frontiers | Research Topics

Exploring the emotional landscape: cutting-edge technologies for emotion assessment and elicitation

Edited by
Javier Marín-Morales, Alessia Celeghin, Enrico Vezzetti
and Federica Marcolin

Coordinated by
Ivonne Castiblanco Jimenez

Published in
Frontiers in Psychology
Frontiers in Computer Science

ISSN 1664-8714
ISBN 978-2-8325-5914-7
DOI 10.3389/978-2-8325-5914-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Exploring the emotional landscape: cutting-edge technologies for emotion assessment and elicitation

Topic editors
Javier Marín-Morales — Universitat Politècnica de València, Spain
Alessia Celeghin — University of Turin, Italy
Enrico Vezzetti — Polytechnic University of Turin, Italy
Federica Marcolin — Polytechnic University of Turin, Italy

Topic coordinator
Ivonne Castiblanco Jimenez — University of Turin, Italy

Citation
Marín-Morales, J., Celeghin, A., Vezzetti, E., Marcolin, F., Castiblanco Jimenez, I., eds. (2025). *Exploring the emotional landscape: cutting-edge technologies for emotion assessment and elicitation*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-5914-7

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 07 January 2025
DOI 10.3389/fpsyg.2024.1542158

OPEN ACCESS

EDITED AND REVIEWED BY
Florin Dolcos,
University of Illinois at Urbana-Champaign, United States

*CORRESPONDENCE
Ivonne Angelica Castiblanco Jimenez
✉ ivonneangelica.castiblancojimenez@unito.it

RECEIVED 09 December 2024
ACCEPTED 20 December 2024
PUBLISHED 07 January 2025

CITATION
Castiblanco Jimenez IA, Marcolin F, Vezzetti E, Marín Morales J and Celeghin A (2025) Editorial: Exploring the emotional landscape: cutting-edge technologies for emotion assessment and elicitation. *Front. Psychol.* 15:1542158. doi: 10.3389/fpsyg.2024.1542158

Editorial: Exploring the emotional landscape: cutting-edge technologies for emotion assessment and elicitation

Ivonne Angelica Castiblanco Jimenez[1]*, Federica Marcolin[2], Enrico Vezzetti[2], Javier Marín Morales[3] and Alessia Celeghin[1]

[1]Department of Psychology, Università degli Studi di Torino, Turin, Italy, [2]Department of Management and Production Engineering, Politecnico di Torino, Turin, Italy, [3]Institute for Human-Centered Technology Research (HUMAN-Tech), Universitat Politècnica de València, Valencia, Spain

KEYWORDS

emotion assessment, affective elicitation, biometrics, physiological data, neuroscience, affective computing, augmented/virtual reality, human-computer interaction

Editorial on the Research Topic
Exploring the emotional landscape: cutting-edge technologies for emotion assessment and elicitation

The challenge of capturing human emotions

The assessment of human emotions in psychological research presents multiple methodological challenges. While traditional measurement approaches have contributed valuable insights, they often face limitations in capturing the complexity of emotional experiences and maintaining ecological validity. Recent technological advances provide new opportunities to address these challenges, enabling more precise and naturalistic approaches to emotion assessment and elicitation.

This Research Topic reunites research that employs emerging technologies for emotion research. The collected works demonstrate how various methodological approaches can overcome traditional limitations while expanding our understanding of emotional processes in both controlled and natural settings. These advances have been demonstrated in various novel technologies applications that are redefining how we study and understand emotions. From virtual reality environments to artificial intelligence and innovative physiological measurements, these progresses are expanding the landscape of emotion research.

Technological advances in emotion assessment

Virtual reality technology has emerged as a powerful tool for psychological research and therapy. Zhang et al. present a systematic investigation of the Apple Vision Pro's capabilities in emotional research and therapeutic applications. Through the integration of multi-sensor technology, high-resolution displays, and remote meeting capabilities, their work highlights VR's potential for creating immersive environments suitable for both emotion assessment and therapeutic interventions.

The integration of artificial intelligence with neurophysiological data represents another significant advancement in emotion research. Wang et al. present an innovative approach using vision transformers for EEG-based emotion classification during music listening tasks. Through advanced neural network architectures, their work advances automated emotion recognition from brain activity patterns.

In a complementary investigation of physiological intervention methods, Cong et al. demonstrate how transcutaneous electrical acupoint stimulation at PC6 can effectively reduce fear and improve emotion regulation, as evidenced through changes in heart rate variability.

The complexity of emotional experiences demands sophisticated measurement approaches. In learning environments, Horvers et al. advance this field by investigating emotional responses to immediate feedback during math problem-solving tasks. Their work combines physiological signals (electrodermal activity, electrocardiogram) with experiential and behavioral measures (self-reports, observations of facial expressions) to provide richer insights into emotional experiences, highlighting the value of multimodal assessment approaches in educational settings.

Emerging research further establishes concrete metrics for emotion detection in learning environments. Lal et al. identify specific physiological and behavioral indicators of learner emotions through careful analysis of skin conductance, temperature, and eye movements. Their findings reveal that measures such as skin conductance response peaks and eye-tracking metrics can effectively distinguish between emotional states, offering practical tools for emotion-aware learning technologies.

The challenge of collecting emotion data in natural settings has led to innovative solutions. Niewiadomski et al. present a framework for physiological data collection based on appraisal theories. Using a wearable device to collect physiological signals (blood volume pulse, electrodermal activity, skin temperature) and movement data, their mobile application detects potentially relevant emotional events and prompts users for self-reports, bridging the gap between laboratory precision and real-world applications.

The effectiveness of emotion regulation strategies varies significantly across different contexts and individuals. Int-Veen et al. examine this complexity through their study of daily emotional regulation practices, revealing important relationships between regulation strategies, self-efficacy, and stress management. Their findings highlight how different regulation approaches may be more or less effective depending on individual characteristics and situational factors.

Physical environments play a crucial role in shaping emotional experiences. Zhang analyses students' emotional attachment to a renovated campus landscape space called the "Heart of Forest," a 12,000-square-meter public activity area at Beijing Forestry University. Through a mixed-method approach combining emotional attachment scales with sentiment analysis, this study reveals how landscape design can enhance students' emotional attachment and wellbeing in educational settings.

Challenges and opportunities in emotion research

The multimodal approaches presented in this Topic highlight specific challenges in data harmonization and interpretation. The integration of EEG measurements, physiological interventions, and eye-tracking metrics requires careful consideration of different temporal and spatial scales. Applications in educational and therapeutic settings illustrate the need to balance technical capabilities with practical implementation. Furthermore, as monitoring extends into daily-life contexts, questions of privacy and informed consent become central to research design and implementation.

This Research Topic demonstrates significant progress in emotion research methodology through technological innovation. Each study advances our understanding while highlighting the importance of maintaining methodological rigor in real-world applications. The convergence of virtual reality, physiological sensing, and artificial intelligence creates new possibilities for understanding emotional processes. As these technologies continue to develop, future research must address both the technical challenges of data integration and the ethical considerations of continuous emotional monitoring. This field moves forward not just through technological advancement, but through careful consideration of how these tools can best serve our understanding of human emotional experience.

Author contributions

IC: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. FM: Conceptualization, Writing – review & editing. EV: Writing – review & editing. JM: Writing – review & editing. AC: Conceptualization, Writing – review & editing.

Acknowledgments

The guest editors would like to thank all authors and reviewers who contributed to this Research Topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 25 September 2023
DOI 10.3389/fpsyg.2023.1250441

OPEN ACCESS

EDITED BY
Javier Marín-Morales,
Universitat Politècnica de València, Spain

REVIEWED BY
Juan Luis Higuera-Trujillo,
Universitat Politècnica de València, Spain
Pongsakorn Suppakittpaisarn,
Chiang Mai University, Thailand

*CORRESPONDENCE
Ruoshi Zhang
✉ zhang_rs@bjfu.edu.cn

RECEIVED 30 June 2023
ACCEPTED 28 August 2023
PUBLISHED 25 September 2023

CITATION
Zhang R (2023) Research and evaluation on students' emotional attachment to campus landscape renewal coupling emotional attachment scale and public sentiment analysis: a case study of the "Heart of Forest" in Beijing Forestry University.
Front. Psychol. 14:1250441.
doi: 10.3389/fpsyg.2023.1250441

Research and evaluation on students' emotional attachment to campus landscape renewal coupling emotional attachment scale and public sentiment analysis: a case study of the "Heart of Forest" in Beijing Forestry University

Ruoshi Zhang*

School of Landscape Architecture, Beijing Forestry University, Beijing, China

In the era of stock renewal, the construction of university campuses in China's first-tier cities has shifted from demolition and construction to renewal and upgrading, in which public landscape space is the main environment for students' daily life, learning and entertainment. Especially during the outbreak of the recent COVID-19 epidemic, it has become an important way for students to interact with nature and obtain emotional healing. In the existing studies, there is a lack of discussion on the correlation between the spatial characteristics of the updated campus landscape and students' emotional attachment, and there are few quantitative studies. Based on this, this paper takes the "Heart of Forest" landscape space as an example, and integrates multi-dimensional quantitative methods including emotional attachment scale and public semantic analysis to study and evaluate the characteristics of landscape space that affect students' emotional attachment. The results show that: (1) Overall, the landscape space renewal of the Heart of Forest provides students with positive emotional experiences and effectively enhances students' emotional attachment as well as sense of belonging to the campus. (2) Among them, the material characteristics of the site including nature-related elements, materials, structures play a positive role in promoting the vast majority of students in the process of establishing emotional attachment, which is particularly obvious for students majoring in landscape, architecture and urban planning. (3) Whether the public social space can effectively provide students with a good emotional experience is closely related to the frequency and purpose of students' use of the space. (4) The interactive characteristics such as changeability and playability fail to promote emotional attachment because of lacking of management and maintenance. The renewal and transformation of the "Heart of Forest" landscape space is generally successful in promoting students' emotional attachment, and provides a reference for the future campus landscape renewal design from different angles. In addition, the quantitative study of emotional attachment constructed in this paper coupled with multi-dimensional data provides a method for the evaluation of students' emotional experience of campus landscape.

KEYWORDS

emotional attachment, campus landscape renewal, emotional attachment scale, public sentiment analysis, Heart of Forest

1. Introduction

In recent years, China's urban development has entered a stage of stock renewal (Yi et al., 2017; Ye et al., 2021). Under this circumstances, in the core areas of first-tier cities such as Beijing, where land is tight, the main way public education spaces development with large areas especially university campuses has also shifted from new construction and reconstruction to renewal and upgrading (Guo W. et al., 2022; Yuanshuo, 2023). Among them, the renewal of public landscape space on campus as the main place for students' extracurricular activities, rest and socialization is closely related to the physical and mental health of college students, and the design of good natural landscape and cultural landscape has also been proved to enhance students' emotional attachment and satisfaction with the campus environment (Lau and Yang, 2009; Andre et al., 2017; Vella-Brodrick and Gilowska, 2022). During the outbreak of COVID-19 in the past 3 years, the campus landscape space has especially become an important psychological healing space and emotional anchor point for students whose life, recreation and learning activities are confined to the campus space (Biswas and Sen, 2020; Lai et al., 2020; Rousseau and Deschacht, 2020; Liu et al., 2022). Therefore, the research and evaluation of the landscape space transformation on campus from the perspective of emotional attachment are of great value to both the improvement of students' wellbeing and the future campus construction in the context of stock renewal in China.

In the existing research on the relationship between landscape transformation and emotion, most of the researches focus on urban scale public spaces, and the types and ages of the research population are widely distributed, and few focuses on landscape renewal and young student groups on campus (Kara, 2013; Kong et al., 2022; Zeng et al., 2022; Yan et al., 2023). Relating researches mainly focuses on landscape design methods, and the discussion of emotions is mostly reflected in qualitative descriptions, lacking objective quantitative evaluation. This may be also due to the complexity and diversity of emotions. At this time, the multidisciplinary emotional attachment scale method and semantic analysis method of social media data with help of big data technologies provide opportunities to improve the quantitative research of emotional attachment (Zhang, 2023).

Based on this, the study takes the "Heart of the Forest" landscape space in Beijing Forestry University as an example. The space was renovated and put into use in 2020. The study couples the top-down targeted research using emotional attachment scale and the bottom-up emotional semantic analysis using social media big data, to quantify the emotional attachment of students to different landscape features and spatial elements in the landscape renewal and transformation of universities. On the one hand, the research complements the lack of quantitative research methods for the emotional attachment between college students and landscape space, and on the other hand, it also provides a more scientific and objective reference for the campus landscape renewal design that cares for students' emotional experience in the era of stock renewal.

2. Literature review

2.1. The effect of natural landscape to students' emotional attachment on campus

The value of natural landscape for human well-being and sustainability has been widely studied and recognized. Back to 1995, the Biophlia Hypothesis proposed by Wilson had assumed that human had innate need for natural environment as they born and evolved with it (Kellert and Wilson, 1995). Later, psychological evolutionary theories including Stress Reduction Theory (SRT) and Attention Restoration Theory (ART) have been developed based on this hypothesis, the former suggests that visiting and contacting with nature helps shift stress states to positive emotional states while discourage negative thoughts, and the later suggests that natural environment help improve cognitive performance and restore attention (Kaplan and Kaplan, 1989; Kaplan, 1995; Jiang et al., 2021). Steven Kaplan and Rachel Kaplan further emphasizes that unlike urban environment, nature is filled with more intriguing stimuli which may grabs attention in a bottom-up fashion. Its ability to establish emotional connection with people can be analyzed in terms of its evolutionary significance. Based on this, the importance of human beings themselves has been highlighted (Kaplan et al., 1998), and the relevant research on the impact of nature on people has begun to further deepen into people's multi-dimensional perception of nature, such as through viewing, smelling, touching, hearing, tasting, etc., people can get different degrees of emotional healing from the natural environment. For example, merely viewing plants or even pictures of nature has been proved to promote people's positive emotion (Lee et al., 2009; Honold et al., 2016). Research from neuroscience also proves this (Van den Berg, 2008). Besides, interacting with natural plants, touching natural materials and hearing natural sounds has also significantly enhanced people's positive experience (Koga and Iwasaki, 2013; Ratcliffe, 2021; Rickard and White, 2021). Nowadays, the supply of natural landscape such as urban greenspaces plays an important role in daily public life, as they serve as a place for exercise, relaxing, as well as natural exposures (McPherson, 1992; Van Leeuwen et al., 2010; Zhu et al., 2021). During COVID-19 pandemic especially city lockdowns, they have become the main space for both people's physical and psychological healing, and their recreational usage is increasing due to social limitations brought by such public health crisis. Multiple researches have reported that greenspace in cities such as parks provided positive contributions to self-reported well-being during the COVID-19 pandemic (Xie et al., 2020; Guo X. et al., 2022). Therefore, greenspace design and construction has become an important humanized renewal method in the current era of stock renewal.

However, few have delved specifically into the interaction between the greenspaces on university campuses and students' physical and mental health. It has been proved by multidisciplinary researchers that a positive university environment creates conditions for students' effective learning, academic achievements, fewer emotional problems and promotes students' overall wellbeing, of which the interaction with the environment itself plays an important role (Brandisauskiene et al., 2021; Vella-Brodrick and Gilowska, 2022). This is particularly prominent in research on students' emotional demands on the university environment during the COVID-19 pandemic (Liu et al., 2022). Therefore, exploring the relationship between the specific characteristics of campus green space and students' emotional attachment is conducive to further promoting the design and renewal methods of a sustainable campus environment that truly cares for students' wellbeing.

2.2. The definition and evaluation of emotional attachment between people and landscape

The study of emotional attachment is promoted by multidisciplinary researchers, including those involved in psychology, human geography, architecture, landscape architecture and urban planning (Scannell and Gifford, 2010; Lewicka, 2011; Ekkekakis, 2013; Adams, 2015; Scannell and Gifford, 2017; Liu et al., 2020; Birnbaum et al., 2021; Arnberger et al., 2022; Zhang et al., 2022). It discusses the relationship between people and certain environment based on emotion, cognition, and behavior. Emotions are the primary focus (Altman and Low, 2012; Adams, 2015). For different research objects, researchers have constructed multi-angle and multi-dimensional theoretical frameworks. Among them, the place attachment theory (PA) systematically and completely discusses the influence factors, as well as indicators of the emotional attachment between people and specific environment characteristics (Lewicka, 2011). Based on this, researchers in different fields have carried out different explorations. In the field of psychology and human geography, researchers explore the mechanism of the establishment of emotional attachment and construct multi-dimensional attachment measurement scales, of which different indicators promoting emotional attachment are emphasized (Williams and Roggenbuck, 1989; Altman and Low, 2012; Scannell and Gifford, 2013; Adams, 2015). Meanwhile, researchers from architecture, urban planning, landscape architecture and other fields related to environmental studies and design pay more attention to the different roles of the built environment and the natural environment on emotional attachment. For example, the place-based phenomenological research and the exploration of the meaning of people-built environment interactions are conducted to study people's sense of identity with buildings or cities (Tuan, 1979; Carmona, 2019; Feng et al., 2021; Arnberger et al., 2022). As for natural environment, the above mentioned Stress Reduction Theory (SRT) and Attention Restoration Theory (ART) have started to raise the importance of nature to emotion. Subsequently, some scholars construct models to analyze the relationship between place attachment and the natural environment, such as the intrinsic relationship between environmental preference, place attachment and restorative evaluation, and the resilience of the natural environment with a high degree of place attachment is higher than that of the urban environment (Lin et al., 2019; Liu et al., 2020; Birnbaum et al., 2021; Feng et al., 2021; Zhang, 2021; Arnberger et al., 2022). Whether in China or abroad, most of their research subjects are concentrated in urban renewal and landscape renewal. This is precisely because, in the context of stock renewal, the response of place attachment to the contradiction between people and land, the preservation of the context of the place and the promotion of public participation are of great value. This also supports the value of this theory for this study.

The method of using the scale to measure people's emotion is derived from the emotional attachment theories represented by Place Attachment Theory, and the scales used in different researches are constructed based on different research objects, problems or goals. Based on existing multidisciplinary researches, the measurement of attachment to landscape mainly includes the following three aspects: the dimension of emotional attachment between people and the landscape (positive or negative), the degree of emotional attachment, and the emotional attachment to specific landscape characteristics. They also correspond to the different goals of the researchers under different background.

The first two goals could be achieved with the help of the Positive and Negative Analysis Scale (PANAS) and the Place Attachment Scale. The positive and negative analysis scale (PANAS) is widely used to explore whether the strength of attachment to a certain space is closely related to the positive and negative emotions. It consists of 10 positive and 10 negative emotional adjectives, and is measured by Likert 5-point scoring method, of which 1 point indicates that there is almost no such emotional experience, and 5 points indicates that such emotional experience is strong. For the latter, the person-process-place (PPP) framework of place attachment measuring scale has shown high reliability and validity and has been used by scholars in multi-scale environment as well as landscape researches in the past 2 years. The Place Attachment Scale consists of 20 items, based on Likert's 7-point scoring method, of which 1 point indicates the least disagreement with the narrative content of the question and 7 points indicates strong agreement with the narrative content of the question.

Meanwhile, in order to explore the emotional influence of specific characteristics of the targeted landscape, researchers often construct scales based on research needs. In the field of built environment design and research, semi-structured measurement scales are formed based on the pre-study of the specific landscape characteristics. The scales mainly include the physical characteristics, social characteristics and interactive characteristics. The specific factors for each part vary slightly from study to study depending on the object. The Likert scale is also used to quantitatively measure the specific intensity of each characteristic, with a score of 1 indicating almost no emotional attachment to this characteristic and a score of 7 indicating a strong emotional attachment.

2.3. The application of semantic analysis in the evaluation of emotional attachment between people and landscape

Although the above emotional attachment scale provides a means to quantify the dimension and degree of emotion, the amount of data obtained is still limited, and its top-down construction logic cannot completely avoid the researcher effect. In the era of information and communication technology, visualizing sentiment based on social media data (SMD) provides a bottom-up quantitative research approach for evaluating human emotions, which is an effective supplement to the top-down scale method (Chakraborty et al., 2020; Jindal and Aron, 2021). SMD serve as a reliable and relatively objective data source to reflect the public's attitudes toward a certain landscape (Barbier and Liu, 2011; Sim et al., 2020). Among them, college students are the most active group of people who use these social platforms to express their feelings. This provides a data basis for the intervention of semantic analysis in this study.

The development and universalization of social media has changed the way people interact with environment, making it useful for exploring public preference for a landscape or place (Wilkins et al., 2020; You et al., 2022). Based on this, sentiment analysis is used to qualify and classify people's emotions during their interaction with landscape, and then define different degree of sentiment values as different emotions, which helps reveal the benefits of landscape. For example, Twitter sentiment and georeferenced Flickr tweets have been used in New York and Sheffield, respectively, to measure well-being brought by urban park and the quality of urban greenspaces. The results have shown that high sentiment in landscape relates to high positive emotions, which implies higher satisfaction (Brindley et al., 2019). Meanwhile, SMD from the two most popular social networking platform in China, Tecent and Sina Weibo, have been used to explore both people's emotional features and preference in different kinds of landscapes, such as emotional enthusiasm enhanced by greenspaces (Guo X. et al., 2022) and positive emotions brought by forest landscape (Zeng et al., 2022). Although the existing research fully supports the effectiveness of semantic analysis in evaluating the emotional attachment characteristics between people and landscapes, few studies have paid special attention to the emotional appeal and evaluation of campus landscape by college students. It is vital to explore and understand these students' emotional demands to improve campus landscape to match needs, especially under the background of stock renewal.

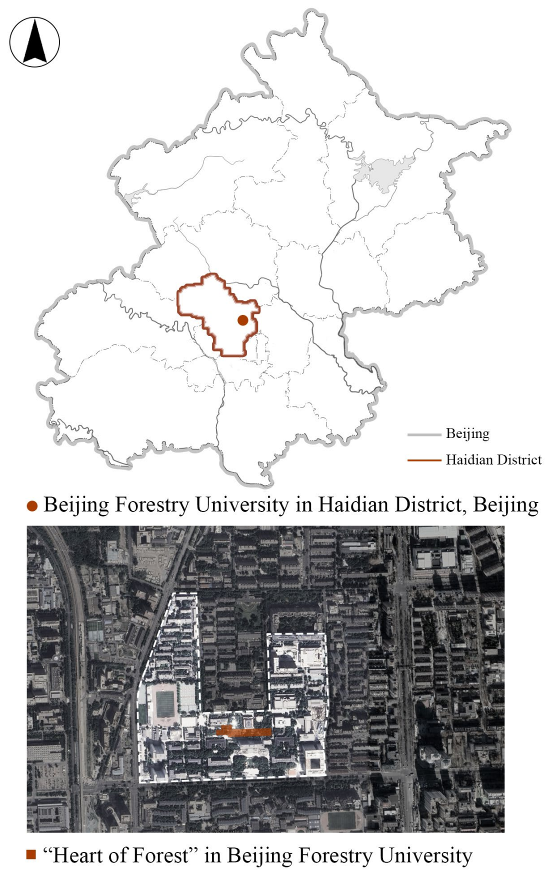


FIGURE 1
The location and area of the Heart of Forest.

3. Materials and methods

3.1. Study area

The study area, "Heart of Forest," is located in the center of Beijing Forestry University, covering a total area of about 12,000 square meters (Figure 1). This campus landscape space is the largest public activity space in Beijing Forestry University, which was completed and put into use in September 2020. Heart of Forest is a landscape renewal project after the demolition of the temporary building of the original school hospital, its transformation and design consists of three aspects: natural land-scape, artificial landscape, and interactive landscape. The natural landscapes include ponds, rainwater harvesting devices, lawns, old elm trees, and garden plants. The cultural and artificial landscape includes pavilions, seats, and a commemorative installation of the former school hospital. The interactive landscape includes herbarium walls, interactive light pillars, and voice recording devices. These landscape spaces promote students' emotional attachment to the Heart of Forest by providing them with a variety of emotional experiences through natural, cultural, historical, and behavioral inter-actions. Since its completion, the Heart of Forest has become the most popular place among students for leisure, activities and communication on the campus of Beijing Forestry University, which makes it a perfect area for this study. Combined with its design concept and landscape composition elements, the landscape characteristics that trigger emotional attachment can be integrated into three categories: material, cultural and interactive, including: material, color, natural elements, form and structure, privacy, diversity, sociability, territoriality, playability, uniqueness, and changeability. This also provides a basis for the construction of the emotional attachment scale of specific landscape composition characteristics in subsequent research. The specific features of each part of the Heart of Forest can be seen in Figure 2.

3.2. Data sources

3.2.1. Emotional attachment scale data

The emotional attachment scale was randomly distributed to students who visited the Heart of Forest to measure the degree and dimension of students' emotion when they were there. The data collection began in October 2022 and lasted about 2 weeks. The time period was under the epidemic of COVID 19 then, in order to ensure safety, almost all the students returning to school studied and lived on campus, and the Heart of Forest became the main public space for students to rest and communicate after class, alleviating the anxiety caused by the epidemic, which especially highlighted the emotional healing value of the Heart of Forest. A total of 159 scales were collected, of which 143 were with valid data. Since the completion of

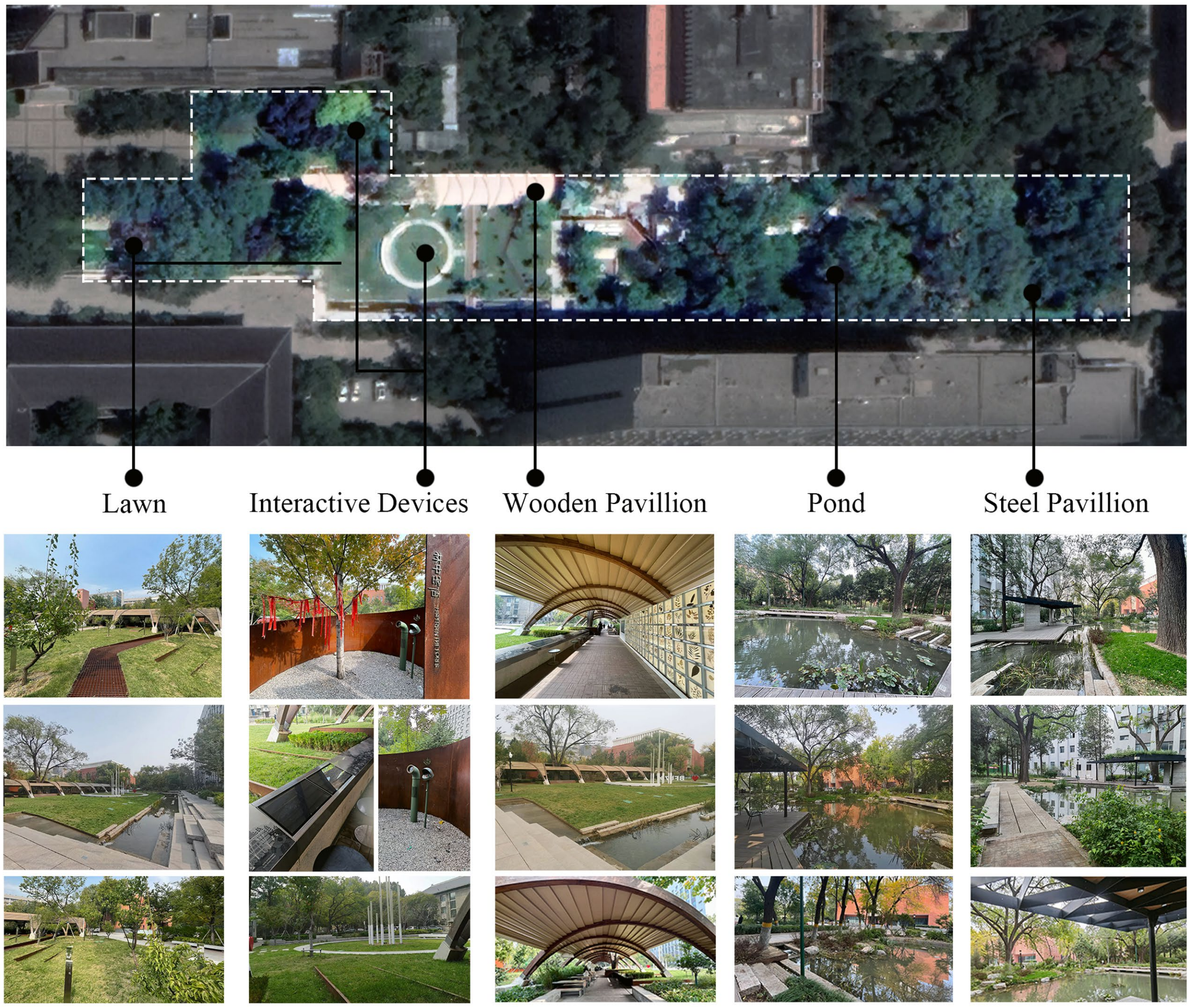


FIGURE 2
The representative features of each part in the Heart of Forest (photos taken by author).

each scale required the semi-structured guidance of a researcher, it took participants about 20 min to complete the scale, and campus activities were seriously limited during the epidemic of COVID-19 then, the scale of available valid data was the maximum amount that can be obtained with limited time and limited manpower. According to previous research, men and women in general have different characteristics of emotional attachment to the same landscape. Meanwhile, as the subject of this study, the Heart of Forest, is located inside the Beijing Forestry University, and its designer is a teacher of the School of Landscape Architecture, so students studying landscape, architecture and urban planning may have different emotional attachment when interacting with landscape. At the same time, whether or not one purposefully interacts with the environment has also been shown to affect the degree of emotional attachment to the environment. Based on this, the study also collected information of participants' major and purpose of visit. The gender, major distribution and whether this is a purposeful visit of the participants can be found in the table. Valid data is entered into IBM SPSS V25.0 (International Business Machines Corporation (IBM), IBM headquarters in Almonk, New York, NY, USA) for subsequent analysis.

