions triggered a higher cognitive effort or vice versa; more sophisticated tasks triggered more positive emotions. While previous research describes emotions as a source of extraneous cognitive load (Plass & Kalyuga, 2019), further studies should investigate this interrelated association and its effect on achievements, especially within interactive virtual learning environments (Dubovi & Lee, 2019; Skulmowski & Xu, 2021). Another interesting finding points to an association between perceived sense of presence and cognitive processing in terms of visual attention. Specifically, higher levels of presence were related to more visual attention toward the relevant choice of medication. This finding is important as it provides evidence that the design of VR learning environments to induce sense of presence can impact the learning process (Dubovi et al., 2017; Witmer & Singer, 1998).

4.4. Research limitations

The current study has several limitations. The research participants were all undergraduate university students learning in a laboratory, which may limit generalization of the results, and future studies should include students on all levels. In addition, the analysis of facial expressions focused on composite emotion models, and it would be interesting if further studies could also evaluate individual facial action units, which have been suggested as predictors of learning gains (Liaw, Yu, Chou, & Chiu, 2021; Sawyer, Smith, Rowe, Azevedo, & Lester, 2017). In this context, there are various additional automated sensor-based measures (e.g., electroencephalography, electrocardiography, and electromyograms) that could be employed for emotion and cognitive process recognition, (Dzedzickis, Kaklauskas, & Bucinskas, 2020; Verma & Tiwary, 2014). Furthermore, while the current study evaluated the different VR phases from a macrolevel approach of simulation sequence, only further studies that code specific behaviors would allow us to identify the actions triggered by VR in the different phases and determine how this impacts engagement and achievements. Additionally, to limit issues of multicollinearity, only variables that demonstrated the strongest and most consistent correlations with the outcome variables were selected for inclusion in the subsequent regression analysis. Though a common approach (Duffy et al., 2018, p. 101150), this might potentially hinder important statistical interactions. Finally, the study focused only on the emotional and cognitive constructs of engagement without examining the behavioral construct of engagement. Again, future studies may be able to incorporate all three types of engagement constructs.

5. Conclusion

The combination of multichannel objectives (eye tracking, facial expressions, and EDA) and subjective data (self-reports) provides a new approach by which important continuous data about learning processes can be captured. The results demonstrate that the extraction of subconscious emotions via facial expression analysis is more sensitive to negative emotional valence (i.e., anger) when explaining students' learning outcomes. In contrast, subjective reporting of emotions of which students were consciously aware was more sensitive to positive emotional valence in explaining their learning outcomes. As such, this study demonstrates that it is crucial to include both automated objective channels and subjective methodologies in order to provide a more holistic and comprehensive perspective of complex learning processes and learning outcomes.

An additional contribution of this study is related to the instructional design. Higher

Psycho-physiological granularity measures revealed that cognitive and emotional engagement levels vary according to the learning activity, and specifically that more engagement is triggered when declarative learning is incorporated with procedural learning. VR learning environments represent a promising technology that is now being widely adopted in the educational field. One reason for their popularity is that these environments provide flexible access to both declarative and procedural learning within the same medium (Fromm et al., 2021).

Granular analysis using psycho-physiological sensors revealed that students who experienced higher levels of anger or focused their visual attention on irrelevant information were prone to lower achievements. These insights can contribute to the development of intelligent actionable personalized feedback that will be able to respond to psycho-physiological signals and support learners (Mills, Gregg, Bixler, & D'Mello, 2021). Consequently, real-time interventions that respond to a student's visual attention flow or emotional intensity may be employed to improve instructional design and potentially improve learning gains.

Funding

This work was supported, in part, by the Joy Ventures Academic Grant.

Declaration of competing interest

None.

Acknowledgements

The author would like to thank engineer Sergey Bondarchuck with whom the VR learning environment was designed and programmed, Liat Liberman for assisting with data collection, and Dr Ruth Moont for reviewing and editing the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2022.104495.

References

Ahn, B. T., & Harley, J. M. (2020). Facial expressions when learning with a queer history app: Application of the control value theory of achievement emotions. *British Journal of Educational Technology, 51*(5), 1563–1576.

Ainley, M. (2012). Students' interest and engagement in classroom activities. In *Handbook of research on student engagement* (pp. 283–302). Springer.

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education, 125*, 413–428.

Anderson, J. R. (2015). *Cognitive psychology and its implications* (8th ed.). Macmillan.