3.2.2. Social media data

The text data associated with the Heart of Forest from Sina Weibo and Little Red Book (XiaoHongShu) was used for emotional semantic analysis in this study. Sina Weibo (Beijing, China) is the most popular platform for people to distribute and access information, especially college students. According to 2020 Weibo User Development Report, Weibo's monthly active users and daily active users reached 511 million and 224 million respectively, of which the post-90s and post-00s accounted for more than 80%, which makes it one of the most visited websites in mainland China as well as a perfect platform for accessing Weibo post data of students (Sina Weibo, 2020). Weibo post data refers to text information posted by Weibo users at a given place, which in this research is the Heart of Forest in Beijing Forestry University. The data is used to quantify students' emotions when they take series activities in the Heart of Forest. These data were obtained from the Sina Weibo application programming interface (API)[1] with

1 https://m.weibo.cn

the help of Python (version 3.7). The data cover the time range from September 1, 2020 when the Heart of Forest was open to students, to October 30, 2022. Same as Weibo, Little Red Book is a lifestyle sharing platform, which also allows users to share short videos and photos about fashion, beauty, food, travel, and much more. According to Qiangua's 2022 XiaoHongshu Active User Portrait Trend Report (2022), the number of monthly active users of Little Red Book reached 200 million in 2022, with 72% of users being born in the 1990s. The data obtained from Little Red Book include log and comment text relating to the Heart of Forest in Beijing Forestry University, which were collected from the API of https://www.xiaohongshu.com/ using Python programming language. The data cover the same time range of Weibo post data above (from September 1, 2020 to October 30, 2022).

The original Weibo and Little Red Book text contains various redundant information, including punctuation marks, subject tags, hyperlinks, and @other-users tags in Weibo. Python was used in data cleaning process to eliminate errors while improve the efficiency of word sentiment in sentiment analysis. The cleaned data were visualized using SPSS, and the location tagged with the Heart of Forest of Beijing Forestry University was selected to obtain 3,702 Weibo and Little Red Book data.

3.3. Build a methodology

3.3.1. Emotional attachment scale design and analysis

Based on the aforementioned review of the application of the emotional attachment scale, firstly, the purpose of this study should be clarified. There are three goals to quantitatively evaluate the emotional attachment between students and landscape of the Heart of Forest: the dimension of emotional attachment between students and the Heart of forest (positive or negative), the degree of emotional attachment, and the emotional attachment to specific characteristics consisting of the Heart of Forest.

Therefore, the Positive and Negative Analysis Scale (PANAS) and the Place Attachment Scale were chosen to achieve the first and second goals, and a newly constructed scale aiming at detailed characteristics of the Heart of Forest were designed to achieve the third goal. The former two scales met the aim of this study both in content and logic by emphasizing the importance of the environment itself in the attachment process. In order to avoid a cultural gap and improve the clarity of the scale, a Chinese version of the two scales was revised. Meanwhile, in order to explore the emotional influence of specific characteristics of the landscape of the Heart of Forest, a semi-structured measurement scale was formed based on the pre-study of landscape characteristics in the Heart of Forest. As a result, the physical characteristics including material, color, natural plants, form and structure, and social characteristics including privacy, diversity, sociability, territoriality, and interactive characteristics that including playability, uniqueness, and changeability were chosen as indicators.

IBM SPSS V25.0 (International Business Machines Corporation (IBM), IBM headquarters in Almonk, New York, NY, USA) was used for data processing after the data collection using the above mentioned scales. The analysis consisted of three steps.

The first step was the description of basic characteristics of the data representing comprehensive degree and intensity of attachment. Arithmetic average was used to describe the extent of positive (A_L) and negative emotion (A_N), place attachment (A_{PA}), and overall attachment to landscape characteristics of the Heart of Forest (A_{LC}). Standard deviation was used to describe the magnitude of the difference in participants' positive (S_P) and negative emotion (S_N), place attachment (S_{PA}), and attachment to overall landscape characteristics of the Heart of Forest (S_{LC}). The formula was as follows:

$$A_{(P,N,PA,LC)} = \frac{1}{n}\sum_{i=1}^{n} b_{(P,N,PA,LC)_i} \tag{1}$$

$$S_{(P,N,PA,LC)} = \sqrt{\frac{\sum_{i=1}^{n}\left(b_{(P,N,PA,LC)_i} - \overline{b_{(P,N,PA,LC)}}\right)^2}{n-1}} \tag{2}$$

where $b_{(P,N,PA,LC)i}$ represents the score of the *i*th participating student on each item in the four components of the above emotional attachment scale: positive affect (b_P), negative affect (b_N), place attachment (b_{PA}), and attachment to overall landscape characteristics (b_{LC}). Among them, b_P is the arithmetic average of the scores scored by the 10 words on the PANAS scale that describe positive emotions, b_N is the arithmetic average of the scores scored by the 10 words on the PANAS scale that describe negative emotions, b_{PA} is the arithmetic mean of the scores scored by all items on the PA scale, and b_{LC} is the arithmetic mean of the scores scored by each specific landscape characteristic on the attachment intensity to landscape characteristics scale. n is 144 in this study.

Additionally, the correlation between positive and negative affect, place attachment, and specific landscape characteristics was explored using correlation analysis (Pearson) based on SPSS. The assessment of the normality of data is conducted before correlation analysis using SPSS. The above arithmetic average result of positive affect, negative affect, place attachment and scores of the intensity of attachment to each landscape characteristic of each participant are used here as specific value to calculate the Pearson correlation coefficient r:

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}} \tag{3}$$

where the variables to both x and y are the same, including arithmetic average result of positive affect, negative affect, place attachment and scores of the intensity of attachment to each landscape characteristic of each participant.

Last but not least, with the help of exploratory factor analysis (EFA) in SPSS, the spatial characteristics contributing to attachment of the Heart of Forest were summarized into systematic spatial dimensions. Exploratory factor analysis (EFA) is used to identify complex interrelationships among items and group items that are part of unified concepts. There's no "*a priori*" assumptions have been made about the relationships among the emotional attachment factors. Extraction method in the analysis is Principal Component Analysis (PCA) with Varimax rotation method.

3.3.2. Sentiment analysis based on sentiment dictionaries

Identifying students' opinions and preference on school landscape transformation plays an important role in updating the existing campus environment and planning the future campus of Beijing Forestry University under the background of stock renewal. In this study, the textual content of Weibo and Little Red Book was used to analyze the sentiments of students who had experiences in the landscape of Heart of Forest from bottom up to avoid the researcher effect as well as provide support for the scale analysis above. Text sentiment analysis is a widely used tool for extracting subjective information from natural language texts and classify them into sentiment categories, such as positive, negative and neutral. The two most commonly used sentiment methods in text sentiment analysis are analysis using sentiment dictionaries and machine learning. According to the feature of the texts in this study, sentiment analysis methods based on sentiment dictionaries were chosen to collect and analyze students' attitudes and sentimental responses to different landscape characteristics of the Heart of Forest. Specifically, frequency in text messages reflects the attention of students to relevant landscape characteristics, and the size of emotional computing data reflects the intensity of students' emotional experience. The process of sentiment semantic analysis is shown in Figure 3.

Texts about the Heart of Forest were selected from the above mentioned two public social media platforms: Weibo and Little Red Book. Python was used to perform data mining and cleaning of the collected texts. A total of 3,702 valid sample data points were obtained. The NLPIR-Parser big data semantic intelligent analysis platform was used to preliminarily segment the comment content, and after filtering out meaningless words, statistical results for high-frequency words were obtained, and high-frequency words with the same semantics were merged.

High-frequency words classified using different evaluation scales and A Dictionary of Chinese Praise and Blame Words were screened by intersection with the Simplified Chinese Version of the National Taiwan University Sentiment Dictionary, the Chinese Emotional Vocabulary Ontology Database, and the China HowNet database (CNKI). Then, similar words were merged using A New Chinese Synonym Forest. After screening, the intensity of each word was used to define its emotional intensity, and the evaluation indicators were divided into four types: commendatory words (intensity: +1, +2, +3), derogatory words (intensity: −1, −2, −3), neutral words (intensity: 0), and special words (words not included in the dictionary). The emotional intensities of commendatory, derogatory, and neutral words not included in the dictionary were uniformly marked as +3, − 3, and 0, respectively.

3.3.3. Methodology framework

Figure 4 illustrate the framework of this study. The study focuses on the research and evaluation on students' emotional attachment to campus landscape renewal coupling emotional attachment scale and public sentiment analysis. The landscape of Heart of Forest in Beijing Forestry University is chosen as research object. The study adopts the logical path of "criteria establishment-methodology construction-results processing" to evaluate the emotional attachment of the updated landscape space and comprehensively analyze the value of specific landscape characteristics with the results of scale analysis and sentiment analysis.

4. Results

4.1. The results of emotional attachment scale analysis

4.1.1. The overall emotional attachment status to the Heart of Forest

The data of all participants and data based on gender and major differentiation of emotional attachment to the Heart of Forest is listed in Table 1. As for attachment degree represented by Place Attachment, the results showed that regardless of gender and major, students generally established a relatively strong place attachment to the Heart of Forest, as the average degree exceeded the mean value, while the value of standard deviation of the total data indicated that this

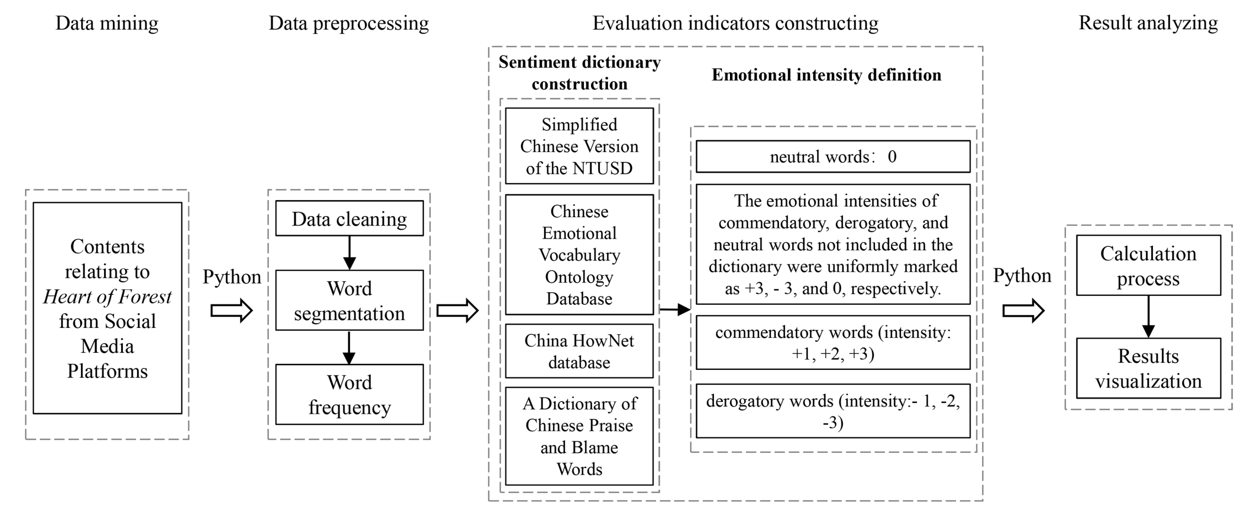


FIGURE 3
Sentiment analysis process.

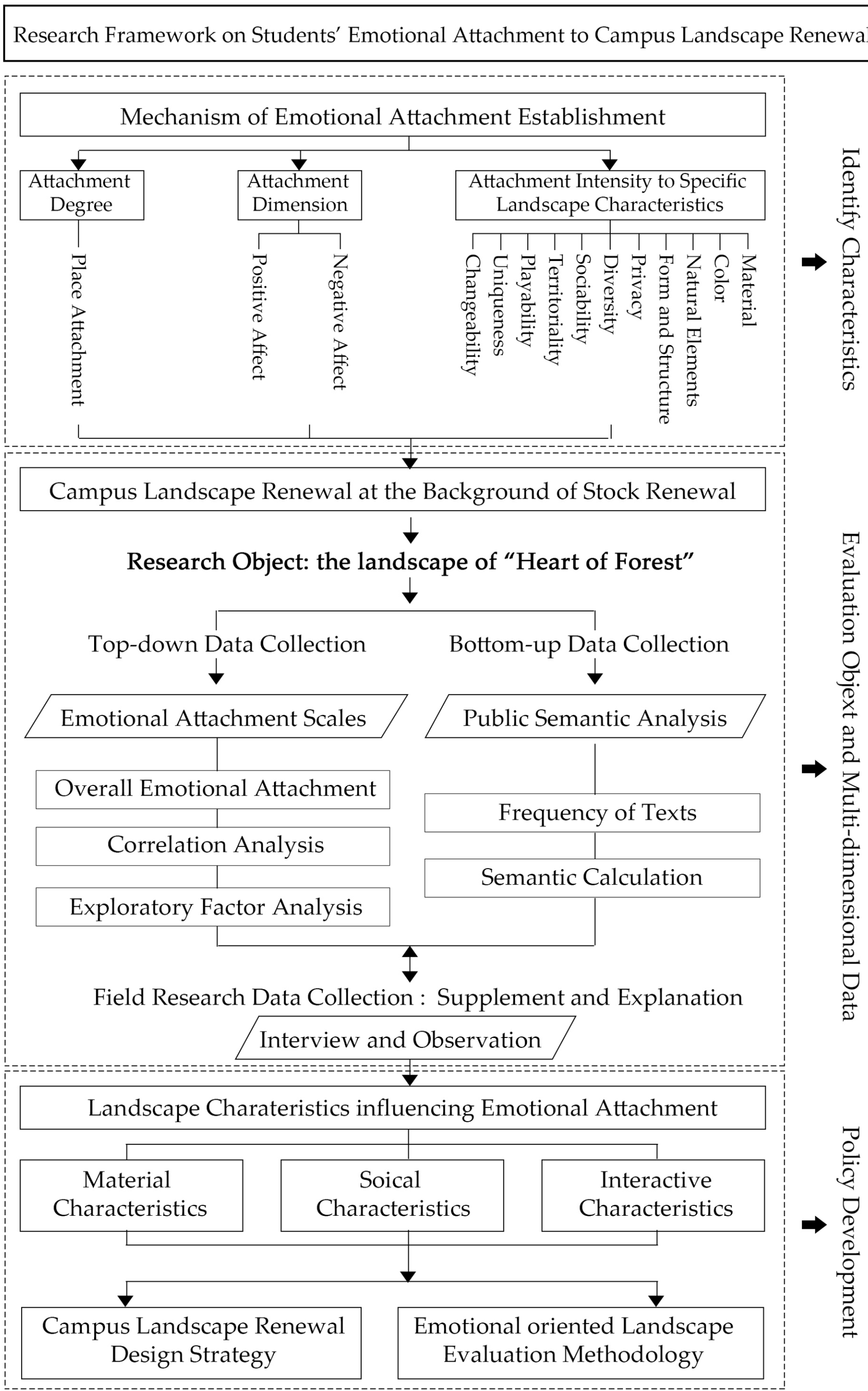


FIGURE 4
The framework of the study.

attachment degree was quite different between different student. Among them, female students had stronger place attachment than male students, while students majoring in environmental design related majors had significantly stronger attachment to the Heart of Forest than students of other majors. As for attachment dimension including positive affect and negative affect, the positive affect obtained by students was overall more pronounced than the negative affect. Among them, female students have stronger positive emotions and weaker negative emotions than male students. As for students' attachment intensity to landscape characteristics of the Heart of Forest, the results of mean value showed that they generally have a strong emotional experience of the characteristics of the landscape and environment themselves, which once again proves the emotion induced value of the specific characteristics of this designed landscape. For different genders, male students had a stronger emotional experience of landscape characteristics, while for different majors,

TABLE 1 The overall emotional attachment status.

Emotional attachment	Total		Male		Female		Environment design related majors	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Positive affect	2.88	0.85	2.84	0.91	2.89	0.82	2.93	0.79
Negative affect	1.39	0.59	1.60	0.88	1.31	0.38	1.34	0.46
Place attachment	4.42	0.87	4.39	0.85	4.43	0.88	4.65	0.84
Attachment intensity to landscape characteristic	5.31	0.94	5.35	0.87	5.29	0.98	5.49	0.87

students majoring in environmental design related majors continued to show a stronger emotional tendency toward landscape itself.

4.1.2. The correlation analysis results

The correlation analysis was processed to further explore the extent to which each specific landscape characteristic contributed to the degree and dimension of emotional attachment. Based on the result of p value for the S-W test and Q-Q plot of the tests of normality, the 14 variables are approximately normally distributed. The correlation coefficients have been calculated among all 14 objects. The results of the overall participants, male, female, and participants from environment design related majors are shown in Tables 2–5. According to Table 2, It was evident that students' overall attachment to the Heart of Forest was positively correlated with positive affect (0.553**). The results describing both the material and social characteristics of the landscape were significantly positively correlated with place attachment, as they all have a relatively high correlation coefficient at 0.01 (double-tailed). This demonstrated the value of the landscape design and construction of the campus environment renewal with respect to emotional attachment. Comparing the correlation strength between the material characteristics of the Heart of Forest as well as the social and interactive characteristics and emotional attachment, it could be found that the social and interactive characteristics have a greater effect on emotional attachment than the material characteristics, as five of the correlation factors of the former were greater than 0.5, including diversity, sociability, territoriality, playability, and uniqueness (0.587**, 0.561**, 0.537**, 0.651**, 0.560**), while in the latter, only the formal and structural characteristics are greater than 0.5 (0.561**), and the natural element characteristics are about 0.5 (0.498**). This showed that in the campus landscape renewal, providing students with comfortable activity venues and setting up interactive and diverse landscape facilities were conducive to promoting students' emotional experience. Focusing on the correlation calculation results between the detailed landscape features, it could be found that the form and structure were significantly correlated with the materials, colors and natural elements in the landscape space of the Heart of Forest, meanwhile the materials and colors were significantly correlated with natural elements, which indicated that the use of natural materials and the method of expressing the texture of the material itself during the design and construction of landscape was conducive to the establishment of its emotional association with users, such as the wooden pavilion and the interactive space enclosed by rusty steel plates in the Heart of Forest. Meanwhile, it could be found that diversity and materials was significantly correlated, indicating that the combination of various materials that made up the landscape of the Heart of the Forest, such as wood, rusty steel plate, natural stone, and natural elements such as water, plants, pebbles, etc., effectively promoted the formation of emotional attachment. The significant correlation between the uniqueness of and the material, form and structure of the Heart of Forest further indicated that the pavilions and spatial nodes composed of the latter were important driving forces for the formation of uniqueness, and thus enhance the emotional experience of students.

Tables 3, 4 revealed the different mechanisms of emotional attachment to the Heart of Forest among students of different genders. According to the correlation factor data results, the positive emotions of male and female students significantly promoted their emotional attachment to landscape space, while the specific characteristics of landscape promoted the emotions of male students less than that of female students. As for male students, landscape materials and privacy were basically not correlated with their emotional attachment, while color, diversity and changeability were correlated with their emotional attachment but were not as significant as those of female students. Meanwhile, the playability of landscape space in the Heart of Forest have a stronger effect on the emotional attachment of male students than that of female students, and according to the correlation factor between specific landscape features, it can be inferred that the natural elements in the landscape space contribute to the formation of this playful experience. In addition, the color, natural elements and form and structure of the landscape spaces consisting of the Heart of Forest had shown almost no correlation with the sense of territoriality generated by male students, while for female students, the correlation was significant, which also reflected the gender difference of environmental characteristics in promoting the establishment of emotional attachment.

Table 5 described the characteristics and internal mechanism of emotional attachment to the Heart of Forest of students majoring in environment design related majors. Same as the above, these students' overall attachment to the Heart of Forest was positively correlated with positive affect and most of the landscape characteristics. Among them, unlike the total sample data results, privacy and changeability are the two landscape characteristics that are weaker in correlation with the emotional attachment of students studying environmental design related majors than other characteristics. In addition, surprisingly, the material characteristics of the Heart of Forest such as color, material, natural elements, form and structure were not significantly correlated with the positive emotions of design students, especially color and natural elements hardly promoted their positive emotional experiences here. The correlation calculation results between material and social and interactive characteristics presented by the data of this group of students helped us to further explore how the design and construction method of the Heart of Forest affect these

10.3389/fpsyg.2023.1250441

TABLE 2 Correlation coefficients of total participants.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Place attachment	1													
2	Positive affect	0.553**	1												
3	Negative affect	−0.023	0.196*	1											
4	Material	0.445**	0.266**	−0.197*	1										
5	Color	0.459**	0.241**	−0.137	0.678**	1									
6	Natural elements	0.498**	0.253**	−0.252**	0.678**	0.646**	1								
7	Form and structure	0.561**	0.291**	−0.188*	0.642**	0.682**	0.685**	1							
8	Privacy	0.342**	0.224**	−0.058	0.362**	0.295**	0.242**	0.353**	1						
9	Diversity	0.587**	0.414**	−0.139	0.617**	0.515**	0.570**	0.578**	0.443**	1					
10	Sociability	0.561**	0.371**	−0.007	0.340**	0.334**	0.467**	0.497**	0.260**	0.565**	1				
11	Territoriality	0.537**	0.257**	−0.198*	0.502**	0.320**	0.383**	0.407**	0.332**	0.549**	0.416**	1			
12	Playability	0.651**	0.402**	−0.046	0.404**	0.261**	0.349**	0.406**	0.286**	0.570**	0.509**	0.457**	1		
13	Uniqueness	0.560**	0.393**	−0.183*	0.633**	0.565**	0.619**	0.692**	0.371**	0.663**	0.521**	0.529**	0.590**	1	
14	Changeability	0.416**	0.256**	−0.046	0.454**	0.380**	0.384**	0.484**	0.446**	0.512**	0.328**	0.392**	0.445**	0.560**	1

*$p < 0.05$ (2-tailed), **$p < 0.01$ (2-tailed).

TABLE 3 Correlation coefficients of male participants.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Place attachment	1													
2	Positive affect	0.525**	1												
3	Negative affect	−0.087	0.120	1											
4	Material	0.304	−0.026	−0.324*	1										
5	Color	0.368*	0.226	−0.160	0.614**	1									
6	Natural elements	0.459**	0.075	−0.430**	0.566**	0.556**	1								
7	Form and structure	0.450**	0.131	−0.503**	0.452**	0.506**	0.632**	1							
8	Privacy	0.129	0.150	−0.052	0.181	0.167	0.053	0.297	1						
9	Diversity	0.377*	0.233	−0.249	0.471**	0.441**	0.465**	0.424**	0.339*	1					
10	Sociability	0.536**	0.234	−0.255	0.329*	0.452**	0.611**	0.637**	0.212	0.643**	1				
11	Territoriality	0.467**	0.080	−0.286	0.420**	0.119	0.134	0.211	0.040	0.348*	0.217	1			
12	Playability	0.672**	0.341*	−0.286	0.374*	0.288	0.485**	0.306	0.068	0.492**	0.455**	0.511**	1		
13	Uniqueness	0.421**	0.063	−0.475**	0.570**	0.527**	0.491**	0.595**	0.442**	0.596**	0.564**	0.382*	0.626**	1	
14	Changeability	0.345*	0.194	−0.139	0.463**	0.379*	0.349*	0.342*	0.444**	0.487**	0.437**	0.220	0.372*	0.646**	1

*$p < 0.05$ (2-tailed), **$p < 0.01$ (2-tailed).

10.3389/fpsyg.2023.1250441

TABLE 4 Correlation coefficients female participants.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Place attachment	1													
2	Positive affect	0.566**	1												
3	Negative affect	0.040	0.323**	1											
4	Material	0.505**	0.405**	−0.107	1										
5	Color	0.500**	0.251*	−0.172	0.710**	1									
6	Natural elements	0.522**	0.351**	−0.073	0.741**	0.701**	1								
7	Form and structure	0.606**	0.360**	−0.035	0.723**	0.748**	0.731**	1							
8	Privacy	0.440**	0.266**	−0.126	0.452**	0.349**	0.354**	0.368**	1						
9	Diversity	0.661**	0.489**	−0.068	0.677**	0.548**	0.627**	0.636**	0.495**	1					
10	Sociability	0.584**	0.437**	0.109	0.355**	0.289**	0.433**	0.446**	0.265**	0.557**	1				
11	Territoriality	0.562**	0.330**	−0.177	0.537**	0.398**	0.504**	0.472**	0.455**	0.617**	0.483**	1			
12	Playability	0.644**	0.428**	0.164	0.419**	0.258**	0.294**	0.449**	0.386**	0.595**	0.546**	0.441**	1		
13	Uniqueness	0.620**	0.551**	0.106	0.665**	0.591**	0.692**	0.751**	0.350**	0.692**	0.534**	0.593**	0.578**	1	
14	Changeability	0.441**	0.282**	0.001	0.455**	0.382**	0.409**	0.526**	0.451**	0.522**	0.298**	0.445**	0.471**	0.537**	1

*$p < 0.05$ (2-tailed), **$p < 0.01$ (2-tailed).

10.3389/fpsyg.2023.1250441

10.3389/fpsyg.2023.1250441

TABLE 5 Correlation coefficients of participants majoring in environment design related majors.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Place attachment	1													
2	Positive affect	0.534**	1												
3	Negative affect	0.052	0.246*	1											
4	Material	0.475**	0.255*	−0.084	1										
5	Color	0.500**	0.171	−0.131	0.620**	1									
6	Natural elements	0.466**	0.214	−0.179	0.677**	0.620**	1								
7	Form and structure	0.521**	0.261*	−0.066	0.601**	0.569**	0.692**	1							
8	Privacy	0.246*	0.264*	−0.108	0.344**	0.205	0.273*	0.388**	1						
9	Diversity	0.554**	0.401**	−0.069	0.621**	0.447**	0.526**	0.621**	0.528**	1					
10	Sociability	0.471**	0.365**	−0.110	0.473**	0.442**	0.495**	0.630**	0.346**	0.617**	1				
11	Territoriality	0.484**	0.294**	−0.157	0.525**	0.341**	0.361**	0.334**	0.249*	0.422**	0.320**	1			
12	Playability	0.662**	0.449**	0.062	0.421**	0.358**	0.353**	0.466**	0.254*	0.547**	0.469**	0.368**	1		
13	Uniqueness	0.602**	0.456**	−0.113	0.590**	0.538**	0.628**	0.723**	0.329**	0.663**	0.648**	0.496**	0.657**	1	
14	Changeability	0.280*	0.336**	0.016	0.426**	0.199	0.402**	0.397**	0.368**	0.486**	0.423**	0.349**	0.424**	0.589**	1

*$p < 0.05$ (2-tailed), **$p < 0.01$ (2-tailed).

students social and interactive experience. For example, materials, form and structure and landscape diversity showed significant positive correlations, and natural elements, form and structure and landscape uniqueness showed significant positive correlations, indicating to a certain extent that the relevant built environment characteristics that make up the Heart of Forest promoted the design students' experience of its diversity and uniqueness. While the difference between the total sample data and the design students' sample data revealed that professional design education may influence students' emotional experience of the built environment.

4.1.3. The exploratory factor analysis results

Exploratory factor analysis aimed to condense and refine the above spatial factors of landscape characteristics with minimal information loss, and explore the characteristics that play a major role in the emotional attachment process of the landscape characteristics of the Heart of Forest from the bottom up, including both material characteristics and social and interactive characteristics. KMO and Bartlett spherical tests were carried out and the results showed that their structural characteristics were good and met the conditions of factor analysis.

Based on the gravel plot and the rotational component matrix, the main factors (principal components) that played a role in the establishment of emotional attachment to landscape characteristics of the Heart of Forest could be extracted into three (Table 6). The main Factor 1 with the strongest explanatory effect on the establishment of emotional attachment was composed of color, natural elements, form and structure, material, and uniqueness, which mainly describes the built environment characteristics of Heart of Forest's landscape space and its unique positioning in the campus, indicating that Heart of Forest 's landscape space ontology has an important impact on the promotion of users' emotional experience, and the formation of this emotional experience was closely related to its uniqueness. Among them, color and natural elements had the strongest explanatory effect on this main factor, which was consistent with the characteristics of the landscape space of the Heart of Forest: a variety of plants, water bodies, natural material settings and the use of natural colors. The main Factors 2 that have a relatively strong explanatory effect on the establishment of emotional attachment include playfulness, sociability, territoriality and diversity, which described the interesting interaction characteristics and localization characteristics of the Heart of Forest, and revealed that providing interesting social activity opportunities and creating a place spirit exclusive to campus were conducive to the formation of students' emotional attachment. The main Factors 3 that had the weakest explanation for the establishment of emotional associations included privacy and changeability. Among them, the effect of privacy on emotional attachment was weak, which was consistent with the results of the aforementioned correlation analysis and the characteristics of the Heart of Forest as an open campus landscape space. However, changeability did not play an effective role in the formation of emotional attachment, indicating that many interactive devices set up in the Heart of Forest did not play an effective role in promoting interactive activities, which might be closely related to its difficult to be detected in spatial distribution and lack of maintenance in subsequent operations.

TABLE 6 Main factors promoting the establishment of emotional attachment to landscape characteristics of the Heart of Forest.

Specific landscape characteristics	Factor 1	Factor 2	Factor 3
Color	0.958		
Natural elements	0.863		
Form and structure	0.769		
Material	0.769		
Uniqueness	0.460		
Playability		0.868	
Sociability		0.824	
Territoriality		0.573	
Diversity		0.512	
Privacy			0.915
Changeability			0.616

4.2. The results of sentiment analysis

Using a top-down analysis of the landscape spatial characteristics consisting of the Heart of Forest, combining the above scales to explore the emotional attachment strength of specific landscape characteristics, the relevant texts were located and extracted through words associated with landscape material characteristics, social and interactive characteristics, and design characteristics (Table 7). Then, the text of the reviews was subjected to word segmentation, integration, screening, and assignment calculations to obtain the students' emotional attention to and emotional experience intensity in different landscape spaces of the Heart of Forest.

The frequency of texts helped to understand students' emotional attention to different landscape characteristics. In general, the high frequency of text associated with students' feeling and natural element of landscape itself indicated that natural elements in Heart of Forest triggered students' diverse emotional experiences. At the same time, it could be found that in addition to the material elements in the landscape space, experiences related to animals had also become important emotional triggers. Meanwhile, the low frequency of text associated with social events, especially couple dating, indicated the effect of social activities on students' attachment to the Heart of Forest is weaker than that of the landscape itself, which also proved the success of the landscape transformation as the environment was highlighted.

The sentiment and semantic calculation results for the reviews reflected the positive and negative directions of emotional experiences by the positive and negative numerical values, respectively, and reflected the intensity of emotional experiences by the magnitude of the numerical values (Figure 5). In general, the overall emotional intensity values for the semantic evaluations of the landscape characteristics were all positive, indicating that the emotional experience of students in the Heart of Forest was positive and pleasant. And the "Feeling" had the highest scores, which again proved the positive emotional trigger effect of this landscape space. The students intentionally or unintentionally experienced emotional changes in the process of interactions between individuals and sites and were derived from both the material and social and interactive landscape characteristics. This was consistent with the results of the above scale.

TABLE 7 Frequency of representative texts of different landscape characteristics.

Review classification	Related landscape characteristics	Representative text	Frequency
Material	Natural element	Forest	721
		Plant	
		Tree	
		Lawn	
		Birds chirp	
		Insect sound	
		Mosquito	
		Pond	
		Water	
		Swamp	
		Swan lake	
	Time element	Night	362
		Lamplight	
		Spring	
		Autumn	
		Summer	
		Winter	
Social and Interactive	Territoriality	Campus	659
		Memory	
		Old hospital	
		Monuments	
		Campus culture	
	Changeability	Interactive games	458
		Device	
		Sound collection	
		Experience	
		Sensor	
	Sociability	Single	242
		In love	
		Couple	
		Date	
		Party	
		Discuss	
		Delicacies	
	Feeling	Wish	863
		Expression	
		Beauty	
		Nice	
		Emotion	
		Full of life	
		Mysterious	
		Romantic	

(Continued)

TABLE 7 (Continued)

Review classification	Related landscape characteristics	Representative text	Frequency
Design-related	Design strategy	Landscape	397
		Space	
		Designer	
		Gardens	
		Create	

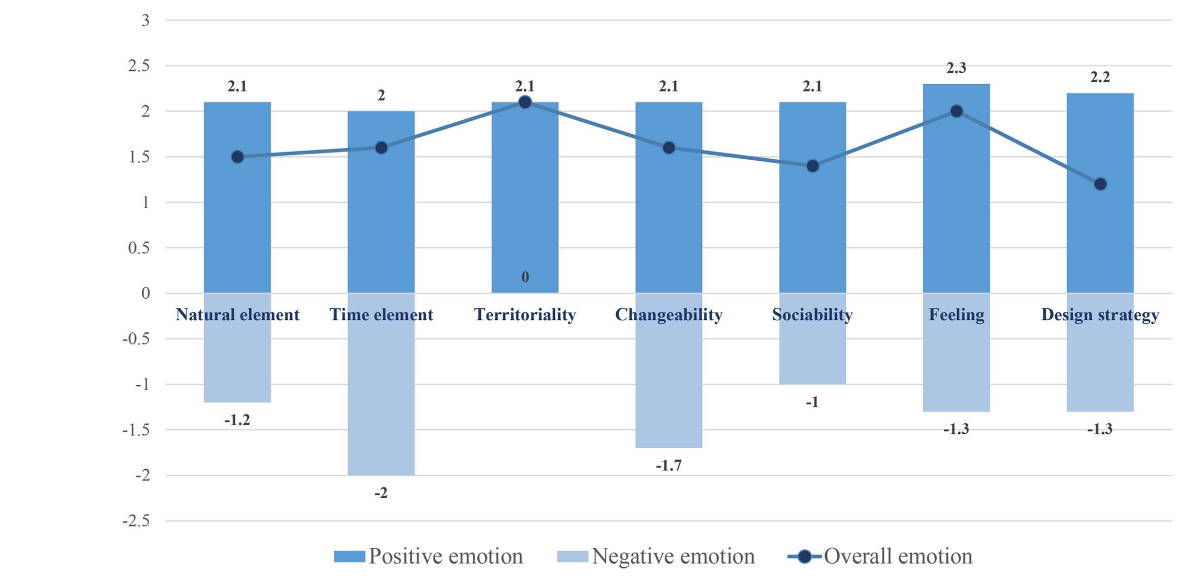


FIGURE 5
The semantic calculation results.

"Feelings" had the highest positive score, and according to its representative texts, it could be inferred that the beauty and romantic atmosphere of landscape characteristics triggered a strong emotional experience during students' studying in the Heart of Forest. "Design Strategy" also received a relatively high score, indicating that the creation of the landscape had played a positive role in promoting the emotional experience of the students, which complemented a part not mentioned in the previous scale. Meanwhile, the frequent occurrence of the word "designer" in this part once again proved that Professor Cai, who designed the Heart of Forest, had promoted many students' love for this landscape with his personal charm, which was also consistent with the results of the qualitative interview latter. "Territoriality" had also got a high positive sentiment score and is the only one characteristic without a negative sentiment score. According to its representative texts, the creation of the school's history, memory and culture-related landscape spaces had become an important emotional anchor point, which once again revealed the importance of the spiritual value of the campus landscape as a "place" to the establishment of students' emotional attachment. This result also complemented the details not described in the territoriality section of the aforementioned scale. This further confirmed the value of the design concept of "Heart of Forest": to carry and reproduce campus memories through landscape renewal.

As for negative emotional semantics analysis results, "Time element" and "Changeability" had relatively high negative scores. Combined with observation and in-depth interviews, for the former, season-related mosquitoes and water accumulation caused by heavy rainfall were the main causes of negative emotions. For the latter, the physical and sound interaction facilities originally set up by the designer could not be used normally due to lack of maintenance, which brought a negative experience from the interactive and changeable aspects of the Heart of Forest.

4.3. The supplementation and corroboration of interview and observation

The qualitative research method based on observation method and interview method further excavates the emotional triggering characteristics of landscape space that are not revealed by the aforementioned scale data and big data by observing the interaction between students and space at different times in the Heart of Forest and conducting in-depth interviews with students, and at the same time interprets and explains the data results. The qualitative results showed that, overall, the design of the Heart of Forest was relatively successful in terms of promoting emotional attachment. The lawn bathed in sunlight, the pleasant birdsong, the growing trees, and the water surface of the pond were all widely loved by students. Diverse interactive facilities and public spaces for dating and dining had greatly enriched students' campus life, especially during the COVID-19. In addition, the different landscape styles of the four seasons and day and night of the venue also left users with rich time and space memories.

Qualitative research had also further revealed some problems that site design might trigger negative emotions among users. For example, the most frequently cited problem was the excessive mosquito insect in the summer venue, which created a negative experience for students who wanted a leisure experience. Secondly, although the facilities in the site were abundant, but such as the specimen wall of "Secret Language in the Forest" and some lighting facilities of the site were often damaged, which brought great inconvenience to students. Most students were unfamiliar with the use of interactive facilities such as the "central light column," resulting in the lack of interactive experience that landscape spaces could provide. In addition, due to the climate of Beijing, the pavilions of the Heart of Forest were less windproof in winter, and the large cold wind and cold climate prevented students from getting a good experience in winter. A small number of students who majored in landscape related majors also reflected the problem of insufficient uniqueness of the landscape, and looked forward to more diversified innovations in design forms. Some students also put forward higher requirements for the balance of landscape places and self-study spaces, as well as insufficient privacy, water accumulation on rainstorm days, and insufficient plant maintenance and cleaning. The interviews had also revealed some new characteristics of the site that promoting students' positive emotions. For example, we found that purposeful activities could effectively enhance students' emotional attachment to the landscape space of the Heart of Forest. In addition, the specimen wall in the "Museum in the Forest" and the diverse plant planting in the landscape made the Heart of Forest a new place for teaching practice in the garden and plant related courses, creating a variety of use scenarios for the site.

On the one hand, the results once again support the calculation results of the aforementioned scale and big data, and on the other hand, provide more detailed descriptions of the spatial characteristics of the landscape that promote emotional experience. You may insert up to 5 heading levels into your manuscript as can be seen in "Styles" tab of this template. These formatting styles are meant as a guide, as long as the heading levels are clear, Frontiers style will be applied during typesetting.

5. Discussion

The study combines the top-down emotional attachment scale and bottom-up public semantic analysis to collect multi-dimensional emotional interaction data between people and the Heart of Forest to study and evaluate the emotional association characteristics of this landscape space. Observations and interviews are used throughout the data collection and analysis process to complement and interpret quantitative results. The research reveals the factors that promote or weaken students' emotional attachment to the Heart of Forest from different aspects including material, social and interactive characteristics of the landscape. On the one hand, the result illustrates the positive emotional promotion value of naturalized landscape to college students and the healing effect of nature, especially under the background of COVID-19 epidemic. And on the other hand, it also proves that positive emotional experience is conducive to strengthening students' emotional attachment to the campus environment, so as to help students obtain stronger place identity. In general, the research results not only support the existing research, but also provide new references for the study of emotional attachment between people and the environment from both the theoretical and methodological level.

At the theoretical level, consistent with previous studies, the healing effect of natural landscapes on negative emotions and the triggering effect of positive emotions have been confirmed again (Kaplan and Kaplan, 1989; Kaplan, 1995). Specifically, nature-related elements, materials, structures and spaces that are easy to get close to nature have important value in providing a good emotional experience. It is achieved by intentionally or unintentionally triggering students to interact with nature, such as watching, touching, listening, etc. (Lee et al., 2009; Koga and Iwasaki, 2013; Honold et al., 2016; Ratcliffe, 2021; Rickard and White, 2021). From the perspective of the material composition of space design and construction, the material characteristics of the site play a positive role in promoting the vast majority of students in the process of establishing emotional attachment (Zhang, 2021; Zhang, 2023). Besides, what's new of this study is that this promotion effect is particularly obvious for students majoring in landscape, architecture and urban planning, while students in other majors are relatively more interested in the social and interactive characteristics of the Heart of Forest, such as playfulness and changeability as important dimensions for them to gain emotional attachment. It reveals that the emotional experience provided by space is closely related to students' understanding and cognition. This emotion may be also closely related to the admiration of the designer himself from students studying design-related majors.

Meanwhile, the social characteristics of landscape space in the establishment of emotional attachment is influenced by more subjective and objective factors outside the built environment. This is consistent with the value of the social dimension on attachment establishment in the aforementioned study of place attachment (Scannell and Gifford, 2010; Lewicka, 2011; Lin et al., 2019). For example, whether the public social space created by the designer can effectively provide students with a good emotional experience is closely related to the frequency and purpose of their use of the space (Liu et al., 2020; Arnberger et al., 2022). In addition, for the specific group of college students, multimodal data further illustrate that activities related to teaching and research are not as strong as leisure and entertainment-related activities such as dating and parties to promote place attachment, the negative emotions caused by them are often more obvious, indicating that academic pressure may affect the emotional healing value of landscape space.

In addition, compared with existing research (Birnbaum et al., 2021), this study also reveals the possibility and effectiveness of using new technologies to enhance the emotional attachment between people and the landscape environment in the digital age. For the interactive characteristics such as changeability and playability in the Heart of Forest, the results show that interesting interactive facilities and digital interactive installations originally set up in the design fail to provide students with a positive emotional experience as expected, but become a factor that triggers negative emotions due to the lack of maintenance and complicated use in the later stage. This also shows that before intervening in emerging technologies and new space types in campus landscape space design, it is necessary to conduct more detailed research and consideration of the installation itself and the interactive experience it can provide.

At the methodological level, with the support of big data technology in the digital age (Barbier and Liu, 2011; Chakraborty et al., 2020; Jindal and Aron, 2021), this study updates and improves the research logic dominated by the qualitative dominance of traditional emotional attachment by integrating multimodal data, providing a more scientific, objective and generalizable method (Zhang, 2023). Through the coupling analysis of top-down scale data, bottom-up big data mining and public semantic analysis, and the qualitative results obtained from in-depth interviews and observations, it can be found that overall, the data obtained in different dimensions are mostly consistent, and the different aspects of emotional attachment between people and landscape space are emphasized. The scale data is collected by researchers based on the theory of emotional attachment and purposefully constructed scales according to the research question (Adams, 2015), and its analysis logic is relatively systematic, which is the main research method to explore the mechanism of emotional attachment of this study. The results of qualitative studies are in some ways more in-depth and specific and can provide interpretation and supplement to the scale data. Big data analysis is bottom-up (Wilkins et al., 2020; Zhang, 2021), although the impact of the researcher effect on the data results is avoided, but the content involved in the data is relatively broad, focusing more on the overall evaluation of students' emotional experience, so it is necessary to combine the description of the spatial characteristics of the landscape in the aforementioned scale to carry out further semantic analysis.

The study improves on the shortcomings in the current post occupancy evaluation of campus renewal spaces especially from the perspective of students' emotional attachment (Feng et al., 2021; Zhang et al., 2022). The results also provide references for the future design and policy practice. Meanwhile, the study avoids the disadvantages of traditional research methods which are always limited to few data, low efficiency and high costs (Ekkekakis, 2013). It promotes the convergence of bottom-up and top-down quantitative methodological logic while enables large-scale automatic evaluation of landscape spatial quality and public sentiment with the help of sentiment analysis and other technologies. The multi-dimensional analysis and evaluation can reveal the ability of specific landscape characteristics that help promote students' emotional attachment more completely.

However, there are still limitations to this study which can be promoted in the future. For example, the completeness and rationality of the scale need to be further improved. The landscape characteristics consisting of the emotional attachment scale are not fully representative of the space and may affect the data results. Furthermore, in addition to students, faculty, staffs and their families on campus should also be considered as an important part of landscape users. The number of the scale collected should be expanded after the impact of COVID-19. Also, while the study has tried to screen users who identify as students as data sources, Sina Weibo and Little Red Book (XiaoHongShu) users may also include people other than students, which has an impact on the representativeness and accuracy of public sentiment results in the Heart of Forest. And the choice of social media platform also affects the results of sentiment semantic analysis. Therefore, we will update the emotional attachment scale in the future to include more spatial factors that may influence people's emotional experience. Meanwhile, we will focus on more diverse population on campus. The collection of SMD will expand to more platforms and more dimensions to extract a more accurate picture of public attitudes toward renovated campus landscape.

6. Conclusion

The research characterizes the mechanisms of students' emotional attachment to different landscape dimensions and spatial characteristics of the "Heart of Forest" through emotional attachment scale and sentiment analysis of text from Sina Weibo and Little Red Book. Based on the correlation and exploratory factor analysis of scale data, semantic analysis of big data and phenomenological analysis of qualitative data, the factors that mainly trigger students' emotional attachment to the landscape space are revealed. Meanwhile, the landscape characteristics and their values that specifically affect the intensity and dimension of emotional attachment are identified. The coupling analysis of multi-dimensional data further reveals the differences and specific characteristics of different landscape components in emotional attachment for students of different majors and genders. The research results once again prove that natural landscape still has irreplaceable emotional value and healing effect in the information age, and at the same time, it also improves the shortcomings of classical research in dealing with new problems in the digital age at the methodological level. Its main values are manifested in the following three aspects: (1) The real natural environment and materials are still of great importance in the construction and renewal of the public environment of colleges and universities, especially for the mental health and positive emotions of college students. Digitalization has not diminished its existential value, but has highlighted its preciousness. Multimodal data of the study objectively proved that the material characteristics of the landscape space of the Heart of Forest play the most important role in the formation of emotional attachment. Among them, natural materials, colors, structures and biodiversity have the most significant effect. (2) The emotional value of the social characteristics of landscape space in colleges and universities is closely related to whether it can provide students with the activity space that students subjectively demand. When the social activities carried by the space are actively chosen by students, it can often positively promote students' positive emotional attachment to the space itself. Compared with academic activities, leisure activities have a more positive effect on the establishment of emotional attachment, and the higher the frequency of activities, the deeper the degree of emotional attachment. (3) As for the interactive characteristics of the renewal landscape, whether they can provide students with interesting and gamified space experience as expected by the designer to enhance students' positive emotional is closely related to the comprehensibility, ease of operation and subsequent maintenance of the interactive devices.

Overall, the renovation of the Heart of Forest has been relatively successful in providing a good emotional experience for students, promoting their emotional attachment to school, and becoming an important emotional healing place for them especially during the COVID-19 epidemic. Due to the specificity of the research subject, the methodological framework and analysis process in this study can be used as a reference for the study of renewal and improvement of landscape environment and public space of other colleges and universities, with policy and practical guidance.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

RZ responsible for conceptualization, methodology, software, resources, data curation, writing, visualization, and funding acquisition.

Funding

This work was supported by the National Natural Science Foundation of China under Grant No. 52208005 and Beijing Social Science Foundation under Grant No. 22GLC063.

Acknowledgments

Thanks for Shujie Zhao's help with the scale measurement. Thanks for Yang Song, Weiyue Duan, Zhikai Zheng, Yuling Liu for their help of distributing and collecting scales. Thanks for Weiyue Duan and Zhikai Zhao for their photography of the Heart of Forest.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adams, P. C. (2015). Place attachment: advances in theory, methods and applications. *Geogr. Rev.* 105, 129–132. doi: 10.1111/j.1931-0846.2014.12061.x

Altman, I., and Low, S. M. *Place attachment*: Berlin: Springer Science & Business Media. (2012).

Andre, E. K., Williams, N., Schwartz, F., and Bullard, C. (2017). Benefits of campus outdoor recreation programs: a review of the literature. *J. Outdoor Recreat. Educ. Leadersh.* 9:7491. doi: 10.18666/JOREL-2017-V9-I1-7491

Arnberger, A., Budruk, M., Schneider, I. E., and Stanis, S. A. W. (2022). Predicting place attachment among walkers in the urban context: the role of dogs, motivations, satisfaction, past experience and setting development. *Urban For. Urban Green.* 70:127531. doi: 10.1016/j.ufug.2022.127531

Barbier, G., and Liu, H. (2011). "Data mining in social media," in *Social network data analytics*. United States: Springer, 327–352.

Birnbaum, L., Wilhelm, C., Chilla, T., and Kröner, S. (2021). Place attachment and digitalisation in rural regions. *J. Rural. Stud.* 87, 189–198. doi: 10.1016/j.jrurstud.2021.09.015

Biswas, R., and Sen, S. (2020). Urban eco-psychological attitude during COVID-19 'lockdown': a survey. *Int. J. Creat. Res. Thoughts* 8

Brandisauskiene, A., Buksnyte-Marmiene, L., Cesnaviciene, J., Daugirdiene, A., Kemeryte-Ivanauskiene, E., and Nedzinskaite-Maciuniene, R. (2021). Sustainable school environment as a landscape for secondary school students' engagement in learning. *Sustainability* 13:11714. doi: 10.3390/su132111714

Brindley, P., Cameron, R. W., Ersoy, E., Jorgensen, A., and Maheswaran, R. (2019). Is more always better? Exploring field survey and social media indicators of quality of urban greenspace, in relation to health. *Urban For. Urban Green.* 39, 45–54. doi: 10.1016/j.ufug.2019.01.015

Carmona, M. (2019). Place value: place quality and its impact on health, social, economic and environmental outcomes. *J. Urban Des.* 24, 1–48. doi: 10.1080/13574809.2018.1472523

Chakraborty, K., Bhattacharyya, S., and Bag, R. (2020). A survey of sentiment analysis from social media data. *IEEE Trans. Comput. Soc. Syst.* 7, 450–464. doi: 10.1109/TCSS.2019.2956957

Ekkekakis, P. *The measurement of affect, mood, and emotion: A guide for health-behavioral research*. Cambridge: Cambridge University Press. (2013).

Feng, Y., Wang, K., Zhang, Y., and Li, X. (2021). An overview of research on rural public space and place attachment: concepts, logics, and associations. *Chin. Landsc. Archit.* 37, 31–36. doi: 10.19775/j.cla.2021.02.0031

Guo, W., Ding, Y., Yang, G., and Liu, X. (2022). Research on the indicators of sustainable campus renewal and reconstruction in pursuit of continuous historical and regional context. *Buildings* 12:1508. doi: 10.3390/buildings12101508

Guo, X., Tu, X., Huang, G., Fang, X., Kong, L., and Wu, J. (2022). Urban greenspace helps ameliorate People's negative sentiments during the COVID-19 pandemic: the case of Beijing. *Build. Environ.* 223:109449. doi: 10.1016/j.buildenv.2022.109449

Honold, J., Lakes, T., Beyer, R., and van der Meer, E. (2016). Restoration in urban spaces: nature views from home, greenways, and public parks. *Environ. Behav.* 48, 796–825. doi: 10.1177/0013916514568556

Jiang, B., He, J., Chen, J., Larsen, L., and Wang, H. (2021). Perceived green at speed: a simulated driving experiment raises new questions for attention restoration theory and stress reduction theory. *Environ. Behav.* 53, 296–335. doi: 10.1177/0013916520947111

Jindal, K., and Aron, R. (2021). A systematic study of sentiment analysis for social media data. *Materials Today*. doi: 10.1016/j.matpr.2021.01.048

Kaplan, S. (1995). The restorative benefits of nature: toward an integrative framework. *J. Environ. Psychol.* 15, 169–182. doi: 10.1016/0272-4944(95)90001-2

Kaplan, R., and Kaplan, S.. *The experience of nature: A psychological perspective*. Cambridge: Cambridge University Press (1989).

Kaplan, R., Kaplan, S., and Ryan, R.. *With people in mind: design and management of everyday nature*. Washington, DC: Island Press. (1998).

Kara, B. (2013). Landscape design and cognitive psychology. *Procedia Soc. Behav. Sci.* 82, 288–291. doi: 10.1016/j.sbspro.2013.06.262

Kellert, S. R., and Wilson, E. O.. *The biophilia hypothesis*. Washington, DC: Island Press (1995).

Koga, K., and Iwasaki, Y. (2013). Psychological and physiological effect in humans of touching plant foliage-using the semantic differential method and cerebral activity as indicators. *J. Physiol. Anthropol.* 32, 1–9. doi: 10.1186/1880-6805-32-7

Kong, L., Liu, Z., Pan, X., Wang, Y., Guo, X., and Wu, J. (2022). How do different types and landscape attributes of urban parks affect Visitors' positive emotions? *Landsc. Urban Plan.* 226:104482. doi: 10.1016/j.landurbplan.2022.104482

Lai, K. Y., Webster, C., Kumari, S., and Sarkar, C. (2020). The nature of cities and the COVID-19 pandemic. *Curr. Opin. Environ. Sustain.* 46, 27–31. doi: 10.1016/j.cosust.2020.08.008

Lau, S. S., and Yang, F. (2009). Introducing healing gardens into a compact university campus: design natural space to create healthy and sustainable campuses. *Landsc. Res.* 34, 55–81. doi: 10.1080/01426390801981720

Lee, J., Park, B.-J., Tsunetsugu, Y., Kagawa, T., and Miyazaki, Y. (2009). Restorative effects of viewing real Forest landscapes, based on a comparison with urban landscapes. *Scand. J. For. Res.* 24, 227–234. doi: 10.1080/02827580902903341

Lewicka, M. (2011). Place attachment: how far have we come in the last 40 years? *J. Environ. Psychol.* 31, 207–230. doi: 10.1016/j.jenvp.2010.10.001

Lin, G., Wu, A., and Cai, K. (2019). Research on place attachment: concept, Progress and trend. *Chin. Landsc. Archit.* 35, 63–66. doi: 10.19775/j.cla.2019.10.0063

Liu, S., Ji, Y., Li, J., Peng, Y., Li, Z., Lai, W., et al. (2022). Analysis of Students' positive emotions around the green space in the university campus during the COVID-19 pandemic in China. *Front. Public Health* 10:10. doi: 10.3389/fpubh.2022.888295

Liu, Q., Wu, Y., Xiao, Y., Fu, W., Zhuo, Z., van den Bosch, C. C. K., et al. (2020). More meaningful, more restorative? Linking local landscape characteristics and place attachment to restorative perceptions of Urban Park visitors. *Landsc. Urban Plan.* 197:103763. doi: 10.1016/j.landurbplan.2020.103763

McPherson, E. G. (1992). Accounting for benefits and costs of urban greenspace. *Landsc. Urban Plan.* 22, 41–51. doi: 10.1016/0169-2046(92)90006-L

Qiangua's 2022 Xiaohongshu Active User Portrait Trend Report. (2022). Available at: https://www.qian-gua.com/information/detail/1591 (Accessed June 10, 2023).