Andreassi, J. L. (2010). Electrodermal activity (EDA) and behavior. In *Psychophysiology: Human behavior and physiological response* (5 ed., pp. 314–340). Psychology Press. https://doi.org/10.4324/9780203880340.

Armougum, A., Orriols, E., Gaston-Bellegarde, A., Joie-La Marle, C., & Piolino, P. (2019). Virtual reality: A new method to investigate cognitive load during navigation. *Journal of Environmental Psychology, 65*, 101338.

Baceviciute, S., Lucas, G., Terkildsen, T., & Makransky, G. (2021). Investigating the redundancy principle in immersive virtual reality environments: An eye-tracking and EEG study. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12595. *n/a*(n/a).

Baceviciute, S., Terkildsen, T., & Makransky, G. (2021). Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in Virtual Reality. *Computers & Education, 164*, 104122.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190*(1), 80–91.

van Bommel, R., Stieger, M., Visalli, M., de Wijk, R., & Jager, G. (2020). Does the face show what the mind tells? A comparison between dynamic emotions obtained from facial expressions and temporal dominance of emotions (TDE). *Food Quality and Preference, 85*, 103976.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology, 49*(1), 1017–1034.

Brown, K. G., & Ford, J. K. (2002). Using computer technology in training: Building an infrastructure for active learning. *Creating, implementing, and managing effective training and development*, 192–233.

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*. Springer.

Chang, Yu Mei, & Lai, Chin Lun (2021). Exploring the experiences of nursing students in using immersive virtual reality to learn nursing skills. *Nurse Education Today, 97*. https://doi.org/10.1016/j.nedt.2020.104670

Chen, S., Epps, J., Ruiz, N., & Chen, F. (2011). Eye activity as a measure of human mental effort in HCI. In *Proceedings of the 16th international conference on Intelligent user interfaces*.

Cheng, M. T., Huang, W. Y., & Hsu, M. E. (2020). Does emotion matter? An investigation into the relationship between emotions and science learning outcomes in a game-based learning environment. *British Journal of Educational Technology, 51*(6), 2233–2251.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Choi, Dong Hwa, Dailey-Hebert, Amber, & Simmons Estes, Judi (2016). *Emerging Tools and Applications of Virtual Reality in Education*. IGI Global.

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement*. Springer.

Cobb, D. P., Jashami, H., & Hurwitz, D. S. (2021). Bicyclists' behavioral and physiological responses to varying roadway conditions and bicycle infrastructure. *Transportation Research Part F: Traffic Psychology and Behaviour, 80*, 172–188.

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, Analytic, Automated (AAA) measurement of engagement during learning. *Educational Psychologist, 52*(2), 104–123. https://doi.org/10.1080/00461520.2017.1281747

Daily, S. B., James, M. T., Roy, T., & Darnell, S. S. (2015). EngageMe: Designing a visualization tool utilizing physiological feedback to support instruction. *Technology, Instruction, Cognition and Learning, 10*(2).

Deci, E. L., & Ryan, R. M. (1985). ***Intrinsic motivation and self-determination in human behavior***. Plenum. https://doi.org/10.2307/2070638

Dede, C. J., Jacobson, J., & Richards, J. (2017). Introduction: Virtual, augmented, and mixed realities in education. In *Virtual, augmented, and mixed realities in education* (pp. 1–16). Springer.

Dindar, M., Jarvela, S., Ahola, S., Huang, X., & Zhao, G. (2020). Leaders and followers identified by emotional mimicry during collaborative learning: A facial expression recognition study on emotional valence. *IEEE Transactions on Affective Computing*.

D'Mello, S. (2017). In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *Emotional learning analytics* (pp. 115–127).

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction, 22*(2), 145–157.

Dubovi, I., & Lee, V. R. (2019). Instructional support for learning with agent-based simulations: A tale of vicarious and guided exploration learning approaches. *Computers & Education, 142*, 103644.

Dubovi, I., Levy, S. T., & Dagan, E. (2017). Now I know how! the learning process of medication administration among nursing students with non-immersive desktop virtual reality simulation. *Computers & Education, 113*, 16–27.

Dubovi, Ilana (2022). Learning with virtual reality simulations: direct versus vicarious instructional experience. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2022.2042313

Duffy, M. C., Lajoie, S. P., Pekrun, R., & Lachapelle, K. (2018). *Emotions in medical education: Examining the validity of the Medical Emotion Scale (MES) across authentic medical learning environments*. Learning and Instruction.