Ratcliffe, E. (2021). Sound and soundscape in restorative natural environments: a narrative literature review. *Front. Psychol.* 12:570563. doi: 10.3389/fpsyg.2021.570563

Rickard, S. C., and White, M. P. (2021). Barefoot walking, nature connectedness and psychological restoration: the importance of stimulating the sense of touch for feeling closer to the natural world. *Landsc. Res.* 46, 975–991. doi: 10.1080/01426397.2021.1928034

Rousseau, S., and Deschacht, N. (2020). Public awareness of nature and the environment during the COVID-19 crisis. *Environ. Resour. Econ.* 76, 1149–1159. doi: 10.1007/s10640-020-00445-w

Scannell, L., and Gifford, R. (2010). Defining place attachment: a tripartite organizing framework. *J. Environ. Psychol.* 30, 1–10. doi: 10.1016/j.jenvp.2009.09.006

Scannell, L., and Gifford, R. (2013). Personally relevant climate change: the role of place attachment and local versus global message framing in engagement. *Environ. Behav.* 45, 60–85. doi: 10.1177/0013916511421196

Scannell, L., and Gifford, R. (2017). The experienced psychological benefits of place attachment. *J. Environ. Psychol.* 51, 256–269. doi: 10.1016/j.jenvp.2017.04.001

Sim, J., Miller, P., and Swarup, S. (2020). Tweeting the high line life: a social media Lens on urban green spaces. *Sustainability* 12:8895. doi: 10.3390/su12218895

Sina Weibo. (2020). Annual Sina Weibo User Report 2020. Available at: https://data.weibo.com/report/reportDetail?id=456 (Accessed June 11, 2023).

Tuan, Y.-F. (1979). "Space and place: humanistic perspective" in *Philosophy in geography*. eds. S. Gale and G. Olsson (Berlin: Springer), 387–427.

Van den Berg, A.. Restorative effects of nature: towards a neurobiological approach. Proceedings of the 9th International Congress of Physiological Anthropology. (2008).

Van Leeuwen, E., Nijkamp, P., and de Noronha, V. T. (2010). The multifunctional use of urban greenspace. *Int. J. Agric. Sustain.* 8, 20–25. doi: 10.3763/ijas.2009.0466

Vella-Brodrick, D. A., and Gilowska, K. (2022). Effects of nature (greenspace) on cognitive functioning in school children and adolescents: a systematic review. *Educ. Psychol. Rev.* 34, 1217–1254. doi: 10.1007/s10648-022-09658-5

Wilkins, E. J., Smith, J. W., and Keane, R. (2020). Social media communication preferences of National Park Visitors. *Appl. Environ. Educ. Commun.* 19, 4–18. doi: 10.1080/1533015X.2018.1486247

Williams, D. R., and Roggenbuck, J. W. Measuring place attachment: Some preliminary results. NRPA Symposium on Leisure Research, San Antonio, TX. (1989).

Xie, J., Luo, S., Furuya, K., and Sun, D. (2020). Urban parks as green buffers during the COVID-19 pandemic. *Sustainability* 12:6751. doi: 10.3390/su12176751

Yan, T., Jin, H., and Jin, Y. (2023). The mediating role of emotion in the effects of landscape elements on thermal comfort: a laboratory study. *Build. Environ.* 233:110130. doi: 10.1016/j.buildenv.2023.110130

Ye, L., Peng, X., Aniche, L. Q., Scholten, P. H., and Ensenado, E. M. (2021). Urban renewal as policy innovation in China: from growth stimulation to sustainable development. *Public Adm. Dev.* 41, 23–33. doi: 10.1002/pad.1903

Yi, Z., Liu, G., Lang, W., Shrestha, A., and Martek, I. (2017). Strategic approaches to sustainable urban renewal in developing countries: a case study of Shenzhen, China. *Sustainability* 9:1460. doi: 10.3390/su9081460

You, S., Zheng, Q., Chen, B., Xu, Z., Lin, Y., Gan, M., et al. (2022). Identifying the spatiotemporal dynamics of Forest ecotourism values with remotely sensed images and social media data: a perspective of public preferences. *J. Clean. Prod.* 341:130715. doi: 10.1016/j.jclepro.2022.130715

Yuanshuo, X. (2023). Renewal strategy of learning space in colleges and universities under the digital trend: a case study of old library of North China University of Technology. *J. Landsc. Res.* 15

Zeng, X., Zhong, Y., Yang, L., Wei, J., and Tang, X. (2022). Analysis of Forest landscape preferences and emotional features of Chinese Forest recreationists based on deep learning of geotagged photos. *Forests* 13:892. doi: 10.3390/f13060892

Zhang, R. (2021). Research on affective-oriented brutalist architecture:two cases of campus libraries of the University of California. *New Archit.* 2, 62–67.

Zhang, R. (2023). Integrating ergonomics data and emotional scale to analyze People's emotional attachment to different landscape features in the Wudaokou Urban Park. *Front. Archit. Res.* 12, 175–187. doi: 10.1016/j.foar.2022.06.007

Zhang, R., Dai, Y., and Sun, Z. (2022). Research and evaluation of the mountain settlement space based on the theory of "Flânuer" in the digital age —— taking Yangchan Village in Huangshan City, Anhui Province, as an example. *J. Asian Archit. Build. Eng.* S1, 1–17. doi: 10.1080/13467581.2023.2217884

Zhu, X., Gao, M., Zhang, R., and Zhang, B. (2021). Quantifying emotional differences in urban green spaces extracted from photos on social networking sites: a study of 34 parks in three cities in northern China. *Urban For. Urban Green.* 62:127133. doi: 10.1016/j.ufug.2021.127133

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 02 November 2023
DOI 10.3389/fpsyg.2023.1280213

OPEN ACCESS

EDITED BY
Ivonne Castiblanco Jimenez,
Polytechnic University of Turin, Italy

REVIEWED BY
Elena Carlotta Olivetti,
Polytechnic University of Turin, Italy

*CORRESPONDENCE
Zhihui Zhang
✉ zhihui.zhang@upc.edu

RECEIVED 19 August 2023
ACCEPTED 17 October 2023
PUBLISHED 02 November 2023

CITATION
Zhang Z, Giménez Mateu L and Fort JM (2023) Apple Vision Pro: a new horizon in psychological research and therapy. *Front. Psychol.* 14:1280213. doi: 10.3389/fpsyg.2023.1280213

Apple Vision Pro: a new horizon in psychological research and therapy

Zhihui Zhang*, Lluis Giménez Mateu and Josep M. Fort

Escola Tècnica Superior d' Arquitectura de Barcelona, Universitat Politècnica de Catalunya, Barcelona, Spain

KEYWORDS

Apple Vision Pro, virtual reality, telepsychotherapy, emotion recognition, psychological experiments

Virtual Reality (VR) harbors immense potential for advancing psychological therapy and emotional research, despite presenting several challenges (Riva, 2022). The release of Apple Vision Pro has unveiled new opportunities in the realm of psychological emotion research, particularly with its facial expression system enabling facial emotion recognition within a virtual reality environment. We will elucidate the new perspectives Apple Vision Pro brings to Psychological Research and Therapy by delving into its Multi-Sensor Technology, High Resolution, and Remote Scene Meeting Capabilities.

Multi-sensor advancements for VR facial emotion recognition

Measuring emotions in virtual reality (VR) has mainly involved using electroencephalogram (EEG) devices (Suhaimi et al., 2020; Pinilla et al., 2021). However, the conflict between the mobility associated with VR experiences and the stillness required for reliable EEG measurements casts doubts on the viability of EEG-based VR emotion measurement. The limitations of EEG-based emotion measurement in VR have led to the exploration of alternative methods for more reliable emotion recognition. One such alternative emerged with the introduction of facial recognition technologies, which promise a less obtrusive and more natural means of gauging emotional responses compared to EEG. The pivotal moment arrived with the debut of Fove 0 in 2017, which pioneered the incorporation of eye-tracking technology in VR, marking the onset of a new era in VR Emotion Recognition technology. The advent of eye-tracking provided a fresh perspective and an additional layer of data to better understand users' emotional responses within virtual environments. Since then, a plethora of devices, each with their distinct features, capabilities, and limitations for research applications, have emerged, as detailed in Table 1.

However, the reliance on eye-tracking alone has shown intrinsic limitations, particularly in capturing the full spectrum of emotional and psychological reactions (Levitan et al., 2022). The intersection of facial recognition technology and VR is gaining traction, aiming to provide a more nuanced understanding of emotional responses within virtual realms. Devices like Meta Quest Pro attempted to integrate facial recognition technologies but faced technical and applicative constraints, limiting their utility in comprehensive emotion recognition research. In contrast, the 2023 launch of Apple Vision Pro represents a significant leap forward by seamlessly combining high-fidelity sensor technology with facial recognition, unlocking new potential for emotion recognition in VR. Uniquely, Apple Vision Pro utilizes individualized facial scans of each user, as opposed to the generalized facial simulation technology employed by other notable devices like those from Meta and Vivo (Zhang et al., 2023). With an intricate array of 12 cameras and 5 sensors, Apple Vision Pro ensures nuanced tracking of facial expressions, presenting a detailed framework for emotion detection. This robust sensor framework not only ensures accurate and personalized facial recognition and expression transmission, marking a substantial advancement in VR facial emotion recognition.

Enhanced realism for experimental and therapeutic applications

Traditional psychological experiments often grapple with the challenge of ecological validity due to the artificial laboratory settings. The Apple Vision Pro, boasting a superior resolution of 3680x3140, offers remarkable scene simulation capabilities, enabling more realistic and immersive environments for both experimental and therapeutic applications. This heightened realism, facilitated by the device's high-definition rendering, paves the way for more authentic and valid research outcomes. For instance, in exposure therapy, therapists can harness the Apple Vision Pro to simulate real-life scenarios in a controlled yet authentic setting, aiding patients in gradually confronting and overcoming their fears (Carlin et al., 1997; Krijn et al., 2004; Barrett et al., 2023). The enhanced resolution and field of view not only deepen users' immersion in highly detailed and broad virtual environments but also significantly improve the ecological validity of psychological and emotional research within VR, heralding a promising trajectory for future innovations and applications in VR-based emotional and behavioral studies.

Potential of 3D rendering and facial expressions in telepsychotherapy

Telepsychotherapy has brought convenience to those unable to undergo face-to-face therapy. However, traditional methods like video calls fall short in terms of realism and immersive experience (Kocsis and Yellowlees, 2018; Poletti et al., 2020; Rosen et al., 2020). The introduction of Apple Vision Pro has revolutionized this domain. Its capability to render 3D scenes and characters, coupled with the realistic portrayal of facial expressions, offers immense potential in remote psychotherapy. Such detailed representation allows for a genuine, immersive therapeutic environment, providing a tangible interactive space for both therapists and patients. The realistic depiction of facial expressions, critical for psychotherapy, enables therapists to gauge patients' emotions and responses accurately, thus potentially enhancing the effectiveness of remote therapeutic interventions (Wiederhold and Wiederhold, 2009).

TABLE 1 Comparison of technical specifications and tracking capabilities of virtual reality devices.

Release year	Name	Positional tracking	Resolution	Refresh rate	Tracking capability
2017	Fove 0	Yes	1,280×1,440	70 Hz	Eye-tracking
2017	Deepoon VR E3	Yes (360° laser room-scale positioning)	1,280×1,440	70 H	Eye-tracking
2018	VRgineers XTAL	Yes (AR Tracking, Optitrack and Lighthouse)	2,560×1,440	70 Hz	Leap Motion hand-tracking, eye-tracking
2018	StarVR One	Inside-out markered (Lighthouse)	1,830×1,464	90 Hz	Tobii eye tracking
2019	Varjo VR-1	Yes	1,920×1,080	60 Hz	Eye-tracking
2021	Varjo Aero	Yes	2,880×2,720	90 Hz	Eye-tracking
2019	HTC Vive Cosmos	Yes	1,440×1,700	90 Hz	ye-tracking, Mouth tracking(VIVE Facial Tracker 1camera)
2021	Varjo VR-3	Yes	2,880×2,720	90 Hz	Eye-tracking
2022	Meta Quest Pro	Inside-out	1,800×1,920	90 Hz	Eye tracking, Face tracking, Mouth tracking (5 external (one color passthrough), 5 internal sensors)
2023	PlayStation VR2	Yes	2,000×2,040	90 Hz	Eye-tracking
2023	Apple Vision Pro	Inside-out	3,680×3,140	90 Hz	Eye tracking, Face tracking, Mouth tracking (12 cameras, 5 sensors)

Confronting ethical and operational challenges

Apple's products are known for their maturity and stability. The Apple Vision Pro, despite its high performance, comes with a hefty price tag, making it unaffordable for many. Additionally, the use of individualized facial scans presents two challenges: increased operational difficulty and time in conducting related research experiments, and privacy concerns compared to using virtual avatars. Addressing these issues while capitalizing on the technological advancements of VR emotion recognition remains a critical task for the industry.

Conclusion

The Apple Vision Pro serves as a beacon of advancement in psychological research and therapy, integrating high-fidelity sensor technology and realistic 3D rendering. Looking ahead, further technological enhancements of Apple Vision Pro might include more precise emotion recognition algorithms and real-time data analytics, offering refined tools for understanding human emotions and behavior in virtual environments. The potential integration of these advancements could bolster telepsychotherapy and exposure therapy, among other therapeutic applications, providing more personalized and effective treatment solutions.

The technological prowess of Apple Vision Pro not only addresses current challenges in the field but also lays a foundation for future exploration. As the realms of technology and psychological sciences continue to merge, the potential for new, innovative methodologies in therapy and research expands. The Apple Vision Pro represents a significant stride toward a tech-driven era in psychological research and therapy, with its potential to unveil deeper insights into the human psyche and contribute to the betterment of mental health services worldwide.

Author contributions

ZZ: Conceptualization, Writing—original draft, Writing—review & editing. JF: Methodology, Supervision, Writing—review & editing. LG: Methodology, Supervision, Writing—review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Barrett, A., Pack, A., Monteiro, D., and Liang, H. N. (2023). Exploring the influence of audience familiarity on speaker anxiety and performance in virtual reality and real-life presentation contexts. *Behav. Inf. Technol.* 1–13. doi: 10.1080/0144929X.2023.2186145

Carlin, A. S., Hoffman, H. G., and Weghorst, S. (1997). Virtual reality and tactile augmentation in the treatment of spider phobia: a case report. *Behav. Res. Ther.* 35, 153–158. doi: 10.1016/S0005-7967(96)00085-X

Kocsis, B. J., and Yellowlees, P. (2018). Telepsychotherapy and the therapeutic relationship: Principles, advantages, and case examples. *Telemed J. E. Health* 24, 329–334. doi: 10.1089/tmj.2017.0088

Krijn, M., Emmelkamp, P. M., Olafsson, R. P., and Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: a review. *Clin. Psychol. Rev.* 24:259–281. doi: 10.1016/j.cpr.2004.04.001

Levitan, C. A., Rusk, I., Jonas-Delson, D., Lou, H., Kuzniar, L., Davidson, G., et al. (2022). Mask wearing affects emotion perception. *Iperception*. 13, 91. doi: 10.1177/20416695221107391

Pinilla, A., Garcia, J., Raffe, W., Voigt-Antons, J. N., Spang, R. P., and Möller, S. (2021). Affective visualization in virtual reality: an integrative review. *Fron. Virt. Real.* 2, 630731. doi: 10.3389/frvir.2021.630731

Poletti, B., Tagini, S., Brugnera, A., Parolin, L., Pievani, L., Ferrucci, R., et al. (2020). Telepsychotherapy: a leaflet for psychotherapists in the age of covid-19. A review of the evidence. *Couns. Psychol. Q.* 34, 345–347. doi: 10.1080/09515070.2020.1769557

Riva, G. (2022). Virtual reality in clinical psychology. *Compreh. Clini. Psychol.* 91, 6. doi: 10.1016/B978-0-12-818697-8.00006-6

Rosen, C. S., Glassman, L. H., and Morland, L. A. (2020). Telepsychotherapy during a pandemic: a traumatic stress perspective. *J. Psychother. Integr.* 30, 174–187. doi: 10.1037/int0000221

Suhaimi, N. S., Mountstephens, J., and Teo, J. (2020). Eeg-based emotion recognition: a state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* 2020, 426. doi: 10.1155/2020/8875426

Wiederhold, B. K., and Wiederhold, M. D. (2009). A review of virtual reality as a psychotherapeutic tool. *CyberPsychol. Behav.* 1, 45–52. doi: 10.1089/cpb.1998.1.45

Zhang, Z., Fort, J. M., and Giménez Mateu, L. (2023). Facial expression recognition in virtual reality environments: challenges and opportunities. *Front. Psychol.* 14, 1280136. doi: 10.3389/fpsyg.2023.1280136

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 11 January 2024
DOI 10.3389/fcomp.2023.1285690

Check for updates

OPEN ACCESS

EDITED BY
Alessia Celeghin,
University of Turin, Italy

REVIEWED BY
Marco Viola,
University Institute of Higher Studies in Pavia, Italy
Joan Cruz,
Escuela Colombiana de Ingenieria Julio Garavito, Colombia

*CORRESPONDENCE
Radoslaw Niewiadomski
✉ radoslaw.niewiadomski@unige.it

RECEIVED 30 August 2023
ACCEPTED 18 December 2023
PUBLISHED 11 January 2024

CITATION
Niewiadomski R, Larradet F, Barresi G and Mattos LS (2024) Self-assessment of affect-related events for physiological data collection in the wild based on appraisal theories. *Front. Comput. Sci.* 5:1285690. doi: 10.3389/fcomp.2023.1285690

Self-assessment of affect-related events for physiological data collection in the wild based on appraisal theories

Radoslaw Niewiadomski[1]*, Fanny Larradet[2], Giacinto Barresi[3] and Leonardo S. Mattos[2]

[1]Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genoa, Genoa, Italy, [2]Advanced Robotics, Istituto Italiano di Tecnologia, Genoa, Italy, [3]Rehab Technologies Lab, Istituto Italiano di Tecnologia, Genoa, Italy

This paper addresses the need for collecting and labeling affect-related data in ecological settings. Collecting the annotations in the wild is a very challenging task, which, however, is crucial for the creation of datasets and emotion recognition models. We propose a novel solution to collect and annotate such data: a questionnaire based on the appraisal theory, that is accessible through an open-source mobile application. Our approach exploits a commercially available wearable physiological sensor connected to a smartphone. The app detects potentially relevant events from the physiological data, and prompts the users to report their emotions using a novel questionnaire based on the Ortony, Clore, and Collins (OCC) Model. The questionnaire is designed to gather information about the appraisal process concerning the significant event. The app guides a user through the reporting process by posing a series of questions related to the event. As a result, the annotated data can be used, e.g., to develop emotion recognition models. In the paper, we analyze users' reports. To validate the questionnaire, we asked 22 individuals to use the app and the sensor for a week. The analysis of the collected annotations shed new light on self-assessment in terms of appraisals. We compared a proposed method with two commonly used methods for reporting affect-related events: (1) a two-dimensional model of valence and arousal, and (2) a forced-choice list of 22 labels. According to the results, appraisal-based reports largely corresponded to the self-reported values of arousal and valence, but they differed substantially from the labels provided with a forced-choice list. In the latter case, when using the forced-choice list, individuals primarily selected labels of basic emotions such as anger or joy. However, they reported a greater variety of emotional states when using appraisal theory for self-assessment of the same events. Thus, proposed approach aids participants to focus on potential causes of their states, facilitating more precise reporting. We also found that regardless of the reporting mode (mandatory vs. voluntary reporting), the ratio between positive and negative reports remained stable. The paper concludes with a list of guidelines to consider in future data collections using self-assessment.

KEYWORDS

emotion recognition, appraisal theories, data collection, self-assessment, physiological data, ecological setting, basic emotions

1 Introduction

Several techniques for emotion recognition from facial expression (Fasel and Luettin, 2003), speech (El Ayadi et al., 2011), full-body motion (Kleinsmith and Bianchi-Berthouze, 2013), tactile gestures (Niewiadomski et al., 2022), physiological signals (Jerritta et al., 2011) and other data sources have been studied intensively for at least two decades. Independently of the chosen method, all of them require creation of appropriate datasets. Therefore, several approaches exist for affect-related data collection and annotation (for recent surveys see Larradet et al. (2020) and Section 2).

Building datasets to train such models is often performed in laboratory setting by purposely inducing emotions to subjects at specific time intervals. This allows experimenters to control the stimuli and reduce the number of contextual factors that may influence the subjects' reactions. Different types of stimuli are used such as sounds, images and videos (Miranda-Correa et al., 2021) but also more complex techniques including virtual reality experience (Chirico et al., 2018; Marín-Morales et al., 2020; Dozio et al., 2022), playing various types of games (Niewiadomski et al., 2013; Bassano et al., 2019), and interaction with social robots (Redondo et al., 2023). To this date, more rare are studies that have attempted to build real-life (not induced) emotions datasets, i.e., collections of affect-related data, outside of the lab, in reaction to every-day events. In the literature, the terms "in the wild" (Dhall et al., 2013), "in the fray" (Healey et al., 2010), and "in real-life" (Devillers et al., 2005) are used to describe approaches where experimenters do not have direct control over the emotion elicitation process. Subjects are typically monitored during their everyday activities over extended periods of time to gather their most natural reactions.

Difficulties in building the affect-related datasets in ecological settings, e.g., establishing the ground truth, are well documented in the literature. Proper data segmentation and labeling is one of the main challenges. Despite all the challenges around their creation, *in the wild* datasets should be preferred in the light of results showing that humans' expressive behaviors and physiological reactions may differ between naturalistic settings and laboratory environments (Wilhelm and Grossman, 2010; Xu et al., 2017).

This paper proposes a novel method for collecting and labeling affect-related data in ecological settings. Our system is based on a commercially available sensor and a purposefully developed mobile application (app) that is used to collect physiological data and corresponding self-assessments. The app can work in two modalities: in the first modality (called mandatory) the user is prompted to report on potentially affect-related events detected by the app based on the analysis of the physiological signals acquired by the wearable sensor. In the second modality (called voluntary) the user can report affect-related events whenever they find it suitable. To collect the self-reports on affect-related events we designed the questionnaire based on the appraisal theory. We call this method **appraisal-based self-assessment questionnaire (ABSAQ)** in the remainder of the paper. The users report their *subjective* evaluations of the significant events, for example, they may state whether the event had positive or negative consequences, or if it was confirmed or not. We believe that using this method may have several advantages over the traditional methods such as self-reports based on (1) dimensional models, and (2) forced-choice lists of labels. The commonly used approach of self-assessment on emotional states is based on dimensional models [e.g., Russell's two-dimensional model of arousal and valence (Russell, 1980)]. Unfortunately, dimensional models do not provide exhaustive information about affective states. For instance, it is often stated that some very different (in terms of elicitation causes) emotional states such as anger and fear are placed very close to each other in a dimensional model (e.g., Russell's model). We expect that when comparing to reports based on dimensional models, our method (ABSAQ) should provide more detailed information about the emotional state of the user and its causes. Thus, in this paper, we investigate (research question 1) whether the reports obtained with the proposed method are consistent with reports based on dimensional models, while also providing more detailed information.

We also compare our method with a traditional approach based on a forced-choice list. By utilizing our method, self-reporting of affective events can be made easier and more efficient, particularly when compared to a method that relies on the forced-choice list with a lengthy list of labels (i.e., more than 6 Ekmanian emotions). This holds especially true when self-reports are collected in the wild. One appraisal theory (e.g., Ortony et al., 1990) can explain more than 20 different emotions. Thus, in the paper, we compare the self-reports obtained with ABSAQ (our proposed method) and a similarly extensive forced-choice list. In particular, in real-life situations, the number and diversity of emotional reactions are undoubtedly broader compared to data collected in-the-lab conditions through the controlled procedure, which often focus on a limited number of specific emotional experiences and utilize a small set of stimuli (such as videos, images, or sounds). Thus, in the wild, it is particularly important that the self-assessment procedure covers a possibly wide range of emotional experiences, which, in the case of a forced-choice list, means creating one long list of emotional labels. Consequently, we expect (research question 2) to observe substantial differences between ABSAQ reports and the self-assessments based on a forced-choice list. Additionally, in-the-lab data collections usually assume that specific stimulus (e.g., an image of a spider) will elicit a specific emotional reaction (e.g., fear). Using ABSAQ, the participants report their subjective evaluations, and thus the same event (i.e., stimulus) may correspond to different emotions. This method can be more suitable to report real-life experiences.

In the paper, we also investigate (research question 3) whether the way of self-reporting (mandatory, voluntary) may influence the quantity of positive and negative emotions reports. The role of positive emotions was widely studied (Fredrickson, 2001). Several intervention programs encourage or facilitate positive emotions awareness and experience (Moskowitz et al., 2021). We expect that participants are more willing to report positive emotions through voluntary reports and "feel obliged" to report negative ones through mandatory reports. In other words, we check whether the reporting method may influence the quantity of reported positive/negative events. Addressing this question is crucial for the development of innovative tools and methods for collecting affect-related data.

Finally, it is important to stress that the focus of this work is on methodological issues related to the (physiological) data collection and self-reporting. We do not provide any new model for emotion classification, nor do we claim that all emotional states considered in this paper can be differentiated in terms of physiological reactions.

The rest of the paper is organized as follows: after presenting the related works in Section 2, we present our mobile app in Section 3, and the dataset in Section 4. The analysis of the self-reports is presented in Section 5 which is followed by the general discussion in Section 6. We describe briefly the physiological signals dataset collected within this study in Section 7 and conclude the paper in Section 8.

2 Related works

2.1 Applications of appraisal theories

Appraisal theories have been largely used in affect-related studies. For instance, Conati and Zhou (2002) implemented a probabilistic model, using Dynamic Decision Networks, to recognize student's emotional states in an educational game context. For this, they followed the OCC appraisal theory (Ortony et al., 1990) and considered the students' goals and personality. In the video game context, Johnstone (1996) analyzed the relation between acoustic features of the player's vocal responses and the manipulations of some appraisals following Scherer's Component Process Model (Scherer, 2009). The same appraisal manipulations are addressed by van Reekum et al. (2004) in the context of a simple video game to study physiological reactions. In Bassano et al. (2019) a Virtual Reality (VR) game and a software platform collecting the player's multimodal data, synchronized with the VR content, are used to build a dataset. The game used was designed according to the emotion elicitation process described by Roseman's appraisal theory. In Meuleman and Rudrauf (2021) the authors used VR consumer games to elicit emotions in participants in-the-lab conditions. They asked participants to self-report appraisal components, physiological reactions, feelings, regulation, action tendencies, as well as emotion labels and dimensions. Using multivariate analyses, they discovered the relation between reported labels and affect components.

2.2 Methods for emotional self-reporting in the wild

According to Scherer (2005), existing techniques for emotional state self-reporting can be divided into two groups: free response and fixed-response labeling. While the first group allows for a higher precision of labeling [custom labels (Isomursu et al., 2007), verbal reports (Muaremi et al., 2013)], it makes it difficult to develop machine learning recognition models due to a potentially wide range of emotion labels selected by users. Constrained solutions include the usage of a finite list of labels (e.g., Nasoz et al., 2004) or dimensional models such as valence-arousal (e.g., Healey et al., 2010) or pleasure-arousal-dominance (e.g., Kocielnik et al., 2013). More user-friendly techniques may be used for reporting such as emoticons (Meschtscherjakov et al., 2009). Affect dimensions are often reported through the Self-Assessment Manikin (SAM) method (Isomursu et al., 2007) or through 2D point maps (Carroll et al., 2013).

In Schmidt et al. (2018), guidelines are provided for emotional labeling in the wild by comparing the results of different methods. A combination of manual reports and automatically triggered prompts is advised, as well as providing the means to the user to manually correct the timespan of an emotional event. Unlike Schmidt et al. (2018), which used time-based trigger, in this study prompting based on physiological cues (Myrtek and Brügner, 1996) was used and an experimenter-free data gathering protocol was implemented.

2.3 Methods for emotional physiological data collection

Emotion recognition from physiological data collected in-the-lab has been studied by different research groups (Shu et al., 2018). Most of the studies use measurements of Heart Rate (HR), Skin Conductance (SC), ElectroDermal Activity (EDA), Galvanic Skin Response (GSR), Skin Temperature (ST), and Respiration. The combinations of several signals, e.g., HR, EDA, and ST, have also been studied (e.g., Nasoz et al., 2004). Studies using data collected in ecological settings are rare, and most of them focus primarily on stress detection (Plarre et al., 2011; Hovsepian et al., 2015; Gjoreski et al., 2017) and moods (Zenonos et al., 2016). In real-life settings, the physiological data labeling and segmentation (i.e., defining the start and end of an emotion) are the main challenge (Healey et al., 2010). A few studies used mobile apps to collect both physiological data and affect-related states. Healey et al. (2010) conducted a real-life experiment using a mobile phone app to study different labeling methodologies for physiological data collection. They collected data and self-reports in the form of discrete labels and dimensional models (valence and arousal) and drew attention to some difficulties linked to self-reporting.