Dzedzickis, A., Kaklauskas, A., & Bucinskas, V. (2020). Human emotion recognition: Review of sensors and methods. *Sensors, 20*(3), 592.

Ekman, P. (1992). Are there basic emotions? *Psychol. Rev., 99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Ekman, P., & Rosenberg, E. L. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. USA: Oxford University Press.

Emerson, A., Cloude, E. B., Azevedo, R., & Lester, J. (2020). Multimodal learning analytics for game-based learning. *British Journal of Educational Technology, 51*(5), 1505–1526.

Felnhofer, A., Kothgassner, O. D., Schmidt, M., Heinzle, A.-K., Beutl, L., Hlavacs, H., et al. (2015). Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. *International Journal of Human-Computer Studies, 82*, 48–56.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Friedrich, M., Rußwinkel, N., & Möhlenbrink, C. (2017). A guideline for integrating dynamic areas of interests in existing set-up for capturing eye movement: Looking at moving aircraft. *Behavior Research Methods, 49*(3), 822–834.

Fromm, J., Radianti, J., Wehking, C., Stieglitz, S., Majchrzak, T. A., & vom Brocke, J. (2021). More than experience?-On the unique opportunities of virtual reality to afford a holistic experiential learning cycle. *The Internet and Higher Education, 50*, 100804.

Ghanbarzadeh, R., Ghapanchi, A. H., Blumenstein, M., & Talaei-Khoei, A. (2014). A decade of research on the use of three-dimensional virtual worlds in health care: A systematic literature review. *Journal of Medical Internet Research, 16*(2), e47.

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist, 50*(1), 14–30.

Greene, J. A., Seung, B. Y., & Copeland, D. Z. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers & Education, 76*, 55–69.

Guan, J.-Q., Wang, L.-H., Chen, Q., Jin, K., & Hwang, G.-J. (2021). Effects of a virtual reality-based pottery making approach on junior high school students' creativity and learning engagement. *Interactive Learning Environments*, 1–17.

Gupta, A., Elby, A., & Danielak, B. A. (2018). Exploring the entanglement of personal epistemologies and emotions in students' thinking. *Physical Review Physics Education Research, 14*(1), Article 010129.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., et al. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology, 96*(3), 403.

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior, 48*, 615–625.