A large number of studies on automatic emotion recognition from physiological signals obtained good recognition rates (Jerritta et al., 2011) but very few of the proposed methods were then tested on data collected in the wild. Wilhelm and Grossman (2010) presented the risks of that approach by comparing physiological signals of in-the-lab induced stress and the ones occurring in ecological settings (e.g., watching a soccer game). They found the heart rate during the latter greatly superior to the former. Similarly, Xu et al. (2017) considered the validity of using in-the-lab collected data for ambulatory emotion detection. Their findings suggested that EDA, ECG, and EMG greatly differ between real-life and laboratory settings and that using such methodology results in low recognition rates (17%–45%). Thus, these results show that it is important to develop methods to build the datasets in the wild.

3 Appraisal theory-based app for data collection and labeling in the wild

We created a new system for physiological data collection and self-reports to satisfy the following requirements:

1. Can be used to capture the data of spontaneous emotions during daily activities;
2. Is minimally intrusive;
3. Guides the user through a process of reporting relevant events, by acquiring the necessary information to infer the related affective states, and without asking the user to pick any emotional label;
4. Guides the user to provide self-assessments by differentiating emotions from moods;
5. Detects the relevant events from the physiological data and prompts the user about it;
6. Provides a limited set of classes or categories of affective experiences that can be used to develop classification models.

The proposed solution consists of a self-assessment questionnaire based on appraisal theory, a commercially available wearable physiological sensor, and, a state-of-the-art event detection algorithm. A mobile application (app) developed *ad-hoc* allows the user to voluntary report affect-related events as well as report events detected through the physiological data analysis.

3.1 Self-reporting about relevant events

To address the requirements (3), (4), and (6) an questionnaire based on appraisal theory was designed. It serves to acquire the data about the whole appraisal process around the event (see Section 3.3.1 for details on the questionnaire). The questionnaire can be presented in a form of decision tree such that consecutive steps correspond to single appraisals. In this way, instead of a scoring potentially long list of emotion labels (in our experiment we use 22 different states), the participants answer to a set of questions that in most cases are binary (i.e., with answers "yes" or "no"). By collecting information based on such appraisal process, we expected to gather more consistent annotation of corresponding physiological signals.

To address the requirement (5), the app may work in two modalities. In the first modality, by utilizing the existing algorithm proposed by Myrtek and Brügner (1996), the app detects changes (such as additional heart rate) in physiological signals sent in real-time by the sensor. These changes may be related to certain emotional states. When these changes are detected, the app prompts users to provide an evaluation of the event, guiding them through the reporting process.

In the second modality, the user marks significant moments over the day (by pressing the button available on the bracelet of the wearable sensor) and later uses the same questionnaire to annotate the event. Both modalities are available all the time. When the person wears the sensor, the connection is maintained between the app and the sensor.

The reports result in a discrete number of classes (that can be represented by some emotional labels) corresponding to a combination of appraisals. They can, therefore, be used to build emotion classifiers using machine learning techniques. Additionally, the reports provide more information about the event (i.e., details on what led to the emotion). So, they can be potentially used not only for emotion classification but also to train the models that detect single appraisals from physiological data. Such models have rarely been investigated so far (Smith, 1989).

3.2 Sensors

The Empatica E4 bracelet[1] allowed us to fulfill requirements 1 and 2. This medical device was chosen for its sensors relevant to emotion detection: BVP, EDA, and ST as well as kinematic data through a 3D accelerometer. Its small size allows for long data collection without being bothersome. The device comes with an API for mobile applications and an already processed BVP to Inter Beat Interval (IBI). The sensor has also been used in the past for research purposes (Gjoreski et al., 2017).

The iPhone-based (iOS) mobile app uses a Bluetooth connection to collect physiological data from the E4 bracelet.

3.3 The application modules

The mobile app (see Figure 2) is composed of three modules.

3.3.1 The self-assessment module

This module is designed to collect information about relevant emotional events. Using this module, the users first provide the duration of a relevant event. The maximum duration was set to 5 min, because emotions are usually short experiences (compared to moods that can also be reported with the same app, see below for details). The user can manually reduce the event duration time (see Figure 2B). Next, they answer a series of questions (see Figures 2C, D) according to the ABSAQ questionnaire. In the last step, they evaluate the emotion intensity.

To create a questionnaire, the Ortony, Clore, and Collins (OCC) model (Ortony et al., 1990) was chosen as it was successfully used in affective computing applications in the past (Bartneck, 2002; Conati, 2002). Additionally, a set of appraisals in the OCC model, and the representation (i.e., decision tree) match our objectives and are easily understandable even by non-experts. The OCC can explain the elicitation of 22 different emotional states that can be triggered by some events, objects, or agents. It is important to notice that in the OCC model, the authors use the concept of emotion groups, which usually contain more than one label.[2] For example, the *Resentment group* (i.e., being displeased about an event presumed to be desirable to someone else) contains labels such as envy, jealousy, and resentment, while the *Reproach group* (i.e., disapproving someone else's blameworthy actions) contains labels such as appalled, contempt, despise, disdain, indignation, and reproach. In Figure 1 we provide one label for each group.

To collect the information about the relevant events, the participants report valence and arousal using five point scale based on SAM Mannekin questionnaire (Bradley and Lang, 1994).

1 https://www.empatica.com/en-eu/research/e4/ (accessed 4th September 2019).

2 In the remainder of the paper, we use a capital letter, when we refer to an emotion group, e.g., Reproach group.

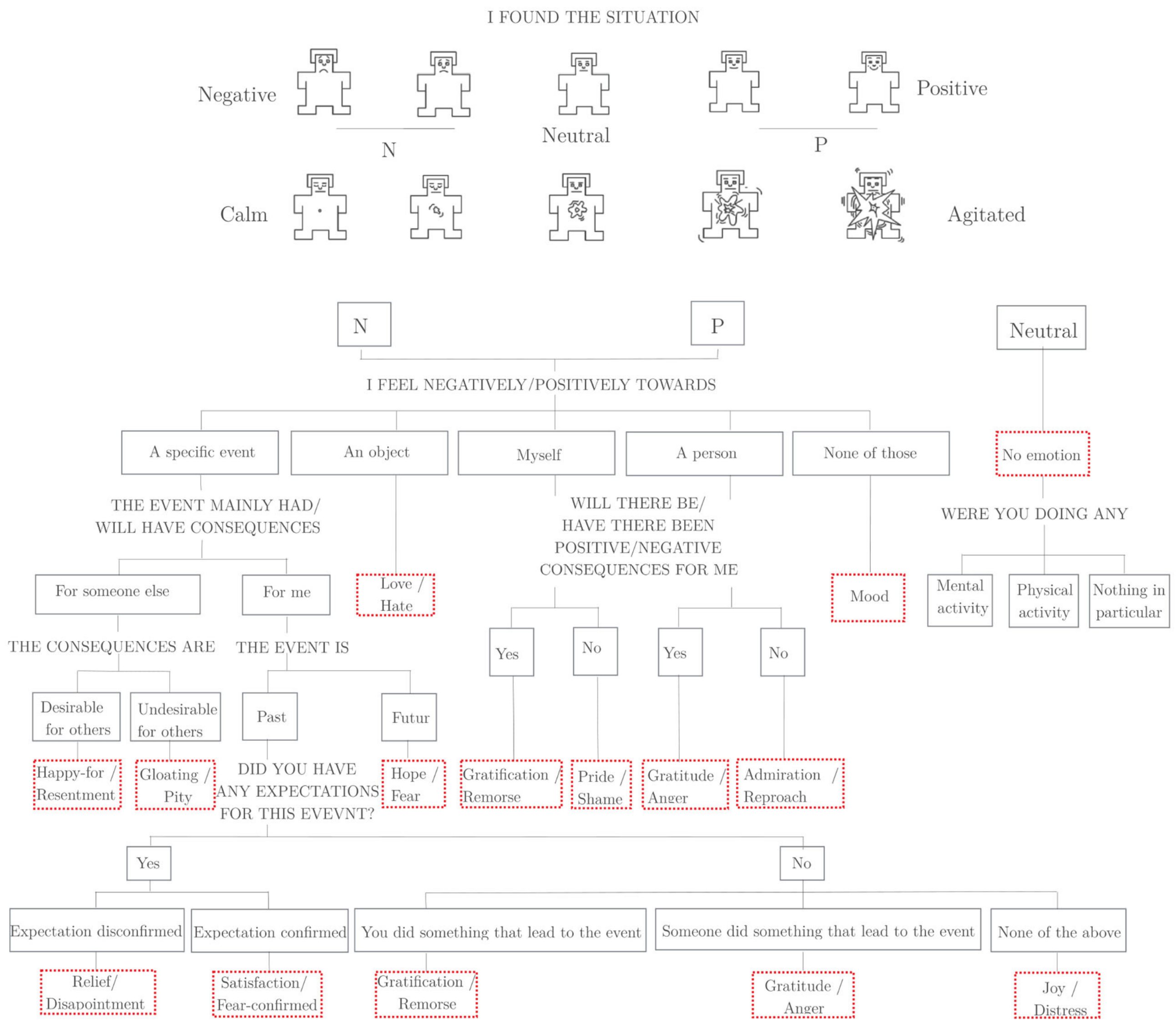


FIGURE 1
ABSAQ questionnaire. The final labels (emotions) are not displayed to the subjects. Red boxes contain two labels: one positive and one negative (the final label is selected according to the valence value reported at the beginning). The upper part of the Figure is based on Bradley and Lang (1994).

Next, the sequence of questions is displayed to the participant, following the structure presented in Figure 1. When designing the questionnaire small changes were introduced in relation to the original model. The main reason was to differentiate mood from emotions. Indeed, according to Clore and Ortony (2013), moods are *unconstrained in meaning*, while emotions are directed at specific objects, events or people. Therefore, a branch was added to report such "unconstrained in meaning" experiences (see Mood branch in Figure 1).

For instance, the person that is expecting a meeting with their boss, might choose the following sequence. First of all, they might consider their emotional state to be negative (valence) and of high arousal. Their emotional state would be caused by a specific event (here a meeting), which mainly had (or will have) *consequences for themself*. The event will take place in *future*. The corresponding emotion group in OCC model is called *Fear emotions* (and fear is an example of emotion belonging to this group). Thus, physiological signals gathered by the app at that moment would be labeled with the above sequence of appraisals which corresponds to Fear emotion group. Obviously, using our method enables individuals to report their personal evaluations, i.e., the same event (here an incoming meeting) could evoke hope in another individual.

The Valence rating was additionally used to identify "no emotion" reports. If the participant picks Valence = 3 (neutral), they are forwarded to "no emotion" part. This is consistent with OCC theory (Ortony et al., 1990), according to which, emotions are valenced experiences.

3.3.2 The event detection module

This module is used to detect relevant events from the data in real-time. The additional heart rate method (Myrtek and Brügner,

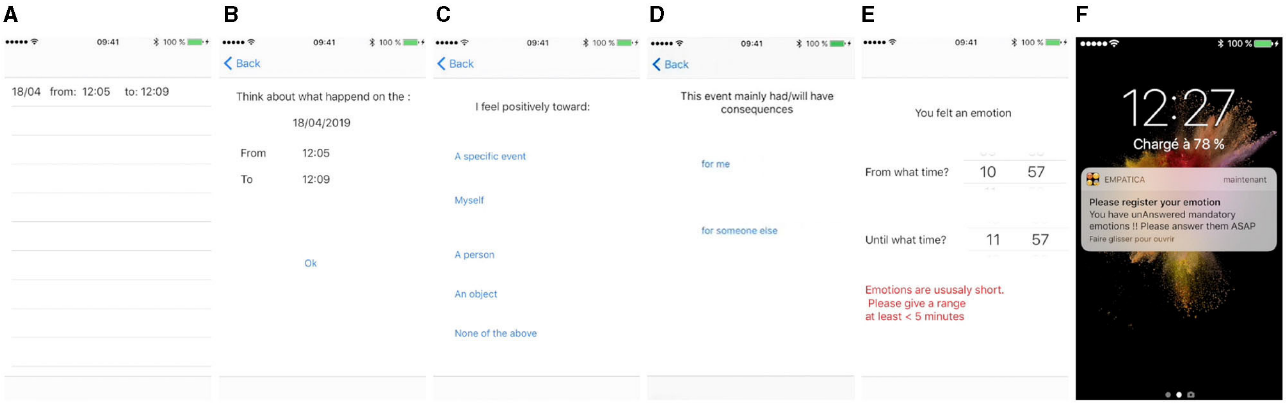


FIGURE 2
Some screenshots of the app: **(A)** mandatory tab with an unreported event on the list, **(B)** first screen of the self-assessment procedure used for mandatory reports, **(C, D)** reporting questions of ABSAQ report, **(E)** providing the voluntary report with a warning about the event duration, **(F)** prompt sent to the user about unreported events.

1996) was used to detect relevant events and prompt the user to report their emotion at this time. It consists in detecting heart rate increases that are unrelated to activity (estimated using the accelerometer). Detected events create a *mandatory events* list, which is always accessible to the user on a separate tab of the app. By implementing this algorithm requirement 5 was fulfilled. The minimum time interval between two mandatory events is fixed at 1 h to avoid the generation of too many related prompts. If two or more events are detected within an hour, only the first event is added to the mandatory list, and the remaining ones are ignored.

3.3.3 The notification module

It reminds the user to wear the device and to report the events from the mandatory event list if any unreported events are left on the list. Reminders are generated every 15 min. after the event. When the connection with the wristband is lost, notifications are generated by phone every 15 s until reconnection.

3.4 The application functionalities

The five tabs are available to the user. Two of them are used to report the events. The first tab called "mandatory" contains a list of automatically detected events for which the reports are not provided yet. When such an event is detected by the event detection module (see Section 3.3.2), a new entry is added to this list (see Figure 2A). Such events are processed according to the procedure described in Section 3.3.1. If there are events on the list, the user receives notification messages (see Figure 2F).

The users can also voluntarily report an undetected event in two different ways. First, they can use the second tab called "voluntary," and select manually the start and end time (see Figure 2E) of an event then continue reporting using the emotion definition module). Secondly, the user can also add an event to the list by pressing the bracelet's button. Reporting about the events immediately (i.e., during the emotional experience) might be difficult, especially when it is related to a strong emotional state. By pressing the button, the users manually add a new entry to the list of events. Such an event will have a precise timestamp corresponding to the moment of pressing the button. The participants can report about the event later (as they do for events detected automatically by the app).

The remaining three tabs of the app allow for a better experience with the app. They are used to temporarily stop the notifications, check the battery level and visualize the reports using graphics.

4 Data collection

The main aim of this study is to validate the new self-reporting procedure in the wild. To assess the utility of our ABSAQ questionnaire, we asked people to use our app for a period of 7 days. The participants were said that they would participate in the physiological data collection in the wild, and thus, they should provide self-reports for all the events detected by the app during this period, but also their additional voluntary reports about other events when they find it appropriate. Importantly, participants were unaware of the real aim of this study (i.e., evaluation of the reporting mode). Instead, they were said that their data would be used to develop novel machine learning models for emotion detection and to improve existing algorithms of the app.

The data collection was performed in two stages. In the first stage, participants used the app described in Section 3. In the second stage, another group of participants used a slightly modified version of the app. An additional question was added at the end of reporting procedure to collect also labels (see research question 2). In more detail, after reporting about the event with ABSAQ, the participants were asked to choose one label from the list containing all the labels that are present in the OCC model (22 labels). This was done to check the differences between the two ways of self-reporting.

As mentioned above, we do not claim that it is possible to recognize all 22 emotions from physiological signals (and in particular from the data collected in this study). Given that our participants were unaware of the challenges associated with emotion recognition from physiological signals, it is reasonable to assume that they believed their data would be utilized to develop emotion recognition models capable of detecting a wide range of emotional labels, potentially up to 22 labels.

4.1 Data collection protocol and annotation format

Twenty-two subjects (seven females) participated in this study: 14 in the stage 1, and 8 in the stage 2. Most of the participants were students or researchers in the early stages of their research careers (avg. age 31) of various nationalities, living and working in the same city (in Italy). The data were collected during a week of their ordinary work/internship/study activities. The experimental procedures follow the IIT ADVR TEEP02 protocol, approved by the Ethical Committee of Liguria Region on September 19, 2017. After signing the informed consent, the subjects wore the Empatica E4 wristband for 7 days. During this time they were asked to report their emotions using the mobile application previously described in Section 3.

The collected reports are of three types: Mood, Emotion, or No emotion. All reports contained a start time, an end time, an optional comment, and a sequence of answers in ABSAQ (Figure 1). The Arousal and Valence are integers between 1 and 5.

In the remainder of the paper, the term "inferred labels" will refer to the labels inferred from the path according to ABSAQ, and "reported labels" will refer to the labels directly provided by the participants (in stage 2). Similarly, "inferred arousal" correspond to the arousal inferred from the answers in ABSAQ, and "reported arousal" is the value explicitly reported by the participant (at the beginning of the questionnaire).

4.2 Data confidentiality

During the data collection, physiological data and self-assessments are gathered and stored on the smartphone without being transmitted to any cloud or similar online service. The smartphone is secured with a system password to prevent unauthorized access by other individuals (e.g., in case the participant loses the device).

Upon completion of data collection, only researchers involved in the study can access the stored data on the smartphone by connecting it physically to a computer and utilizing appropriate software. They download the data to an offline storage device and follow the best practices to maintain data pseudo-anonymity.

5 The analysis of reports

5.1 General results

Twenty-two participants used the app and sensor and reported their emotions. Some participants used the app for less than the suggested seven days. We include their data in the following analyses. Overall, the reports were collected over 133 days. In total,

TABLE 1 The ABSAQ reports.

Emotions			
Appreciation emotions (admiration)	27	Anger emotions (anger)	13
Gloating emotions (gloating)	2	Disappointment emotions (disappointment)	3
Gratitude emotions (gratitude)	45	Distress emotions (distress)	10
Gratification emotions (gratification)	59	Fear emotions (fear)	19
Hope emotions (hope)	37	Fear-confirmed emotions (fear-confirmed)	7
Happy-for emotions (happy-for)	8	Disliking emotions (hate)	13
Joy emotions (joyful)	2	Sorry-for emotions (pity)	8
Liking emotions (love)	25	Remorse emotions (remorse)	10
Pride emotions (pride)	21	Resentment emotions (resentment)	2
Relief emotions (relief)	0	Reproach emotions (reproach)	44
Satisfaction emotions (satisfaction)	27	Self-Reproach emotions (shame)	20
Total positive emotions	253	Total negative emotions	149
Moods			
Positive mood	37	Negative mood	14
Other activities			
Mental effort	65		
Physical activity	32		
Nothing in particular	159		
Total			709

The exemplary emotion labels are provided in parentheses.

709 reports were collected. This corresponds to 32.2 reports on average per person. The highest number of events reported by a single participant was 66, and the lowest was 12 (standard deviation is 13.87). Four hundred two reports described emotional states, 51 moods, 65 mental activities, 32 sport activities, and the remaining 159 did not correspond to any of the above considered categories. Most of the reports (415) were mandatory (i.e., responses to the events detected by the app). One hundred fifty reports were generated after using the button, and the remaining were reported manually in the app. The total number of reports (including emotions and moods) with a positive valence is 290, while the total number of reports with a negative valence is 163.

Table 1 provides detailed information about the labels that are inferred from the questionnaire. We provide the emotion group as well as, one example of a label from each group. In the remaining of the paper, we refer to the whole group by using this label.

Strong differences can be observed in the frequency of label occurrences: gratification group was the most often chosen group (59 times). The group of negative emotions most frequently chosen was Reproach (44 times). The other two frequent groups were:

TABLE 2 The appearance of specific appraisal combinations when merging positive and negative reports (see also Figure 1).

Path	Occurrence
Other agent–no	71
Self agent-yes	59
Event–for me–future	56
None	51
Other agent–yes	45
Self agent–no	41
Object	38
Event–for me–past–yes–confirmed	34
Event–for me–past–no–none of the above	12
Event–for me–past–no–someone did something that lead to the event	12
Event–for me–past–no–you did something that lead to the event	11
Event–for someone else–desirable for others	10
Event–for someone else–undesirable for others	10
Event–for me–past–yes–disconfirmed	3

Gratitude picked 45 times and Hope picked 37 times. On the opposite side, the groups such as Gloating, Joy, and Resentment were chosen very rarely, and Relief was never picked.

There is also a relatively high number of reported events (307 in total) that are not related to emotions. These occurrences may be a result of the introduction of mandatory reports (see Section 5.4 for more details).

After observing strong differences in the choice of emotional categories, we check whether some appraisals are more frequent than others. Therefore, in Table 2, we present the frequency of selecting each appraisal path, without considering the valence of the emotional state associated with the reported event. It can be seen that participants often reported emotions toward other persons. The most frequently chosen path in the questionnaire indicates emotions toward other persons that would not bring (or had brought) any consequences to the reporting person (such as admiration or reproach), while emotions toward others that had or would have some consequences for the reporting person (such as gratitude or anger) were at the fourth place. The second most often chosen path corresponds to the emotions toward self that would bring (or had brought) some consequences to the reporting person (such as gratification or remorse). Next, emotions related to some future events that may not be confirmed (such as hope or fear) were picked 56 times. The other frequent path corresponds to moods (i.e., states not related to any specific person, event not object).

5.2 Dimensional vs. ABSAQ reports

To determine whether the reported arousal and valence align with the answers given in the questionnaire, we calculate the average reported values of arousal and valence for each category. We then compare these values with the arousal and valence values for these labels reported in the literature. For this purpose, we used two works: Whissell (1989) and Hepach et al. (2011) as they consider a huge number of emotional states. In Whissell (1989), the arousal and valence for 107 labels are reported using 7 point scales (from 1 to 7). In Hepach et al. (2011), the arousal and valence for 62 labels are reported using nine point scales.[3] Unfortunately, these two publications do not cover all the emotion categories present in the OCC model. The Fear-confirmed, Gloating, Gratification, and Admiration groups are neither present in Whissell (1989) nor in Hepach et al. (2011).

To compare the values, for each emotional category we check all matching labels from an OCC emotion group in Whissell (1989) and Hepach et al. (2011). If more than one label is present then we compute the average value of arousal and valence by taking into consideration all matching labels. At the same time, we also rescale the resulting average values from Whissell (1989) and Hepach et al. (2011) to the range of values used in our experiment (i.e., 1–5). For example, the Disappointment group in OCC model (i.e., being displeased about the disconfirmation of the prospect of a desirable event) contains five labels: dashed-hopes, despair, disappointment, frustration and heartbroken. Two of these labels are present in Whissell (1989): despairing (arousal 4.1, valence 2.0) and disappointed (arousal 5.2, valence 2.4). First, we compute the average values (here: 4.65 and 2.2), and, next, we scale them to the interval [1, 5], obtaining 3.43 for arousal and 1.8 for valence.

In the last step, we compute the 2D distance between the reported average values of arousal and valence, and the average rescaled values of arousal and valence according to Whissell (1989) and Hepach et al. (2011). All the results are in Table 3.

It is important to recall that when the participants choose the valence value (first step of the questionnaire), their choice determines the remaining path in the report. It is not permitted to report the emotional state with valence equal to 3 (i.e., neutral), as such a state is not considered an emotion in the OCC model. In our questionnaire, a valence value of 3 corresponds to physical or mental activity, rather than an emotion.

As it can be seen in Table 3, distances between the average reported arousal and valence, and the values computed from the inferred labels by using values provided by Whissell (1989) and Hepach et al. (2011) are small for most of emotion groups. The average distance is 0.8 when using (Whissell, 1989) and 0.73 when using Hepach et al. (2011). These two values are similar to the average distance between the values reported by Whissell (1989) and Hepach et al. (2011) for the same set of labels, which is 0.74. It let us believe that the reported arousal and valence are in general consistent with the answers given to the second part of the questionnaire (i.e., appraisal-based part). Consequently, below we discuss only a small number of cases for which these results vary the most.

The strongest differences were observed for: (i) the Hope group, for which reports indicate much higher valence but lower arousal than it is reported in the literature, (ii) the Sorry-for Group (e.g., pity), for which reports show higher arousal but lower valence;

3 The meaning of the valence scale is reversed in Hepach et al. (2011), i.e., it starts "very positive" and it ends at "very negative."

TABLE 3 The reported average arousal and valence, and comparison to the average values in the literature.

Emotion group	Reported values		Whissell (1989)			Hepach et al. (2011)		
	Avg. A	Avg. V	Avg. A	Avg. V	Dist.	Avg. A	Avg. V	Dist.
Appreciation (e.g., admiration)	2.93	4.48						
Anger	3.85	1.46	3.49	2.31	0.92	4.53	1.75	0.74
Disappointment	3.33	1.33	3.43	1.80	0.48	3.89	1.88	0.78
Distress	3.90	1.80	3.08	2.25	0.93	3.69	2.07	0.34
Fear	3.37	1.79	3.97	2.23	0.74	4.02	2.21	0.77
Fear-confirmed	3.71	1.71						
Gloating	5.00	5.00						
Gratitude	2.67	4.64				2.00	4.24	0.78
Gratification	2.08	4.53						
Disliking (e.g., hate)	3.08	1.69	3.17	2.37	0.68	4.26	1.58	1.19
Hope	2.59	4.68	3.20	3.63	1.21	3.19	3.72	1.13
Happy-for	3.25	4.50	3.13	4.60	0.15			
Joy	2.50	4.00	3.69	4.10	1.19	2.45	4.49	0.49
Liking (e.g., love)	3.00	4.52	3.47	3.93	0.75	2.79	4.43	0.23
Sorry-for (e.g., pity)	3.38	1.63	2.73	2.47	1.06	2.98	2.77	1.21
Pride	2.76	4.43	3.47	3.87	0.90	2.56	4.10	0.39
Remorse	3.40	1.40	2.40	1.80	1.08	3.20	2.74	1.35
Resentment	4.00	1.50	4.00	2.20	0.70	3.92	2.08	0.59
Reproach	3.52	1.70	2.87	1.93	0.69	3.73	1.74	0.21
Satisfaction	2.37	4.28	3.07	3.60	0.98			
Self-reproach (e.g., shame)	3.20	1.75	2.91	1.78	0.29	3.39	2.49	0.76

The values in columns 4th, 5th, 7th, and 8th are rescaled. Additionally, the values in column 8th are inverted. "A" states for arousal, "V" for valence, and "Dist" for distance.

and (iii) the Remorse Group. At the same time, we also notice differences between the two sources, namely (Whissell, 1989) and Hepach et al. (2011). This is evident in the case of the Joy and Hate groups, for which the reported values in our study are similar to only one of the two sources.

We also look at the difference between arousal reported by our participants and the one that can be inferred from the labels using Whissell (1989). Here relatively strong differences were observed again for Joy and Remorse groups. When comparing the values of reported arousal with Hepach et al. (2011), noticeable differences occur only for the Disliking group (e.g., hate).

It is important to notice that some of these groups have very few instances (e.g., the Joy group, see Table 1) which may explain not optimal results for these groups.

5.3 Forced-choice vs. ABSAQ reports

The Table 4 provides detailed information about the labels explicitly reported by participants and those that are inferred from the ABSAQ.

Out of the 89 reports considered in stage 2, there were only six instances where the explicitly given label and the label inferred from ABSAQ precisely corresponded. These instances occurred twice for distress and gratitude, and once for joy and satisfaction. Occasionally (three cases) the reports do not match even in terms of valence (i.e., the individuals may have selected a negative emotion while the ABSAQ indicated a positive one, or vice versa).

A detailed analysis of Table 4 shows that participants tended to pick basic or "Ekmanian" labels when providing the reports. Joy was the most often chosen label (14 times). Anger and distress (which can be considered related to sadness) were at the second place (12 times each), and fear was at third place (ex aequo, with happy-for label). However, when considering the ABSAQ, the most frequently chosen appraisal sequences corresponded to negative emotions toward other individuals (e.g., reproach label) and toward oneself (e.g., shame label), as well as positive emotions toward oneself with positive consequences (gratification). It is interesting to see that the labels: reproach, shame, and gratification were not picked even once from the forced-choice list. In general, the labels inferred from ABSAQ are better distributed (standard deviation 3.12) compared to forced-choice list reports (standard deviation 4.18).

When analyzing disagreements between reported emotions and inferred labels, certain patterns were observed to be more frequent than others. For positive emotions,

TABLE 4 Number of reported (rep.) labels directly selected by participants and inferred (inf.) from the questionnaire in stage 2.

Label	Rep.	Inf.	Label	Rep.	Inf.
Admiration	3	3	Anger	12	6
Gloating	0	0	Disappointment	7	0
Gratitude	3	8	Distress	12	4
Gratification	0	10	Fear	7	3
Hope	2	4	Fear-confirmed	0	2
Happy-for	7	1	Hate	1	2
Joy	14	2	Pity	3	3
Love	5	3	Remorse	0	3
Pride	2	4	Resentment	1	0
Relief	4	0	Reproach	0	11
Satisfaction	6	5	Shame	0	9

TABLE 5 Number of positive, negative, mandatory and voluntary reports.

	Inferred Positive	Inferred Negative	Total
Mandatory	141	82	223
Voluntary	112	67	179
Total	253	149	402

the reported label of happy-for often coincided with the Gratitude group, while the reported joy aligned with the Pride and Gratification groups. For negative emotions, the reported label of anger frequently coincided with the Reproach group.

5.4 Mandatory vs. voluntary reports

Additionally, the relation between valence and the mandatory or voluntary character of the report was calculated. Table 5 presents the number of positive, negative, mandatory and voluntary reports.

Contrary to our expectations, the participants reported negative emotions with equal frequency in both reporting modalities: when prompted to do so and when reporting voluntarily. The percentage of mandatory positive reports is 63.2% of all mandatory reports, while the percentage of voluntary positive reports is 62.5% of all voluntary reports.

The relatively high contribution of the voluntary reports was observed for the emotion groups, for which, at least according to the current state-of-the-art, automatic detection from the physiological data is particularly challenging or even impossible: such as Gratitude, Gratification, Reproach, and Admiration. Probably the simple algorithm used to detect significant events (see Section 3.3.2) is not suitable, and, consequently, these emotions were relatively more often reported voluntarily.

5.5 Auxiliary analyses

The mobile application was programmed in such a way that it was possible to identify when subjects changed their mind when reporting their emotional state. For instance, one may select “A specific event," then, once the next question is displayed, go back and “A person" instead. We notice that 32 times participants changed their opinion when reporting.

Additionally, participants were able to add comments when they desired. Such disclosure of personal information was made optional in order to respect the subjects' privacy. In total, only 37 such reports were collected. These comments associated with the inferred label give additional information about the reporting process. Most of the optional comments given by the participants seem to fit the self-assignments, a “software crash" (ID 1) is likely to induce “distress," and “a meeting" to induce fear (ID 3). For instance, the comment “itchy annoying mosquito bite" (ID 4) is interesting. In this case, the participant's emotion appears to be directed toward the reason for the pain, that is the mosquito, as the subject selected the sequence “I feel negatively toward A person," resulting in the inferred label of anger. This highlights the benefits of utilizing appraisal theory for self-assessment.

6 Discussion

In stages 1 and 2, we observed that the reported arousal and valence values do not differ substantially from the values found in the literature corresponding to inferred labels. However, only a very small number of explicitly provided labels in stage 2 correspond to inferred labels. Thus, both results confirm our expectations (research question 1 and 2). Our structured approach allows the user to describe their emotional states by offering a set of relatively simple questions with a limited number of options. In the stage 2, the most often reported emotion is joy, which is considered a generic description of a positive experience. Our belief is that when seeing a list of 22 labels the participants choose the well-known label without considering the subtle differences between different positive states such as gratitude and gratification. A completely different situation arises when considering the ABSAQ questionnaire and inferred labels. In this case, the two most often reported states are gratitude and gratification. A similar preference toward the popular (i.e., basic, Ekmanian) labels is observed when analyzing the negative emotions. In this case, anger is the most frequently reported label (on a pair with distress). The same tendency is observed in the entire dataset: positive emotions toward self or other people with positive consequences (Gratitude and Gratification) are the most frequently inferred emotion groups, along with negative emotions toward others with no consequences (Reproach group). This result shows that there might be a bias toward selecting more well-known emotional labels (such as basic emotions) when they are explicitly listed. This should be taken into consideration in future studies when using a self-assessment procedure. We believe that our approach helps the participants to focus on the possible causes, and in consequence, to report more precisely about their affective state.

Surprisingly, we discovered that there is no difference in terms of positive and negative reports when comes to the modality of reporting (research question 3). Our expectations were not confirmed as in both modalities (voluntary and mandatory) a similar ratio of negative and positive reports were observed. This contradicts our hypothesis that participants would be more eager to voluntarily report positive emotions than negative ones. We did not observe differences in valence ratio between two reporting modalities, and both of them have some important advantages. The mandatory reports were rated as emotions 56% of the time and 78% of the emotional reports were mandatory. On one hand, it means that nearly one quarter of the emotional events would not be annotated, if the voluntary modality was not available. On the other hand, 88% of the mood labels were picked in mandatory modality, which shows that current approach to detect the emotional states is not optimal. Thus, we recommend the researchers to use both modalities (i.e., mandatory and voluntary) in the future data collection studies.

Our study brings several additional interesting observations. First, it can be noticed that a low number of emotional reports per day was collected (three emotions per day were reported in average, and 5.32 reports per day in total, that is, including other activities). The number of reports per day is clearly lower when compared to previous works, e.g., Trampe et al. (2015). In Trampe et al. (2015), the app prompts randomly the user a fixed number of times per day (the number of daily questionnaire requests is preset). Such an approach can result in situations when the app prompts the user to report something, even if they do not experience anything that would be worth reporting. In Trampe et al. (2015) even 90% of reports indicate an emotional experience, and the authors comment on this result stating that "People's everyday life seems profoundly emotional." In our approach, several mandatory reports (i.e., prompted by the app) result in non-emotional experience (e.g., mental activity). In total 36% of the reports concern non-emotional experiences. Thus, we recommend that other researchers also consider introducing the possibility for participants to report on activities unrelated to emotions, even in data collections in-the-wild focusing on emotional states.

Second, some emotions groups are chosen more frequently than others (Figure 3) with Gratification counting a total of 59 reports while groups such as Gloating and Resentment counting only two reports. At first sight, it might be surprising that some well-known labels such as joy are rarely present in the dataset. On the other hand, it has to be acknowledged that in this study we distinguished a high number of positive emotions, compared to other studies. In total, 11 different positive emotional states are considered (while in several other studies, all positive states are covered with one generic label of joy). This result confirms the necessity of a more fine-grained analysis of positive emotional states also in future studies. Adding the possibility to report a variety of positive emotions provides interesting information. The most commonly chosen group, i.e., the Gratification group, encompasses positive states that are directed toward oneself and are associated with having or expecting positive consequences for the reporting person. This result could potentially be attributed to the fact that the majority of our participants were students or researchers in the early stages of their research careers. It is possible that individuals belonging to such a group often experience positive emotions related to personal achievements. We also notice differences in the label frequency between our study and (Trampe et al., 2015). In Trampe et al. (2015) the most frequently reported emotions were joy, followed by love, anxiety, and satisfaction.

In general, positive emotions are reported more often than negative ones. At the same time, it should be noticed that the ratio between reported positive and negative emotions in our study is 1.7, which is far from the postulated relation 3:1 (Fredrickson and Losada, 2005), and also lower than the ratio observed in other studies, e.g., 2.59 in Trampe et al. (2015). The disparity can be observed between participants (see Figure 4) with four individuals who reported more negative than positive states, and one participant (ID12) whose nearly 80% of reports was positive. This result, however, may also be influenced by the specific demographics and occupational situation of the participants.

We also observed that some individuals reported a relatively low number of emotions, such as subject 5 who reported only 11 emotional labels over the course of one week, while others reported a higher number, such as subject 7 who reported 48 emotions (see Figure 5). The observed disparity in the reports could be influenced by differences in the number of emotional stimuli encountered by the individuals during their participation in the experiment. It is, however, known that individuals may vary in their level of emotional awareness, with some being more attuned to their emotions than others (Myrtek et al., 2005). Based on these findings, we recommend to vary the duration of data collection time with respect to this factor, and permit some participants to use the app longer than others in future works.

7 Open dataset

The dataset collected during the experiment described in the previous Section is freely available and can be used by researchers to unravel the challenges of emotion detection in the wild. The data annotation consists of both appraisals and corresponding emotional labels. The dataset includes emotional states that are rarely considered in other publicly available physiological datasets. All the data was collected with the E4 wristband.

The physiological signals had the following frame rates: GSR—4 data points per second, BVP—64 data points per second, ST—4 data points per second, ACC—32 data points per second, IBI—Calculated from BVP, one data point for each BVP peak. No signal post-processing or filtering was applied to the data.

Both the app code (iPhone) and the physiological data gathered during the experiment are freely available at https://gitlab.com/flarradet/epsdi.

8 Conclusions and future works

In this paper, a new tool was proposed to collect physiological signals and self-assessments in ecological settings. Our solution inspired by appraisal theories, allows users to self-report the appraisal process around relevant events. The reports can be of two types: voluntary and requested by the app (when a substantial change in the physiological data is detected). We also performed the data collection with 22 participants who used the app for 133 days in total.

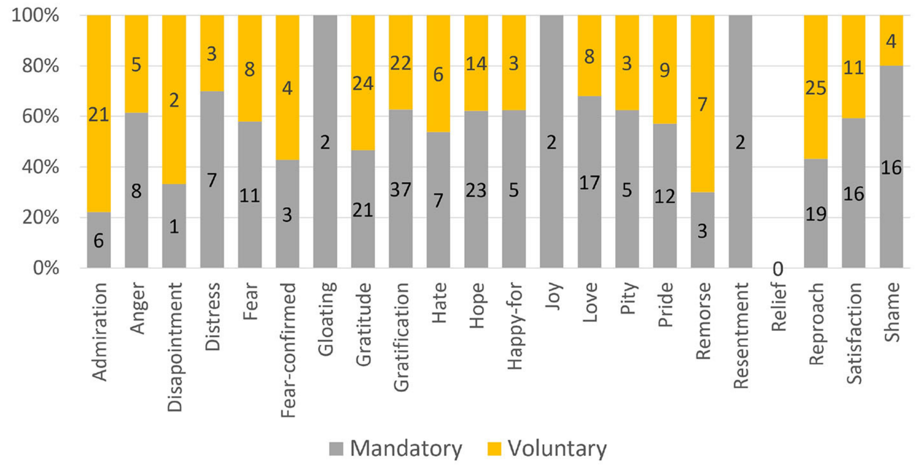


FIGURE 3
The number of voluntary and mandatory reports per label.

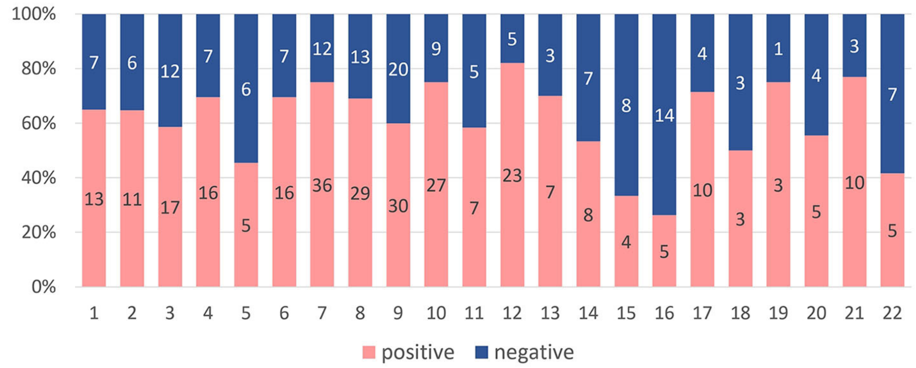


FIGURE 4
The number positive and negative reports per participant.

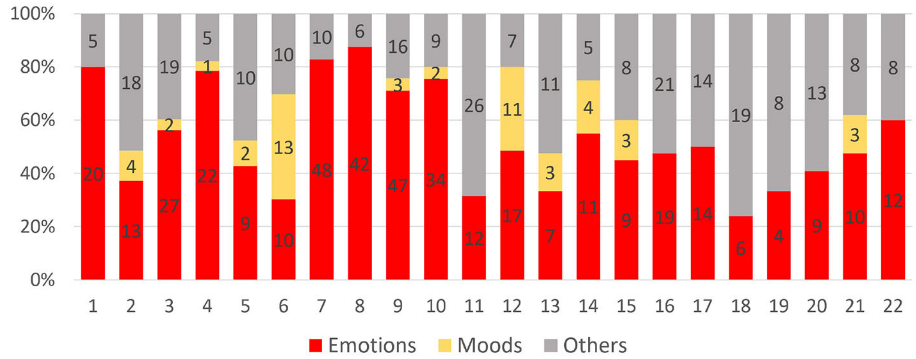


FIGURE 5
The number emotions, moods and others reports per participant.

The proposed data collection methodology proven to be effective in gathering self-assessment on affect-based events in the wild. Self-assessments obtained through our technique are consistent with the reported valence and arousal, but they differ substantially from self-assessments provided with a forced-choice list. We observed that regardless of the reporting mode (mandatory vs. voluntary reporting), the ratio between positive and negative reports is stable. Last but not least, when using a forced-list choice the participants have a tendency to pick labels of basic emotions (e.g., anger, joy), this is not the case for the ABSAQ questionnaire which additionally provides insights into the causes of emotional states.

To our knowledge, this is the first app for data collection and self-assessment based on appraisal theory. It can be used by other researchers, e.g., to extend the dataset. The dataset might be used in the future, e.g., to train specific classifiers, by choosing the relevant subset of the appraisals and emotional labels.

The additional advantage of using appraisal theory for reporting is that the information about the appraisal process can be used to develop models for single appraisals (and not emotional labels). For instance, one can use physiological data to train detection models, focusing on whether an event is confirmed (or not), or whether it is (un)expected (see Mortillaro et al., 2012). It is in line with Scherer's Component Process Model (Scherer, 2009) according to which the behavioral changes correspond to specific appraisals rather than emotional labels. To sum up, our approach allows the researchers to develop two different types of models (for emotions and appraisals).

Some limitations of this work should be mentioned. First of all, we do not claim that it is possible to detect 22 different emotions by using existing sensors (and physiological signals in general). Building the recognition model is out of the scope of this work. Secondly, taking into account the number of participants, and time of recordings, the number of reported events is surprisingly low. Some technical issues revealed during the data collection (e.g., device disconnection) may hand a limited impact on it. Similarly, the event detection algorithm used by the app is rather simple and the event detection should be improved by training appropriate machine learning models, for instance, on the dataset collected in this experiment. Third, this study uses one appraisal theory only. It is envisaged to explore other theories (e.g., Roseman et al., 1996) in future works.

Future work will focus on additional evaluations of the ABSAQ questionnaire. While our main aim is to introduce a new self-assessment method in the wild, we also believe that we need to perform more controlled experiments. The different methods of self-assessment can be compared when events (stimuli) that potentially may elicit emotional states are experimentally controlled (e.g., in the virtual environment, see a preliminary work by Bassano et al., 2019). This would allow us to compare the reports provided by different participants about the specific events. The study presented in the paper was conducted on a homogeneous group of young adults living in Western culture during a week of their ordinary work activity. We recommend replicating the study on different populations. While the main results (i.e., three main research questions) are not likely to be strongly affected by demographic factors, these factors as well as, the main activity or lifestyle might influence some secondary results such as the ratio of reported positive and negative emotions. Some individual's characteristics, such as the ability to correctly interpret one's own emotional states or neurodivergence, might also have a certain impact on results and thus should be studied more thoroughly in the future.

Several applications of this work are envisaged. First of all, the work is part of project TEEP-SLA, which aims at automatically detecting emotions from physiological signals for Amyotrophic Lateral Sclerosis (ALS) patients. The long-term aim is to create a large dataset to be used in emotion recognition from physiological data collected in natural settings. Moreover, people are already willing to report their emotions on mobile apps for the sole purpose of self-monitoring. As wearable sensors become more popular, our approach may enable large data collections, and boost the research on affect recognition. While in this study, the app and ABSAQ are used to collect the physiological data and self-assessments, we believe that the same methodology can be used to collect self-reports for other data sources, e.g., audio data. Consequently, a large number of affective computing applications may benefit from the more accurate and efficient tools for data collection and labeling in the wild.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://gitlab.com/flarradet/epsdi.

Ethics statement

IIT ADVR TEEP02 protocol approved by the Ethical Committee of Liguria Region. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RN: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft, Writing— review & editing. FL: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing—original draft. GB: Writing—original draft, Writing—review & editing, Methodology. LM: Resources, Supervision, Validation, Writing—original draft, Writing—review & editing, Funding acquisition, Project administration.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was partially supported by Fondazione Roma as part of the project TEEP-SLA.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bartneck, C. (2002). "Integrating the OCC model of emotions in embodied characters," in *Workshop on Virtual Conversational Characters* (Princeton, NJ: Citeseer), 39–48.

Bassano, C., Ballestin, G., Ceccaldi, E., Larradet, F. I., Mancini, M., Volta, E., et al. (2019). "A VR game-based system for multimodal emotion data collection," in *Motion, Interaction and Games* (New York, NY: ACM), 38. doi: 10.1145/3359566.3364695

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Carroll, E. A., Czerwinski, M., Roseway, A., Kapoor, A., Johns, P., Rowan, K., et al. (2013). "Food and mood: just-in-time support for emotional eating," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (IEEE), 252–257. doi: 10.1109/ACII.2013.48

Chirico, A., Ferrise, F., Cordella, L., and Gaggioli, A. (2018). Designing awe in virtual reality: an experimental study. *Front. Psychol.* 8, 2351. doi: 10.3389/fpsyg.2017.02351

Clore, G. L., and Ortony, A. (2013). Psychological construction in the OCC model of emotion. *Emot. Rev.* 5, 335–343. doi: 10.1177/1754073913489751

Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. *Appl. Artif. Intell.* 16, 555–575. doi: 10.1080/08839510290030390

Conati, C., and Zhou, X. (2002). "Modeling students' emotions from cognitive appraisal in educational games," in *Intelligent Tutoring Systems*, eds S. A. Cerri, G. Gouardères, and F. Paraguaçu (Berlin: Springer Berlin Heidelberg), 944–954. doi: 10.1007/3-540-47987-2_94

Devillers, L., Vidrascu, L., and Lamel, L. (2005). Challenges in real-life emotion annotation and machine learning based detection. *Neural Netw.* 18, 407–422. doi: 10.1016/j.neunet.2005.03.007

Dhall, A., Goecke, R., Joshi, J., Wagner, M., and Gedeon, T. (2013). "Emotion recognition in the wild challenge 2013," in *Proceedings of the 15th ACM on International Conference on Multimodal Interaction* (New York, NY: ACM), 509–516. doi: 10.1145/2522848.2531739

Dozio, N., Marcolin, F., Scurati, G. W., Ulrich, L., Nonis, F., Vezzetti, E., et al. (2022). A design methodology for affective virtual reality. *Int. J. Hum. Comput. Stud.* 162, 102791. doi: 10.1016/j.ijhcs.2022.102791

El Ayadi, M., Kamel, M. S., and Karray, F. (2011). Survey on speech emotion recognition: features, classification schemes, and databases. *Pattern Recognit.* 44, 572–587. doi: 10.1016/j.patcog.2010.09.020

Fasel, B., and Luettin, J. (2003). Automatic facial expression analysis: a survey. *Pattern Recognit.* 36, 259–275. doi: 10.1016/S0031-3203(02)00052-3

Fredrickson, B. (2001). The role of positive emotions in positive psychology. the broaden-and-build theory of positive emotions. *Am. Psychol.* 56, 218–226. doi: 10.1037/0003-066X.56.3.218

Fredrickson, B., and Losada, M. (2005). The positive affect and the complex dynamics of human flourishing. *Am. Psychol.* 60, 678–86. doi: 10.1037/0003-066X.60.7.678

Gjoreski, M., Luštrek, M., Gams, M., and Gjoreski, H. (2017). Monitoring stress with a wrist device using context. *J. Biomed. Inform.* 73, 159–170. doi: 10.1016/j.jbi.2017.08.006

Healey, J., Nachman, L., Subramanian, S., Shahabdeen, J., and Morris, M. (2010). "Out of the lab and into the fray: towards modeling emotion in everyday life," in *International Conference on Pervasive Computing* (Cham: Springer), 156–173. doi: 10.1007/978-3-642-12654-3_10

Hepach, R., Kliemann, D., Gruneisen, S., Heekeren, H., and Dziobek, I. (2011). Conceptualizing emotions along the dimensions of valence, arousal, and communicative frequency – implications for social-cognitive tests and training tools. *Front. Psychol.* 2, 266. doi: 10.3389/fpsyg.2011.00266

Hovsepian, K., al'Absi, M., Ertin, E., Kamarck, T., Nakajima, M., and Kumar, S. (2015). "cstress: towards a gold standard for continuous stress assessment in the mobile environment," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (New York, NY: ACM), 493–504. doi: 10.1145/2750858.2807526

Isomursu, M., Tähti, M., Väinämö, S., and Kuutti, K. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *Int. J. Hum. Comput. Stud.* 65, 404–418. doi: 10.1016/j.ijhcs.2006.11.007

Jerritta, S., Murugappan, M., Nagarajan, R., and Wan, K. (2011). "Physiological signals based human emotion recognition: a review," in *2011 IEEE 7th International Colloquium on Signal Processing and its Applications* (Penang: IEEE), 410–415. doi: 10.1109/CSPA.2011.5759912

Johnstone, T. (1996). "Emotional speech elicited using computer games," in *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96*, Vol. 3 (IEEE), 1985–1988. doi: 10.1109/ICSLP.1996.608026

Kleinsmith, A., and Bianchi-Berthouze, N. (2013). Affective body expression perception and recognition: a survey. *IEEE Trans. Affect. Comput.* 4, 15–33. doi: 10.1109/T-AFFC.2012.16

Kocielnik, R., Sidorova, N., Maggi, F. M., Ouwerkerk, M., and Westerink, J. H. (2013). "Smart technologies for long-term stress monitoring at work," in *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems* (Porto: IEEE), 53–58. doi: 10.1109/CBMS.2013.6627764

Larradet, F., Niewiadomski, R., Barresi, G., Caldwell, D. G., and Mattos, L. S. (2020). Toward emotion recognition from physiological signals in the wild: approaching the methodological issues in real-life data collection. *Front. Psychol.* 11, 1111. doi: 10.3389/fpsyg.2020.01111

Marín-Morales, J., Llinares, C., Guixeres, J., and Alcañiz, M. (2020). Emotion recognition in immersive virtual reality: from statistics to affective computing. *Sensors* 20, 5163. doi: 10.3390/s20185163

Meschtscherjakov, A., Weiss, A., and Scherndl, T. (2009). Utilizing emoticons on mobile devices within esm studies to measure emotions in the field. *Proc. MME Conjunct. MobileHCI* 9, 3361–3366.

Meuleman, B., and Rudrauf, D. (2021). Induction and profiling of strong multi-componential emotions in virtual reality. *IEEE Trans. Affect. Comput.* 12, 189–202. doi: 10.1109/TAFFC.2018.2864730

Miranda-Correa, J. A., Abadi, M. K., Sebe, N., and Patras, I. (2021). Amigos: a dataset for affect, personality and mood research on individuals and groups. *IEEE Trans. Affect. Comput.* 12, 479–493. doi: 10.1109/TAFFC.2018.2884461

Mortillaro, M., Meuleman, B., and Scherer, K. R. (2012). Advocating a componential appraisal model to guide emotion recognition. *Int. J. Synth. Emot.* 3, 18–32. doi: 10.4018/jse.2012010102

Moskowitz, J. T., Cheung, E. O., Freedman, M., Fernando, C., Zhang, M. W., Huffman, J. C., et al. (2021). Measuring positive emotion outcomes in positive psychology interventions: a literature review. *Emot. Rev.* 13, 60–73. doi: 10.1177/1754073920950811

Muaremi, A., Arnrich, B., and Tröster, G. (2013). Towards measuring stress with smartphones and wearable devices during workday and sleep. *Bionanoscience* 3, 172–183. doi: 10.1007/s12668-013-0089-2

Myrtek, M., Aschenbrenner, E., Brügner, G., and Psychophysiological Research Group at the University of Freiburg (FRG) (2005). Emotions in everyday life: an ambulatory monitoring study with female students. *Biol. Psychol.* 68, 237–255. doi: 10.1016/j.biopsycho.2004.06.001

Myrtek, M., and Brügner, G. (1996). Perception of emotions in everyday life: studies with patients and normals. *Biol. Psychol.* 42, 147–164. doi: 10.1016/0301-0511(95)05152-X

Nasoz, F., Alvarez, K., Lisetti, C. L., and Finkelstein, N. (2004). Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cogn. Technol. Work* 6, 4–14. doi: 10.1007/s10111-003-0143-x

Niewiadomski, R., Beyan, C., and Sciutti, A. (2022). Affect recognition in hand-object interaction using object-sensed tactile and kinematic data. *IEEE Trans. Haptics* 16, 112–117. doi: 10.1109/TOH.2022.3230643

Niewiadomski, R., Mancini, M., Baur, T., Varni, G., Griffin, H., Aung, M., et al. (2013). "MMLI: multimodal multiperson corpus of laughter in interaction," in *Human Behavior Understanding, Volume 8212 of Lecture Notes in Computer Science*, A. Salah, H. Hung, O. Aran, and H. Gunes(Cham: Springer International Publishing), 184–195. doi: 10.1007/978-3-319-02714-2_16

Ortony, A., Clore, G. L., and Collins, A. (1990). *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.

Plarre, K., Raij, A., Hossain, S. M., Ali, A. A., Nakajima, M., Al'Absi, M., et al. (2011). "Continuous inference of psychological stress from sensory measurements collected in the natural environment," in *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks* (IEEE), 97–108.

Redondo, M. E. L., Niewiadomski, R., Rea, F., Incao, S., Sandini, G., Sciutti, A., et al. (2023). Comfortability analysis under a human–robot interaction perspective. *Int. J. Soc. Robot*. doi: 10.1007/s12369-023-01026-9

Roseman, I. J., Antoniou, A. A., and Jose, P. E. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cogn. Emot*. 10, 241–278. doi: 10.1080/026999396380240

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol*. 39, 1161. doi: 10.1037/h0077714

Scherer, K. R. (2005). What are emotions? and how can they be measured? *Soc. Sci. Inf*. 44, 695–729. doi: 10.1177/0539018405058216

Scherer, K. R. (2009). The dynamic architecture of emotion: evidence for the component process model. *Cogn. Emot*. 23, 1307–1351. doi: 10.1080/02699930902928969

Schmidt, P., Reiss, A., Dürichen, R., and Van Laerhoven, K. (2018). "Labelling affective states in the wild: Practical guidelines and lessons learned," in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (New York, NY: ACM), 654–659. doi: 10.1145/3267305.3267551

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., et al. (2018). A review of emotion recognition using physiological signals. *Sensors* 18, 1–41. doi: 10.3390/s18072074

Smith, C. (1989). Dimensions of appraisal and physiological response in emotion. *J. Pers. Soc. Psychol*. 339–353. doi: 10.1037/0022-3514.56.3.339

Trampe, D., Quoidbach, J., and Taquet, M. (2015). Emotions in everyday life. *PLoS ONE* 10, 1–15. doi: 10.1371/journal.pone.0145450

van Reekum, C., Johnstone, T., Banse, R., Etter, A., Wehrle, T., Scherer, K., et al. (2004). Psychophysiological responses to appraisal dimensions in a computer game. *Cogn. Emot*. 18, 663–688. doi: 10.1080/02699930341000167

Whissell, C. M. (1989). "Chapter 5 - The dictionary of affect in language," in *The Measurement of Emotions*, eds R. Plutchik, and H. Kellerman (Cambridge:Academic Press), 113–131. doi: 10.1016/B978-0-12-558704-4.50011-6

Wilhelm, F. H., and Grossman, P. (2010). Emotions beyond the laboratory: theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. *Biol. Psychol*. 84, 552–569. doi: 10.1016/j.biopsycho.2010.01.017

Xu, Y., Hübener, I., Seipp, A.-K., Ohly, S., and David, K. (2017). "From the lab to the real-world: an investigation on the influence of human movement on emotion recognition using physiological signals," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (IEEE), 345–350.