Harmon, Joanne, Pitt, Victoria, Summons, Petter, & Inder, Kerry J. (2021). Use of artificial intelligence and virtual reality within clinical simulation for nursing pain education: A scoping review. *Nurse Education Today, 97*. https://doi.org/10.1016/j.nedt.2020.104700
Hayes, A. F. (2012). *Process: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling*. KS: University of Kansas.
He, W., Boesveldt, S., de Graaf, C., & de Wijk, R. A. (2016). The relation between continuous and discrete emotional responses to food odors with facial expressions and non-verbal reports. *Food Quality and Preference, 48*, 130–137.
Holland, M. K., & Tarlow, G. (1972). Blinking and mental load. *Psychological Reports, 31*(1), 119–127.
Holland, M. K., & Tarlow, G. (1975). Blinking and thinking. *Perceptual & Motor Skills, 41*(2), 403–406.
Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: OUP.
Hong, J., Pi, Z., & Yang, J. (2018). Learning declarative and procedural knowledge via video lectures: Cognitive load and learning effectiveness. *Innovations in Education & Teaching International, 55*(1), 74–81.
Huang, Hsiu-Mei, Liaw, Shu-Sheng, & Lai, Chung-Min (2016). Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. *Interactive Learning Environments, 24*(1), 3–19. https://doi.org/10.1080/10494820.2013.817436
Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies, 23*(4), 1515–1529.
Johnson-Glenberg, M. C. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI, 5*, 81.
Khan, Rishad, Plahouras, Joanne, Johnston, Bradley C., Scaffidi, Michael A., Grover, Samir C., & Walsh, Catharine M. (2019). Virtual reality simulation training in endoscopy: a Cochrane review and meta-analysis. *Endoscopy, 51*(7), 653–664.
Köster, E. P., & Mojet, J. (2015). From mood to food and from food to mood: A psychological perspective on the measurement of food-related emotions in consumer research. *Food Research International, 76*, 180–191.
Krejtz, K., Duchowski, A. T., Krejtz, I., Kopacz, A., & Chrząstowski-Wachtel, P. (2016). Gaze transitions when learning with multimedia. *Journal of Eye Movement Research, 9*(1).
Kruger, J.-L., & Doherty, S. (2016). Measuring cognitive load in the presence of educational video: Towards a multimodal methodology. *Australasian Journal of Educational Technology, 32*(6).
Kulke, L., Feyerabend, D., & Schacht, A. (2020). A comparison of the Affectiva iMotions Facial Expression Analysis Software with EMG for identifying facial expressions of emotion. *Frontiers in Psychology, 11*, 329.
Ladd, G. W., & Dinella, L. M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade? *Journal of Educational Psychology, 101*(1), 190–206. https://doi.org/10.1037/a0013153
Larmuseau, C., Vanneste, P., Cornelis, J., Desmet, P., & Depaepe, F. (2019). Combining physiological data and subjective measurements to investigate cognitive load during complex learning. *Frontline Learning Research, 7*(2), 57–74.
Lee, V. R., Fischback, L., & Cain, R. (2019). A wearables-based approach to detect and identify momentary engagement in afterschool Makerspace programs. *Contemporary Educational Psychology, 59*, 101789.
Li, Lan, Yu, Fei, Shi, Dongquan, Shi, Jianping, Tian, Zongjun, Yang, Jiquan, … Jiang, Qing (2017). Application of virtual reality technology in clinical medicine. *American Journal of Translational Research, 9*(9), 3867–3880.
Liaw, H., Yu, Y.-R., Chou, C.-C., & Chiu, M.-H. (2021). Relationships between facial expressions, prior knowledge, and multiple representations: A case of conceptual change for kinematics instruction. *Journal of Science Education and Technology, 30*(2), 227–238.
Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education, 95*, 174–187.
Liu, H.-C., & Chuang, H.-H. (2011). An examination of cognitive processing of multimedia information based on viewers' eye movements. *Interactive Learning Environments, 19*(5), 503–517.
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.
Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225–236.
Mancini, G., Biolcati, R., Agnoli, S., Andrei, F., & Trombini, E. (2018). Recognition of facial emotional expressions among Italian pre-adolescents, and their affective reactions. *Frontiers in Psychology, 9*, 1303.
Marín-Morales, J., Higuera-Trujillo, J. L., Greco, A., Guixeres, J., Llinares, C., Scilingo, E. P., et al. (2018). Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors. *Scientific Reports, 8*(1), 1–15.
Martins, R., & Carvalho, J. (2015). Eye blinking as an indicator of fatigue and mental load-a systematic review. *Occupational safety and hygiene III, 10*, 231–235.
Matsumoto, R. R., Walker, B. B., Walker, J. M., & Hughes, H. C. (1990). In J. T. Cacioppo, & L. G. Tassinary (Eds.), *Fundamentals of neuroscience* (pp. 58–112). https://psycnet.apa.org/record/1990-98757-003.
McDuff, D., Kodra, E., Kaliouby, R.e., & LaFrance, M. (2017). A large-scale analysis of sex differences in facial expressions. *PLoS One, 12*(4), Article e0173942.
Miller, B. W. (2015). Using reading times and eye-movements to measure cognitive engagement. *Educational Psychologist, 50*(1), 31–42.
Mills, C., Gregg, J., Bixler, R., & D'Mello, S. K. (2021). Eye-mind reader: An intelligent reading interface that promotes long-term comprehension by detecting and responding to mind wandering. *Human-Computer Interaction, 36*(4), 306–332.
Munshi, A., Mishra, S., Zhang, N., Paquette, L., Ocumpaugh, J., Baker, R., et al. (2020). Modeling the relationships between basic and achievement emotions in computer-based learning environments. In *International conference on artificial intelligence in education*.
Ning Woon, A. P., Mok, W. Q., Chieng, Y. J. S., Zhang, H. M., Ramos, P., Mustadi, H. B., et al. (2020). *Effectiveness of virtual reality training in improving knowledge among nursing students: A systematic review, meta-analysis and meta-regression*. Nurse Education Today.
Nourbakhsh, N., Wang, Y., Chen, F., & Calvo, R. A. (2012). Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. In *Proceedings of the 24th Australian computer-human interaction conference*.
O'Brien, H. L., Roll, I., Kampen, A., & Davoudi, N. (2021). Rethinking (Dis)engagement in human-computer interaction. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2021.107109, 107109.
Orquin, J. L., Ashby, N. J., & Clarke, A. D. (2016). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making, 29*(2–3), 103–115.
Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning, 37*(1), 226–241.
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., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101*(1), 115.
Pekrun, R., & Linnenbrink-Garcia, L. (2012). In A. R. S. Christenson, & C. Wylie (Eds.), *Academic emotions and student engagement* (Springer ed., pp. 259–282). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_12.
Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In *International handbook of emotions in education* (pp. 120–141). https://doi.org/10.4324/97802 03148 211.ch7. Routledge.
Peng-Li, D., Mathiesen, S. L., Chan, R. C., Byrne, D. V., & Wang, Q. J. (2021). Sounds Healthy: Modelling sound-evoked consumer food choice through visual attention. *Appetite, 164*, 105264.
Piromchai, Patorn, Avery, Alex, Laopaiboon, Malinee, Kennedy, Gregor, & O'Leary, Stephen (2015). Virtual reality training for improving the skills needed for performing surgery of the ear, nose or throat. *Cochrane Database of Systematic Reviews, 9*. https://doi.org/10.1002/14651858.CD010198.pub2

Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. *Educational Psychology Review, 31*(2), 339–359. https://doi.org/10.1007/s10648-019-09473-5

Renninger, K. A., & Bachrach, J. E. (2015). Studying triggers for interest and engagement using observational methods. *Educational Psychologist, 50*(1), 58–69. https://doi.org/10.1080/00461520.2014.999920

Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and engagement*. Routledge. https://doi.org/10.4324/9781315771045

Sawyer, R., Smith, A., Rowe, J., Azevedo, R., & Lester, J. (2017). Enhancing student models in game-based learning with facial expression recognition. In *Proceedings of the 25th conference on user modeling, adaptation and personalization*.

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3459–3474.

Schotter, E. R., Berry, R. W., McKenzie, C. R., & Rayner, K. (2010). Gaze bias: Selective encoding and liking effects. *Visual Cognition, 18*(8), 1113–1132.

Setianingsih, T., & Nafisah, B. Z. (2021). Memory trick game towards students' enthusiasm in learning grammar. *JISIP (Jurnal Ilmu Sosial dan Pendidikan), 5*(3).

Sharma, K., & Giannakos, M. (2020). Multimodal data capabilities for learning: What can multimodal data tell us about learning? *British Journal of Educational Technology, 51*(5), 1450–1484.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317–1322.

Shorey, Shefaly, & Ng, Esperanza Debby (2021). The use of virtual reality simulation among nursing students and registered nurses: A systematic review. *Nurse Education Today, 98*. https://doi.org/10.1016/j.nedt.2020.104662

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). In *The challenges of defining and measuring student engagement in science* (Vol. 50). Routledge.

Skinner, E. (2016). *Handbook of motivation at school*. Routledge.

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765–781. https://doi.org/10.1037/a0012840

Skulmowski, A., & Xu, K. M. (2021). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review*. https://doi.org/10.1007/s10648-021-09624-7

Stern, J. A., & Skelly, J. J. (1984). The eye blink and workload considerations. In *Proceedings of the human factors society annual meeting*.

Taub, M., Sawyer, R., Lester, J., & Azevedo, R. (2020). The impact of contextualized emotions on self-regulated learning and scientific reasoning during learning with a game-based learning environment. *International Journal of Artificial Intelligence in Education, 30*(1), 97–120.

Thomas, A. W., Molter, F., Krajbich, I., Heekeren, H. R., & Mohr, P. N. (2019). Gaze bias differences capture individual choice behaviour. *Nature Human Behaviour, 3*(6), 625–635.

Tonguç, G., & Ozkara, B. O. (2020). Automatic recognition of student emotions from facial expressions during a lecture. *Computers & Education, 148*, 103797.

Verma, G. K., & Tiwary, U. S. (2014). Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals. *NeuroImage, 102*, 162–172.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225–240.

Woon, Adele Pei Ning, Mok, Wen Qi, Chieng, Ying Jia Shermin, Zhang, Hui Min, Ramos, Patricia, Mustadi, Haryani Binte, & Lau, Ying (2021). Effectiveness of virtual reality training in improving knowledge among nursing students: A systematic review, meta-analysis and meta-regression. *Nurse Education Today, 98*. https://doi.org/10.1016/j.nedt.2020.104655

Xie, K., Heddy, B. C., & Greene, B. A. (2019). Affordances of using mobile technology to support experience-sampling method in examining college students' engagement. *Computers & Education, 128*, 183–198.

Xing, W., Tang, H., & Pei, B. (2019). Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of MOOCs. *The Internet and Higher Education, 43*, 100690.

Yu, Z. (2021). A meta-analysis of the effect of virtual reality technology use in education. *Interactive Learning Environments*, 1–21.