Zenonos, A., Khan, A., Kalogridis, G., Vatsikas, S., Lewis, T., Sooriyabandara, M., et al. (2016). "Healthyoffice: mood recognition at work using smartphones and wearable sensors," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)* (Sydney, NSW: IEEE), 1–6. doi: 10.1109/PERCOMW.2016.7457166

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 01 February 2024
DOI 10.3389/fpsyg.2023.1294386

OPEN ACCESS

EDITED BY
Enrico Vezzetti,
Polytechnic University of Turin, Italy

REVIEWED BY
Andreas Rausch,
University of Mannheim, Germany
Tiina Susanna Törmänen,
University of Oulu, Finland

*CORRESPONDENCE
Anne Horvers
✉ anne.horvers@ru.nl

RECEIVED 14 September 2023
ACCEPTED 19 December 2023
PUBLISHED 01 February 2024

CITATION
Horvers A, Molenaar I, Van Der West H, Bosse T and Lazonder AW (2024) Multimodal measurements enhance insights into emotional responses to immediate feedback. *Front. Psychol.* 14:1294386.
doi: 10.3389/fpsyg.2023.1294386

Multimodal measurements enhance insights into emotional responses to immediate feedback

Anne Horvers*, Inge Molenaar, Heleen Van Der West, Tibor Bosse and Ard W. Lazonder

Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

Adaptive learning technologies often provide students with immediate feedback on task performance. This feedback can elicit various emotional responses, which, in turn, influence learning. Most recent studies capture these emotions by single data streams, contradicting the multi-componential nature of emotion. Therefore, this study investigated 32 university students solving mathematical problems using an adaptive learning technology. Students received immediate feedback on every step in the solution process, after which their physiological, experiential and behavioral responses to this feedback were recorded. Physiological arousal was measured by electrodermal activity, valence was measured by self-reports (experiential), and emotion types were measured by observations of facial expressions (behavioral). Results showed more peaks in electrodermal activity after feedback than was expected based on chance. These responses were comparable in strength after feedback on failure and success. Students' experiential responses conveyed mostly positive valence after feedback on success and mostly negative valence after feedback on failure. Behavioral observations showed more negative than positive emotion types after feedback on failure and more positive than negative emotion types after feedback on success. These results show that physiological arousal is a valuable objective indicator of emotional responses after immediate feedback but should be accompanied by other data streams in order to understand students' emotional responses. Both valence and emotion types can be used for this purpose. These outcomes pave the way for designing adaptive learning technologies that take students' emotions into account.

KEYWORDS

emotional responses, immediate feedback, adaptive learning technologies, physiological arousal, multimodal measurements

1 Introduction

In recent years, computer-assisted learning, intelligent tutoring systems and adaptive learning technologies (ALTs) have increasingly become prevalent in education across the globe (Martin et al., 2020; Baker, 2021). These systems have been developed to promote individual students' learning by providing automated feedback on their task performance (van Lehn, 2011; Aleven et al., 2016; Pardo et al., 2019). This feedback can be delivered in real-time during the learning process, which is known to increase learning effectiveness (Tärning, 2018; Deeva et al., 2021). Most contemporary learning technologies, such as ALTs, go beyond providing immediate feedback by adapting the difficulty of future practice problems to individual students' current task performance. Students' answers are utilized to infer their ability level, which the ALT's underlying algorithm uses to determine the difficulty of the next practice problem (Elo, 1978; Klinkenberg et al., 2011). However, this algorithm is solely based on cognitive achievements and disregards students' emotions (Aleven et al., 2016), which play a crucial role in the learning process and directly influence students' learning (D'Mello, 2017; Pekrun, 2022). For example, negative emotions hamper learning (Götz and Hall, 2013; Loderer et al., 2020) and affect students' effort, perception and use of learning strategies (Eteläpelto et al., 2018; Obergriesser and Stoeger, 2020). Moments of success and failure in learning can elicit a wide range of emotional responses, and the same goes for feedback on learning task performance (Peterson et al., 2015; Lipnevich et al., 2021). In other words, there is good reason to examine the possibility of ALTs taking emotions into account. A first step in that direction is to better understand which emotions are triggered by immediate feedback.

This study aimed to gain insight into students' emotional responses to immediate feedback given by an ALT. Emotions are seen as multi-componential in nature in this study, consisting of the dimensions arousal and valence (Russell, 1980; Pekrun, 2006). This study extends previous research by analyzing multimodal data streams to capture these dimensions: measures of physiological arousal via electrodermal activity (EDA), self-reported valence (experiential) and observations of emotion types via facial expressions (behavioral). Additional innovative features include the instant measurement of emotional responses (rather than at the end of a learning session) to immediate feedback on every step in the solution process (as opposed to feedback on the final solution or delayed feedback). The next sections elaborate on the roles emotions can play during learning, measures of emotions, and the different types of emotional responses to feedback.

1.1 Emotions in learning

Learning and human emotions are reciprocally related: emotions affect learning directly, and success or failure during the learning process influences students' emotions (Pekrun, 2006). Emotions also affect students' effort, perception and use of learning strategies (Eteläpelto et al., 2018; Pekrun, 2022). Emotions can either enhance or impede learning. For example, when students feel frustrated, confused or bored, their learning is negatively impacted, while feelings of enjoyment or pride have a positive influence on students' learning (Götz and Hall, 2013; Loderer et al., 2020). However, overcoming a state of confusion can also benefit learning, which illustrates that the interplay between emotion and learning can vary among students (Baker et al., 2010; Graesser, 2020). That is, some students may prefer easy tasks to avoid negative emotions due to failure, whereas others like to be challenged by difficult tasks and experience fewer emotions when they do not succeed (Baker et al., 2010).

These emotions students experience can be defined from a categorical and dimensional perspective. Categorical theories divide emotions into different types, such as fear, anger, happiness, surprise, disgust and sadness (Ekman, 1999). Each emotion type is associated with a distinct facial expression and action tendencies (Coppin and Sander, 2021). However, it has been argued that these basic emotion types bear little relationship with learning (Kort et al., 2001). Dimensional theories of emotion, by contrast, do have this connection and address the multi-componential nature of emotions by portraying emotions by the continuous components arousal and valence (Russell, 1980; Pekrun, 2006). Arousal indicates the amount of physiological activation of the body that occurs when an emotion is triggered, while valence refers to the pleasantness of an emotion, which can range from positive to negative (Russell, 1980; Pekrun, 2017).

The Control-Value Theory (CVT) integrates these perspectives, specifically focusing on emotions during learning (i.e., achievement emotions) and is widely used in educational research (Pekrun, 2017). This theory argues that achievement emotions can differ in object focus, with a distinction between activity emotions that occur during learning (e.g., boredom during a learning task) and outcome emotions related to success and failure in the past or future (e.g., anxiety related to future failure or pride related to past success) (Jarrell et al., 2017; Pekrun, 2017). Anxiety is, for instance, seen as an emotion with a negative valence, high activation and an outcome focus (Pekrun et al., 2007). The effects of especially positive deactivating and negative activating emotions on learning are complex (Pekrun and Linnenbrink-Garcia, 2014). Experiencing positive deactivating emotions (e.g., relaxation) can reduce students' effort and negatively influence learning, contrary to positive activating emotions (e.g., enjoyment and pride) (Wu and Yu, 2022). Negative activating emotions (e.g., frustration and anxiety) are shown to impede learning, but can also enhance students' effort to perform better (Graesser and D'Mello, 2012; Pekrun, 2017; Cloude et al., 2021; Taub et al., 2021). In this study, emotions are conceptualized as multi-componential using the CVT.

The multi-componential nature of emotions, as described by dimensional theories, points to differences in the expression and experience of these emotions between humans (Harley et al., 2015; Azevedo et al., 2022). This emotional experience is also influenced by a combination of a student's appraisal of a learning situation and the associated emotional response, as students take into account their perceptions of control and evaluations of task value (Pekrun, 2017). Moreover, different psychological subsystems are at play when a student feels anxious, including affective (feeling nervous), cognitive (being worried), motivational (avoidance), expressive (nervous face), and physiological (high bodily activation) processes (Kleinginna and Kleinginna, 1981; Pekrun, 2006; Ortony et al., 2022). Considering emotions as a multi-componential construct and measuring it as such is recommended by recent studies as well (Li et al., 2021). Previous research typically recorded students' physiological, experiential, and behavioral responses to personally meaningful stimuli (Mauss and Robinson, 2009; Horvers et al., 2021). These physiological responses involve the reaction of the body when an emotion is evoked (Dawson et al., 2016). Experiential responses refer to the subjective personal experience of an emotion, and behavioral responses concern the observable behavioral reactions (Mauss and Robinson, 2009). These responses provide the opportunity to measure emotions in a multi-componential way, this approach will be used in this study.

1.2 Emotional responses to feedback during learning

Feedback can cause various physiological, experiential and behavioral responses (Jarrell et al., 2017). Variations in EDA (physiological arousal) can occur when students receive feedback on their performance. For example, Aghaei Pour et al. (2010) found cross-student differences in an unspecified set of physiological features, while Malmberg et al. (2019a) showed that synchrony in the EDA's above-threshold peaks occurred when students discussed collective feedback. Feedback can elicit valence (experiential) ranging from positive to negative and different facial expressions (behavioral) (Peterson et al., 2015; Lipnevich et al., 2021). Evidence regarding the relationship between emotional responses and feedback on success (FOS) and feedback on failure (FOF) is typically mixed, possibly resulting from differences in individual's appraisals of a learning situation (Pekrun, 2017). Some studies concluded that FOF leads to negative emotions, such as frustration, and FOS leads to positive emotions (D'Mello et al., 2010; Lipnevich et al., 2021). Other studies showed that FOF elicits particularly intense and negative emotions (Rowe et al., 2014; Hill et al., 2021), which can linger longer than positive emotions and resurface with greater intensity on future tasks (Hill et al., 2021). These emotional responses can also impact students' actions — that is, positive emotions can motivate students to try harder and improve by facilitating the evaluation of their learning (Pitt and Norton, 2017). When students get feedback that their answer is incorrect, they can become frustrated or anxious and, hence, discouraged to perform the next task (Vogl and Pekrun, 2016). However, these same negative emotions can also motivate some students to perform better on the upcoming task (Vogl and Pekrun, 2016; Lim et al., 2020). Repeated instances of FOF undermine students' sense of control and result in negative emotions (Pekrun, 2006). Feedback that an answer is correct (i.e., FOS) can lead to feeling in control of learning, which can again lead to pride (Lipnevich et al., 2021).

This study adds to the existing body of research by (1) investigating immediate feedback on every step in the solution process, (2) measuring emotional responses during the learning process: continuously and after every instance of feedback, and (3) using a multimodal approach to capture these emotional responses. The above-mentioned insights from previous research and existing theories are mainly derived from studies investigating emotional responses to delayed feedback, either given by teachers when students struggle and ask for help or by technologies after the completion of a full learning session (Hill et al., 2021; Lipnevich et al., 2021). Even though ALTs and other technologies are particularly suitable for providing students with immediate feedback, most recent studies focused on detecting emotions during the learning process without attending to the role of feedback (Loderer et al., 2020; Lal et al., 2021). The studies that do focus on feedback investigated it after a task has been completed (Aghaei Pour et al., 2010; D'Mello et al., 2010), while contemporary technologies also provide opportunities to provide feedback on every step in the solution process (Molenaar, 2022). It remains unclear whether and to what extent these previous insights generalize to situations where students receive immediate feedback. The present study will adopt a fine-grained approach to immediate feedback to answer this question.

Secondly, extant research generally assessed emotional responses at the end of the learning session (Jarrell et al., 2017; Hill et al., 2021) or even after a week's delay (Peterson et al., 2015; Lipnevich et al., 2021). These retrospective measures miss out on the rapidly changing nature of emotions which causes the dynamics of feedback and emotion to happen in less than a second (Pekrun, 2006). As ALTs provide immediate feedback, emotional responses to this feedback should be instantly assessed by fine-grained measures (Pekrun, 2006; D'Mello, 2013). Physiological arousal is a promising measure to instantly capture emotional responses because the intensity of and fluctuations in arousal can be measured in real-time (Jarrell et al., 2017; Malmberg et al., 2019a).

Finally, most previous research on the connection between feedback and emotions used unimodal approaches, for example, by capturing only experiential responses through semi-structured interviews (Lim et al., 2020; Hill et al., 2021) or self-report questionnaires (Jarrell et al., 2017; Lipnevich et al., 2021). However, considering emotions as a multi-componential construct is recommended by recent publications (Li et al., 2021). Using a single data stream has constraints, such as the focus on one aspect of emotion and the subjective nature of self-report data (Pekrun, 2020). In addition, physiological, experiential, and behavioral responses are weakly related, suggesting that each portrays a unique aspect of a person's emotional responses (Mauss and Robinson, 2009), that cannot be measured by the other responses (Egger et al., 2019). For example, physiological arousal can be successfully measured with EDA, whereas valence cannot (Mauss et al., 2005; Mauss and Robinson, 2009; Horvers et al., 2021). As it remains unclear how students emotionally respond to feedback during learning, a multimodal approach seems more appropriate (Peixoto et al., 2015; Egger et al., 2019). By using physiological arousal as measured by EDA (physiological responses), self-reported valence (experiential responses) and emotion types as measured by observations of facial expressions (behavioral responses) after every feedback event, this study aimed to gain a detailed understanding of emotional responses to immediate feedback in ALTs, which paves the way for designing ALTs that take students' emotions into account.

1.3 Research questions and hypotheses

Although ALTs provide students with immediate feedback in real time, hardly any research has been done on emotional responses to this feedback. However, gaining an understanding of these relationships is important for designing ALTs that can take emotions into account. This study, therefore, aimed to investigate emotional responses to feedback given by an ALT immediately after students entered a calculation into the system. Physiological, experiential and behavioral responses were measured in a multimodal approach. Physiological responses were captured by physiological arousal as measured continuously by EDA. Experiential responses were assessed through self-reported valence and behavioral responses by capturing emotion types as measured by observations of facial expressions, both after every instance of feedback. These data streams were analyzed to answer the following research questions: (1) To what extent does immediate feedback trigger peaks in students' EDA (physiological arousal)? (2) Which experiential (valence) and behavioral responses (emotion type) of students are triggered by immediate feedback?

As previous research has shown that feedback can elicit emotions in students and physiological arousal can vary after feedback (D'Mello, 2017; Malmberg et al., 2019a), we expected that that feedback will generate above-threshold peaks (EDA) at an above-chance level (hypothesis 1). As feedback on failure (FOF) elicits particularly strong emotions (Rowe et al., 2014; Hill et al., 2021), students' physiological responses were expected to be stronger. That is, there will be more above-threshold peaks in EDA, a higher amplitude sum and higher mean phasic activity within response window for FOF than for feedback on success (FOS) (hypothesis 2). As most previous research showed that FOF would predominantly elicit negative emotions while FOS would mainly elicit positive emotions (D'Mello et al., 2010; Lipnevich et al., 2021), the third hypothesis predicted that students' experiential responses would indicate predominantly negative valence after FOF and positive valence after FOS. A similar pattern was expected to occur for behavioral responses, meaning that students will exhibit predominantly negative emotion types after FOF and positive emotion types after FOS (hypothesis 4) (D'Mello et al., 2010; Lipnevich et al., 2021).

2 Materials and methods

2.1 Participants

Data was gathered from 36 Dutch university students, but 4 of them had to be excluded from the sample due to technical problems during data collection. The remaining 32 participants were 24 women (75%) and 8 men (25%) aged 18–28 ($M = 21.28$, $SD = 2.67$). They studied at the faculty of arts (25%), faculty of medical sciences (15%), faculty of management (15%), faculty of law (5%), faculty of science (15%), and faculty of social sciences (25%). Thirteen participants were first-year bachelor students (40.6%), 4 were second-year bachelor students (12.5%), 5 were third-year bachelor (15.6%), 5 were fourth-year bachelor (15.6%), and 5 were master students (15.6%). Students signed active consent for participation in this study. The research has been independently reviewed by the Ethics Committee Social Sciences (ECSS) of the Radboud University, and there is no formal objection [ECSW-2020-14].

2.2 Design and procedure

This descriptive study administered a pre-test-only design (Figure 1) and took place in a research laboratory at the students' university. During the main phase of the study, students worked with an ALT to solve mathematical problems using the quadratic formula. Prior to that, they took a pre-test that assessed their prerequisite math knowledge and received video instructions on the quadratic formula. Next, they watched a nature video to establish the EDA baseline. Students then solved three mathematics problems and received immediate feedback on every calculation they entered into the ALT. After every feedback event, students were prompted to report their experiential responses by indicating the valence of their emotional state on a 5-point scale. Their physiological arousal was measured through sweat gland activity using an EDA wristband during the entire session (see section 2.4.4). Behavioral responses were captured by observations of facial expressions indicating emotion types and were done after the session using video recordings of the participants' faces.

2.3 Materials

2.3.1 Adaptive learning technology

The ALT used in this study was AlgebraKiT, a web-based software application for practicing mathematical problem-solving (Figure 2). Students entered every calculation they made to arrive at the final solution of the problem into the ALT. The algorithm behind AlgebraKiT analyzed these steps using the calculation principles taught in secondary education. Based on these principles, the ALT determined if a student's calculation was correct or incorrect and provided students with immediate feedback after every calculation. The task difficulty of tasks was manually adjusted based on students' pre-test scores.

2.3.2 Feedback types

AlgebraKiT could generate two types of immediate feedback, which were labeled feedback on success (FOS), which was given after a student made a correct calculation, and feedback on failure (FOF), which was given after an incorrect calculation. A green circle indicated a correct calculation and a red circle with a white cross indicated an incorrect calculation (Figure 2). In some cases (105 times), the notification that a calculation was incorrect was followed by an explanation.

2.3.3 Learning objective

The learning objective in this study was the quadratic formula $\left(x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}\right)$. Students were asked to solve three practice problems with this formula, which required multiple calculation steps. The first problem had students solve a given quadratic equation. In the second problem, students needed to calculate the intersection of two equations. The third problem was of an applied nature in that the formula was embedded in a cover story. Every problem had three difficulty levels to facilitate variation in success and failure in all students:

Easy: e.g. Given are two equations $f(x)$ and $g(x)$. Solve the equation

$$f(x) = g(x).$$

$$f(x) = 2 - x^2 - 2x \text{ and } g(x) = 5x - 4$$

Intermediate: e.g. Given are two equations $f(x)$ and $g(x)$. Solve the equation

$$f(x) = g(x).$$

$$f(x) = 8 - x(3 - 2x) \text{ and } g(x) = 5x^2 + 2 + 4x.$$

Hard: e.g. Given are two equations $f(x)$ and $g(x)$. Solve the equation

$$f(x) = g(x).$$

$$f(x) = \frac{3}{2}x^2 + \frac{1}{2}x - 7 \text{ and } g(x) = -x\left(\frac{1}{4}x + 2\right) - 5$$

Students were assigned to one of these difficulty levels based on their pre-test score (see section 2.4.3. for more information on the pre-test).

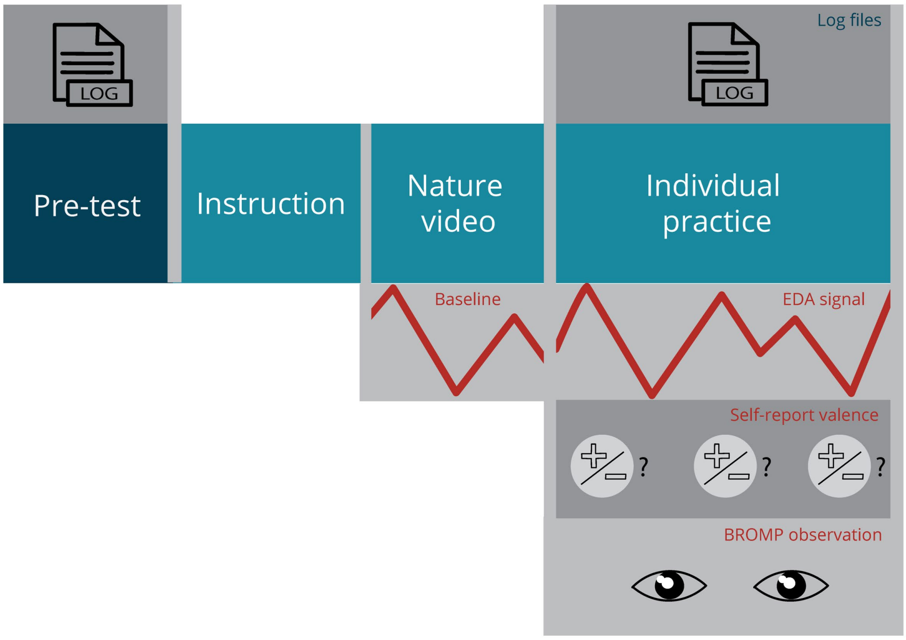


FIGURE 1
Study design.

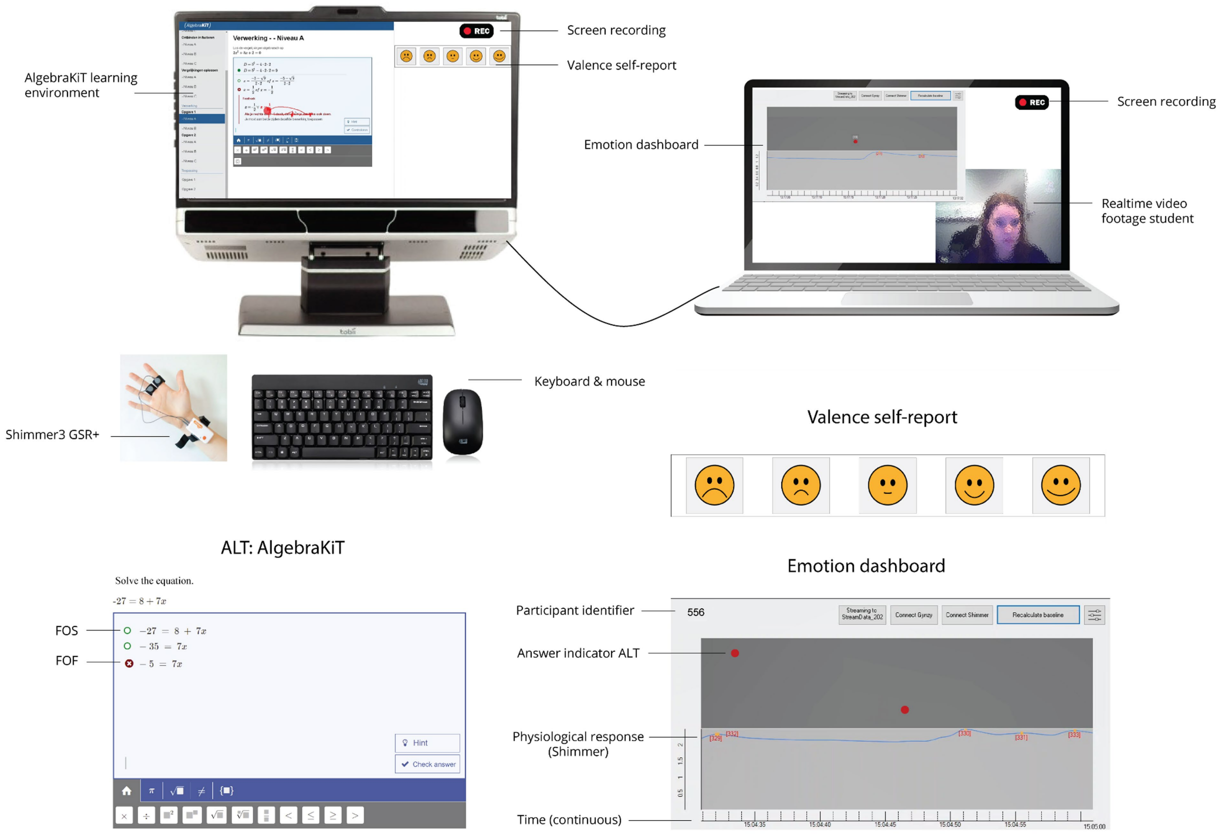


FIGURE 2
Overall set-up of the experiment.

2.3.4 Instruction

Students were taught the quadratic formula via an instruction video taken from a public YouTube account (Jawiskunde, 2017). This 6-min video explained all components of the quadratic formula and demonstrated in two examples how it could be applied in solving math problems.

2.3.5 Nature video

Students watched a 5-min video of different landscape views accompanied by relaxing music (Cat Trumpet, 2019). This video served to establish a baseline for the EDA measurement, which was used to measure students' physiological arousal.

2.3.6 Emotion dashboard

To ensure the synchronization of the data streams, an emotion dashboard was developed (Figure 2). On this dashboard, the researcher could see a student's answers to every calculation step and physiological responses on a continuous timeline. The on-screen position of the answer indicator (red circle) indicated whether a calculation was correct (at the top of the dark gray part of the screen) or incorrect (at the bottom of the dark gray part of the screen).

2.4 Measurements

2.4.1 Logfile data of the ALT

The ALT stored the following information for each participant: student identifier, exercise identifier, timestamp in milliseconds, and correctness of a calculation.

2.4.2 Background characteristics

The following background characteristics of the participants were orally collected by the researcher: gender, age, field of study, year of study, and level of math in high school.

2.4.3 Pre-test

The pre-test measured students' prerequisite math knowledge for the quadratic formula. The test consisted of three items about removing brackets (e.g., remove the brackets: $(x-12)(x+6)$) and three items on factorizing (e.g., $7x^2+63x+140$), increasing in difficulty from easy to hard. One point was awarded for each correct item, so total pre-test scores could range from 0 to 6. Based on students' pre-test scores, they were assigned to either easy, intermediate or hard practice problems in the main part of the session (see section 2.3.3. for more information on the learning objective). When a student had a score of 2 or less, they were assigned easy problems. A score of 3 or 4 resulted in intermediate problems, and a score of 5 or more in hard practice problems.

2.4.4 Measures of emotional responses

2.4.4.1 Physiological responses

Electrodermal activity (EDA), also called skin conductance, was used to measure physiological arousal throughout the learning session. EDA captures the variation of electrical characteristics of the skin due to sweat gland activity (Dawson et al., 2016). This study used the Shimmer3 GSR+; a wearable device fixed on a wristband with two electrodes placed on the middle phalanges of the index and middle finger of the participant's non-dominant hand. The researcher verified the correct placement of the wristband at the beginning of each session (participants had to place the electrodes themselves due to COVID-19 regulations). EDA data was recorded using a sampling rate of 51.2 hertz (51.2 raw values measured in micro Siemens (μS) per second).

2.4.4.2 Experiential responses

Self-reports of valence were used to indicate experiential responses. A 5-point scale was used, ranging from very negative to very positive. Based on the Smiley-o-meter (Read, 2008), all five scale values were visualized by an emoticon. The researcher prompted students to indicate how they were feeling after every feedback event. The self-report tool was visible on a split screen next to the AlgebraKiT screen (Figure 2).

2.4.4.3 Behavioral responses

Observations of students' facial expressions were used to provide insight into the emotion type (Baker et al., 2010). The Baker Rodrigo Ocumpaugh Monitoring Protocol (Ocumpaugh et al., 2015) was used to classify students' facial expressions as either anxiety, boredom, confusion, disappointment, engaged concentration, enjoyment, frustration, relief or surprise (Pekrun, 2006; D'Mello, 2013). Two observers indicated the facial expression they observed instantly after every feedback event using recordings of the participants' faces. Before, observers were trained using videos of students working on individual practice problems. After the training phase, the agreement between the observers was 78.5%. The observers discussed their disagreement and adjusted the coding scheme when they agreed on distinct features in the face. Formal interrater reliability was calculated with Cohen's kappashere was substantial agreement between the two observers, $\kappa = 0.614$, $p < 0.001$ (Landis and Koch, 1977; Hallgren, 2012).

2.5 Coding of the dependent variables

2.5.1 Coding of physiological responses

EDA data were analyzed using the MATLAB Ledalab toolbox (Benedek and Kaernbach, 2010). Movement artifacts were visually identified and manually deleted. No filtering and down-sampling was applied. Feature extraction was obtained via Continuous Deconvolution Analysis (CDA). This analysis divides the EDA signal into a tonic and a phasic component. The tonic component is a slowly varying signal that generates a moving baseline per individual (relatively stable within a few seconds). The phasic component refers to the fast-moving signal, representing faster-changing elements in the EDA signal, i.e., peaks (Braithwaite et al., 2015; Dawson et al., 2016). Data of one participant is shown in Figure 3. Event-related responses after immediate feedback were extracted by this analysis using a response window of plus 4 s, resulting in stimulus-specific features. This response window was chosen because of recommendations in previous research on fast stimuli and EDA latency of 1 to 3 s (Benedek and Kaernbach, 2010; Dawson et al., 2016; Horvers et al., 2021). Three features were extracted from the EDA signal (see Figure 4), all measured within the response window (wrw) based on previous research (Horvers et al., 2021). These features were *the number of above-threshold peaks*, *the amplitude sum of above-threshold peaks*, and *the mean phasic activity.* An increase in EDA could be classified as a significant peak when it is above a certain threshold; 0.01 μs was used as threshold based on previous research (Horvers et al., 2021).

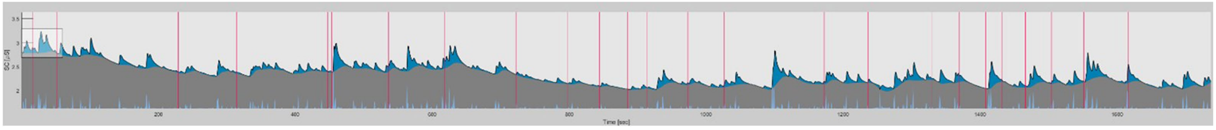


FIGURE 3
Ledalab screenshot of one participant during the whole learning session (www.ledalab.de; Benedek & Kaernbach, 2010). Tonic component (grey), phasic component (blue), and feedback events (red line).

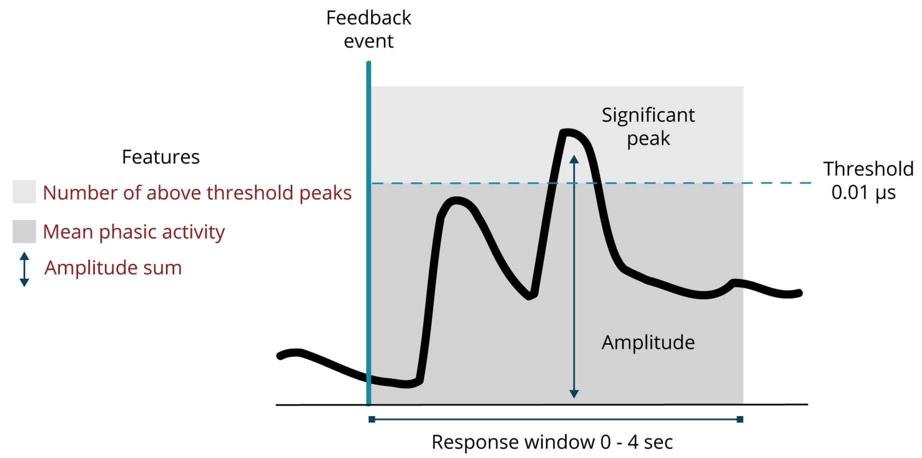


FIGURE 4
EDA features used in this study.

2.5.2 Coding of experiential responses

The five valence options were strong positive, positive, neutral, negative and strong negative. These options were coded according to the 5-point scale: 1 for very negative, 2 for negative, 3 for neutral, 4 for positive, and 5 for very positive. In line with the categorization of emotion types into positive and negative emotion types and neutral facial expressions (see section 2.5.3), the valence options were also merged into three categories, with strong positive and positive in the positive category, strong negative and negative in the negative category and neutral in the neutral category.

2.5.3 Coding of behavioral responses

The observed facial expressions were categorized into negative and positive emotion types and neutral facial expressions. Based on literature, enjoyment (Pekrun and Stephens, 2010; D'Mello, 2013) and relief (Pekrun and Stephens, 2010) were placed in the positive emotion category. Boredom, frustration (Pekrun and Stephens, 2010; D'Mello, 2013), anxiety, disappointment (Pekrun and Stephens, 2010), and confusion (D'Mello, 2013) were placed in the negative emotion category. Engaged concentration was indicated as neutral. Surprise can have negative as well as positive valence (Noordewier and Breugelmans, 2013), due to this ambiguity this emotion was excluded from analyses. After categorization, there was still substantial agreement between the two observers overall ($\kappa = 0.704$, $p < 0.001$), and separately for negative emotion types ($\kappa = 0.792$, $p < 0.001$) and neutral facial expressions ($\kappa = 0.741$, $p < 0.001$). Positive emotion types had a moderate agreement ($\kappa = 0.429$, $p < 0.001$) (Landis and Koch, 1977; Hallgren, 2012). As a next step, the positive and negative emotion types were divided in activating and deactivating emotions (Pekrun et al., 2007; D'Mello et al., 2014). Enjoyment is categorized as positive activating and relief as positive deactivating. Frustration, anxiety and confusion are categorized as negative activating and boredom and disappointment are negative deactivating (Pekrun et al., 2007; D'Mello et al., 2014).

2.6 Data analysis

For the analysis of physiological responses, the number of above-threshold peaks, the amplitude sum of above-threshold peaks, and the mean phasic activity were used. A one-sample t-test analyzed whether immediate feedback triggered an above-threshold peak in EDA at above-chance level. Repeated measures MANOVA was used to examine differences in students' physiological responses between FOF and FOS. Dependent variables were the proportion of above-threshold peaks, the amplitude sum and the mean phasic activity. For the analysis of experiential responses, the proportions of positive valence and negative valence were used, averaged per participant. To examine students' experiential responses, repeated measures ANOVAs with Greenhouse–Geisser correction were run separately for FOF and FOS with the proportions of negative and positive valence as dependent variables. For the analysis of behavioral responses, the proportions of positive emotion types and negative emotion types were used, averaged per participant. A similar approach was used to examine within-subject differences in behavioral responses, with the proportions of negative and positive emotion types as dependent measures. To investigate the difference in deactivating and activating emotion types, repeated measures ANOVAs were used.

TABLE 1 Descriptives of students' physiological responses to feedback.

	Total		FOF		FOS	
	M	*SD*	*M*	*SD*	*M*	*SD*
Proportion of above-threshold peaks	0.75	0.24	0.73	0.27	0.77	0.26
Amplitude sum	0.24	0.26	0.28	0.34	0.23	0.26
Mean phasic activity	0.34	0.35	0.39	0.45	0.32	0.32

FOF, feedback on failure; FOS, feedback on success. $N = 32$.

TABLE 2 Descriptives of students' experiential responses to feedback.

	Total		FOF		FOS	
	M	*SD*	*M*	*SD*	*M*	*SD*
Proportion positive valence	0.47	0.21	0.06	0.10	0.77	0.24
Proportion negative valence	0.30	0.21	0.66	0.30	0.02	0.05
Proportion neutral valence	0.23	0.16	0.28	0.26	0.21	0.24

FOF, feedback on failure; FOS, feedback on success. $N = 32$.

3 Results

Students received immediate feedback 990 times in total. The average number of feedback events per student was 30.94 ($SD = 11.99$). Students received feedback on failure (FOF; $M = 15.03$, $SD = 9.09$) about as often as feedback on success (FOS; $M = 15.56$, $SD = 7.83$).

3.1 Physiological responses

Feedback yielded 1.17 above-threshold peaks in EDA on average (maximum of 5 above-threshold peaks). To investigate whether feedback triggered an above-threshold peak at an above-chance level, a one-sample *t*-test was run. This test indicated that the proportion of above-threshold peaks was significantly higher than chance level (0.50), $t(31) = 5.965$, $p < 0.001$.

Table 1 shows the descriptives of physiological responses for the three features that are extracted from the EDA signal (proportion of above-threshold peaks, amplitude sum and mean phasic activity). The columns show the grand means of these three variables and the means for FOF and FOS. The mean proportion of above-threshold peaks was slightly lower for FOF (0.73) than for FOS (0.77). The average amplitude sum of the above-threshold peaks was 0.24, and slightly higher amplitudes were observed for FOF (0.28) than for FOS (0.23). The same goes for the mean phasic activity. Repeated measures MANOVA showed that these minor differences were not statistically significant, $F(3, 30) = 1.682$, $p = 0.193$, partial $\eta^2 = 0.153$.

3.2 Experiential responses

Students reported the valence of their emotions 26.13 times on average during the learning session ($SD = 11.05$). Valence was indicated after 84.4% of the feedback events, and no indication of valence was given after 15.6% of the feedback events. For FOF, the valence was indicated 11.84 times on average ($SD = 8.03$), and for FOS slightly more often ($M = 14.28$, $SD = 7.04$).

On average, students reported more positive valence (0.47) than negative valence (0.30), although the occurrence rates differed considerably based on the type of feedback (Table 2). Repeated measures ANOVA with Greenhouse–Geisser correction showed a significant difference between the proportion negative and positive valence after FOF, $F(1, 31) = 84.274$, $p < 0.001$, partial $\eta^2 = 0.731$. This result indicates that students predominantly expressed negative valence after FOF. A reverse pattern was found for FOS, where the proportion of positive valence exceeded the proportion of negative valence. Repeated measures ANOVA with Greenhouse–Geisser correction indicated that this difference was statistically significant, $F(1, 31) = 252.996$, $p < 0.001$, partial $\eta^2 = 0.891$.

3.3 Behavioral responses

Students' emotion type was observed 29.00 times on average ($SD = 11.97$). There were slightly higher frequencies for FOS ($M = 15.21$, $SD = 7.77$) than for FOF ($M = 12.65$, $SD = 7.88$). Emotion types were recorded after 92.4% of the feedback events; the remaining 7.6% of the events had missing observations.

Engaged concentration was the prevailing facial expression (0.71) (Table 3). For FOF, a trend was visible with higher proportions of negative emotion types than positive emotion types on average. Repeated measures ANOVA with Greenhouse–Geisser correction showed a significant difference between negative and positive emotion types for FOF, $F(1, 31) = 48.044$, $p < 0.001$, partial $\eta^2 = 0.608$. This indicates that students showed significantly more negative emotion types after FOF than positive emotion types. Students showed significantly more negative activating than negative deactivating emotion types after FOF, $F(1, 31) = 38.277$, $p < 0.001$, partial $\eta^2 = 0.533$. Regarding FOS, the proportions of positive emotion types exceeded the negative emotion types on average. Repeated measures ANOVA with Greenhouse–Geisser correction produced a significant difference between negative and positive emotion types, $F(1, 31) = 10.888$, $p = 0.002$, partial $\eta^2 = 0.260$. This indicates that students showed more positive emotion types after FOS than negative emotion types. Students showed significantly more positive activating than positive deactivating emotion types after FOS, $F(1, 31) = 13.366$, $p < 0.001$, partial $\eta 2 = 0.301$.

4 Discussion

The main purpose of this study was to explore university students' emotional responses to immediate feedback provided by an ALT. The first goal of this study was to investigate students' physiological responses by analyzing to what extent immediate feedback triggers physiological arousal as measured by peaks in EDA. The results indicate that the proportion of above-threshold peaks after feedback

TABLE 3 Descriptives of students' behavioral responses to feedback.

	Total		FOF		FOS	
	M	*SD*	*M*	*SD*	*M*	*SD*
Proportion positive emotion types	*0.10*	*0.11*	*0.03*	*0.06*	*0.15*	*0.18*
Enjoyment	0.09	0.10	0.03	0.06	0.13	0.17
Relief	0.01	0.02	0.00	0.00	0.02	0.04
Proportion negative emotion types	*0.19*	*0.15*	*0.39*	*0.29*	*0.03*	*0.08*
Anxiety	0.01	0.00	0.03	0.12	0.00	0.00
Boredom	0.00	0.01	0.00	0.00	0.00	0.01
Confusion	0.10	0.11	0.18	0.19	0.03	0.08
Disappointment	0.03	0.04	0.06	0.12	0.00	0.00
Frustration	0.06	0.08	0.11	0.14	0.00	0.00
Proportion neutral emotion type	*0.71*	*0.19*	*0.55*	*0.30*	*0.81*	*0.18*
Engaged concentration	0.71	0.19	0.55	0.30	0.81	0.18

FOF, feedback on failure; FOS, feedback on success. $N=32$.

exceeded chance level. Physiological responses were not stronger after FOF than FOS. The second goal of this study was to examine students' experiential and behavioral responses to different types of immediate feedback. The results show that students' experiential responses entailed mostly positive valence after FOS and mostly negative valence after FOF. Regarding behavioral responses, FOF elicited significantly more negative than positive emotion types and significantly more positive than negative emotion types were elicited by FOS. FOF elicited significantly more negative activating than deactivating emotion types, and significantly more positive activating than deactivating emotion types were elicited by FOS.

In line with hypothesis 1, there were more peaks in EDA after immediate feedback than would be expected based on chance. This result indicates that feedback likely elicits an increase in physiological arousal, or at least that variations will occur after feedback, as was found in previous research (Aghaei Pour et al., 2010; Malmberg et al., 2019a). Hypothesis 2 predicted that FOF would elicit stronger physiological responses than FOS, as previous research indicated particularly strong emotions after failure (Rowe et al., 2014; Hill et al., 2021). This hypothesis was not supported by the results, as the three physiological arousal indicators showed an inconsistent pattern. Although there are slightly different values for the amplitude sum of above-threshold peaks and the mean phasic activity, FOF indeed yielded slightly higher values than FOS, but not significantly. In contrast, FOS produced slightly more above-threshold peaks in EDA than FOF, but also not significantly. These divergent results are potentially due to the analysis method, which could have overestimated the number of above-threshold peaks (Thammasan et al., 2020). A possible solution is to use sparse recovery methods or accelerometer data (Kelsey et al., 2018; Thammasan et al., 2020). This can be a result of low power, as a post-hoc power analysis showed that the sample size was too small (power: 21.9%). Moreover, the frequency of peaks could also have been overestimated because of the chosen threshold. In previous research, a threshold of 0.05 μs instead of 0.01 μs is often used as well (Pijeira Díaz, 2019; Malmberg et al., 2019b). However, these studies mostly used older devices for which the standard is 0.05 μs (Horvers et al., 2021). To conclude, electrodermal activity can be a valuable objective indicator of emotional responses after immediate feedback but should at least be accompanied by one other data stream in order to fully understand students' emotional responses. The valence scale used in this study also gained insights into the strength of the emotional response, future research could combine this with the arousal data to gain insights into the strength of emotions.

Previous research on the relationship between feedback and emotions is typically mixed, but most studies found that FOF elicits negative emotions and FOS leads to positive emotions (D'Mello et al., 2010; Lipnevich et al., 2021). The present study replicates these findings and, hence, substantiates hypothesis 3. An interesting result is that students indicated more positive valence overall, even though there were comparable numbers of FOF and FOS events in this study. This is an argument for using multimodal data streams to pinpoint what actually happens after FOF and FOS. Similar results were obtained for students' behavioral responses to immediate feedback. As predicted by hypothesis 4, significantly more negative than positive emotion types occurred in observations following FOF and more positive than negative emotion types in observations following FOS. After both FOF and FOF, activating emotion types occurred significantly more than deactivating emotion types. These results extend existing emotion and feedback theories by showing that emotional responses to immediate and delayed feedback are comparable. Moreover, this study shows that both valence as measured by self-report and observations of emotion types via facial expressions can be used as additional data streams to understand emotional responses.

This study makes a significant scientific contribution to the field of emotions during learning and specifically emotional responses to immediate feedback. The unique contribution of this study is, firstly, its focus on immediate feedback on every calculation students enter into the ALT, as most previous research focused on delayed feedback or feedback after each task (Jarrell et al., 2017; Hill et al., 2021). Secondly, contrary to prior research, which mainly relied on retrospective measures of emotions (Jarrell et al., 2017; Hill et al., 2021), this study measured emotions during the learning process by prompting students to indicate valence and observing their facial expressions after every feedback event and continuously measuring physiological arousal. Lastly, this study uses a multimodal approach instead of a unimodal approach. Most research on the relationship between feedback and emotion used a single data stream (Lim et al., 2020; Hill et al., 2021). This unimodal approach has some constraints, such as the possibility for participants to control their self-reported answers (Pekrun, 2020). The multimodal approach used in this study overcomes these constraints by using continuous measures of arousal (physiological responses), valence self-reports (experiential responses) and observations of emotion type via facial expressions (behavioral responses) to capture emotional responses. This combination of measures is in line with recommendations to view emotions as multi-componential in nature (Harley et al., 2015; Li et al., 2021). This study is one of the first to show how multimodal measurement of emotional responses to immediate feedback in the context of adaptive learning technologies can be performed. However, future research should combine the multimodal data streams in their analyses, to gain even more insight in the multi-componential nature of emotion.

A limitation of this study is that data were collected in the research laboratory. Previous research has substantiated the importance of investigating emotional responses in authentic settings because students may show different responses under controlled circumstances (D'Mello, 2013). Our results may, therefore, not generalize to students' regular classes at the university because not all universities use ALTs in their daily classes and teaching yet, but some examples exist (Gillebaart and Bellinga, 2018). However, 60 to 70% of the pupils in Dutch primary schools use ALTs on a daily basis (Van Wetering et al., 2020; Horvers et al., 2021). Contrary to the ALT used in this study, these primary education ALTs mostly automatically adjust the difficulty of tasks to the ability level of students (Klinkenberg et al., 2011; Van Wetering et al., 2020). Therefore, we recommend replicating the present study with a younger group of learners.

Another potential direction for further research would be to compare physiological responses to feedback to the occurrence and magnitude of other peaks in EDA during the learning process. This could show which variations there are after feedback compared to other moments in learning. An additional suggestion for future research would be to extend all measurements of emotion beyond feedback events, not only physiological arousal. Experiential and behavioral responses could be measured whenever a peak in arousal occurs to investigate learning processes in a fine-grained manner. The multimodal data streams that are used provide insights into a detailed level of emotional responses by using a continuous measure of physiological arousal and self-reports and observations after each calculation step students take. Future research could compare physiological, experiential and behavioral responses to feedback to these responses at other critical moments in learning, such as calculating an answer or receiving a new problem (Fritz et al., 2014).

Adaptive learning technologies currently only use students' cognitive achievements to base the difficulty level of problems and immediate feedback on (Klinkenberg et al., 2011). However, as emotions can both hamper and improve learning, there is an opportunity for ALTs to take students' emotions into account, thereby moving from "cold" to "warm" technologies (Götz and Hall, 2013; Loderer et al., 2020). The results of this study indicate that students mostly show negative emotional responses to FOF and positive emotional responses to FOS, but as these are averages, these responses do not always occur. Sometimes students can react negatively to FOS as well, as also indicated by previous research, for instance, by eliciting boredom when a task is too easy (Pekrun, 2006; Inventado et al., 2011). Addressing these differences in emotional responses to feedback is an important first step to move from cold to warm technologies. The effects of these emotional responses to feedback on learning outcomes should be addressed as well in future research. Moreover, individual differences between students should be taken into account because some students prefer easy tasks to avoid negative emotions and others experience less emotions when they fail than others when challenged by difficult tasks (Baker et al., 2010). A next step for future research is to investigate these individual differences in emotions even more by looking at emotional responses to different difficulty levels of tasks, to ultimately arrive at warm technologies that can accommodate these individual differences.

5 Conclusion

This study is an important first step to understanding students' emotional responses to immediate feedback and consequently moving from cold to warm technologies. Results show that immediate feedback often elicits peaks in EDA and thus high physiological arousal — which is generally taken as a sign that an emotion is triggered — with no differences in strength between feedback on failure or feedback on success. This result indicates that multiple data streams are needed to capture emotional responses. Both experiential responses measured by self-reported valence and behavioral responses measured by observations of emotion types can be used as an additional data source. Feedback on failure elicits predominantly negative emotions, while feedback on success elicits mostly positive emotions. Both feedback on failure and success elicited mostly activating emotions. To conclude, these results imply that emotional responses to immediate feedback can be validly assessed from multimodal data streams, which aligns with the theoretical notion that emotions are multi-componential in nature. These insights provide a good starting point for going from cold to warm technologies that take students' emotions into account.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee Social Sciences (ECSS) of the Radboud University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AH: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. IM: Conceptualization, Methodology, Writing – review & editing. HV: Data curation, Writing – review & editing. TB: Conceptualization, Methodology, Writing – review & editing. AL: Conceptualization, Methodology, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aghaei Pour, P., Hussain, M. S., AlZoubi, O., D'Mello, S., and Calvo, R. A. (2010). "The impact of system feedback on learners' affective and physiological states" In J. Kay & V. Aleven (Eds.), *Proceedings of 10th International conference on intelligent tutoring systems* (Berlin, Germany: Springer-Verlag) 264–273.

Aleven, V., McLaughlin, E. A., Glenn, R. A., and Koedinger, K. R. (2016). "Instruction based on adaptive learning technologies" in R. E. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction* (Routledge) 2, 522–560.

Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., et al. (2022). Lessons learned and future directions of Meta tutor: leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Front. Psychol.* 13, 1–23. doi: 10.3389/fpsyg.2022.813632

Baker, R. S. (2021). Artificial intelligence in education: bringing it all together. OECD digital education outlook 2021 pushing the Frontiers with artificial intelligence, Blockchain and Robots: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots, 43

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., and Graesser, A. C. (2010). Better to be frustrated than bored: the incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *Int. J. Hum.-Comput. Stud.* 68, 223–241. doi: 10.1016/j.ijhcs.2009.12.003

Benedek, M., and Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* 190, 80–91. doi: 10.1016/j.jneumeth.2010.04.028

Braithwaite, J., Watson, D., Jones, R., and Rowe, M. (2015). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments, 1–43. doi: 10.1017.S0142716405050034

Cat Trumpet. (2019). 3 Hours of Amazing Nature Scenery & Relaxing Music for Stress Relief [Video]. YouTube. Available at: https://www.youtube.comwatch?v=BHACKCNDMW8&t=581s

Cloude, E. B., Wortha, F., Dever, D. A., and Azevedo, R. (2021). Negative emotional dynamics shape cognition and performance with Meta tutor: toward building affect-aware systems. 9th International Conference on Affective Computing and Intelligent Interaction, ACII 2021, October.

Coppin, G., and Sander, D. (2021). "Theoretical approaches to emotion and its measurement" in H. L. Meiselman (Eds.), *Emotion measurement* (Woodhead Publishing) 3–37.

D'Mello, S. K. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *J. Educ. Psychol.* 105, 1082–1099. doi: 10.1037/a0032674

D'Mello, S. K. (2017). "Emotional learning analytics" in C. Lang, G. Siemens, A. Wise, and D. Gašević (Eds.) *Handbook of learning analytics* (Beaumont, Alberta: Society for Learning Analytics Research) (SoLAR) 115–127.

D'Mello, S. K., Lehman, B., Pekrun, R., and Graesser, A. (2014). Confusion can be beneficial for learning. *Learn. Instr.* 29, 153–170. doi: 10.1016/j.learninstruc.2012.05.003

D'Mello, S. K., Lehman, B., and Person, N. (2010). Monitoring affect states during effortful problem solving activities. *Int. J. Artif. Intell. Educ.* 20, 361–389. doi: 10.3233/JAI-2010-012

Dawson, M. E., Schell, A. M., and Filion, D. L. (2016). "The electrodermal system" In J. Cacioppo, L. G. Tassinary, G. G. Berntson (Eds.), *Handbook of psychophysiology*. 4th ed

Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., and De Weerdt, J. (2021). A review of automated feedback systems for learners: classification framework, challenges and opportunities. *Comp. Educ.* 162:104094. doi: 10.1016/j.compedu.2020.104094

Egger, M., Ley, M., and Hanke, S. (2019). Emotion recognition from physiological signal analysis: a review. *Electr. Notes Theor. Comp. Sci.* 343, 35–55. doi: 10.1016/j.entcs.2019.04.009

Ekman, P. (1999). "Basic emotions" in T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion*, 98, 45–60.

Elo, A. (1978). *The rating of chess players, past and present* New York: Arco.

Eteläpelto, A., Kykyri, V. L., Penttonen, M., Hökkä, P., Paloniemi, S., Vähäsantanen, K., et al. (2018). A multi-componential methodology for exploring emotions in learning: using self-reports, behaviour registration, and physiological indicators as complementary data. *Frontline Learn. Res.* 6, 6–36. doi: 10.14786/flr.v6i3.379

Fritz, T., Begel, A., Müller, S. C., Yigit-elliott, S., and Züger, M. (2014). Using psycho-physiological measures to assess task difficulty in software development categories and subject descriptors. In Proceedings of the 36th International Conference on Software Engineering, 402–413.

Gillebaart, T., and Bellinga, P. (2018). From co-created learning resources to personalised online courses for blended learning. In looking Back, looking forward. Proceedings of the 10th International Conference on Teaching Statistics (ICOTS10 2018) Kyoto, Japan.

Götz, T., and Hall, N. C. (2013). "Emotion and achievement in the classroom" in J. Hattie & E. M. Anderman (Eds.) *International guide to student achievement* (New York: Routledge), 192–195.

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

Graesser, A. C., and D'Mello, S. K. (2012). Moment-to-moment emotions during reading. *Read. Teach.* 66, 238–242. doi: 10.1002/TRTR.01121

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol.* 8, 23–34. doi: 10.1080/11035896009449194

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., and Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Comput. Hum. Behav.* 48, 615–625. doi: 10.1016/j.chb.2015.02.013

Hill, J., Berlin, K., Choate, J., Cravens-Brown, L., McKendrick-Calder, L., and Smith, S. (2021). Exploring the emotional responses of undergraduate students to assessment feedback: implications for instructors. *TLI* 9, 294–316. doi: 10.20343/teachlearninqu.9.1.20

Horvers, A., Tombeng, N., Bosse, T., Lazonder, A. W., and Molenaar, I. (2021). Detecting Emotions through Electrodermal Activity in Learning Contexts: A Systematic Review. *Sensors* 21, 7869.

Inventado, P. S., Legaspi, R., Suarez, M., and Numao, M. (2011). Predicting student emotions resulting from appraisal of its feedback. *Res. Prac. Technol. Enhanc. Learn.* 6, 107–133.

Jarrell, A., Harley, J. M., Lajoie, S., and Naismith, L. (2017). Success, failure and emotions: examining the relationship between performance feedback and emotions in diagnostic reasoning. *Educ. Technol. Res. Dev.* 65, 1263–1284. doi: 10.1007/s11423-017-9521-6

Jawiskunde. (2017). Abc-formule [Video]. YouTube. Available at: https://www.youtube.com/watch?v=z1Lv5IXnzAE

Kelsey, M., Akcakaya, M., Kleckner, I. R., Palumbo, R. V., Barrett, L. F., Quigley, K. S., et al. (2018). Applications of sparse recovery and dictionary learning to enhance analysis of ambulatory electrodermal activity data. *Biomed. Sig. Process. Control* 40, 58–70. doi: 10.1016/j.bspc.2017.08.024

Kleinginna, P. R., and Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motiv. Emot.* 5, 263–291. doi: 10.1007/BF00993889

Klinkenberg, S., Straatemeier, M., and Van Der Maas, H. L. J. (2011). Computer adaptive practice of Maths ability using a new item response model for on the fly ability and difficulty estimation. *Comp. Educ.* 57, 1813–1824. doi: 10.1016/j.compedu.2011.02.003

Kort, B., Reilly, R., and Picard, R. W. (2001). An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion. In Proceedings-IEEE International Conference on Advanced Learning Technologies IEEE.

Lal, S., Eysink, T. H., Gijlers, H. A., Verwey, W. B., and Veldkamp, B. P. (2021). Detecting emotions in a learning environment: A multimodal exploration in *Proceedings of EC-TEL (Doctoral consortium)*. Springer Cham.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33:159. doi: 10.2307/2529310

Li, S., Zheng, J., Lajoie, S. P., and Wiseman, J. (2021). Examining the relationship between emotion variability, self-regulated learning, and task performance in an intelligent tutoring system. *Educ. Technol. Res. Dev.* 69, 673–692. doi: 10.1007/s11423-021-09980-9

Lim, L. A., Dawson, S., Gašević, D., Joksimović, S., Pardo, A., Fudge, A., et al. (2020). Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: an exploratory study of four courses. *Assess. Eval. High. Educ.* 46, 339–359. doi: 10.1080/02602938.2020.1782831

Lipnevich, A. A., Murano, D., Krannich, M., and Goetz, T. (2021). Should I grade or should I comment links among feedback, emotions, and performance. *Learn. Individ. Differ.* 89:102020. doi: 10.1016/j.lindif.2021.102020

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

Malmberg, J., Haataja, E., Seppänen, T., and Järvelä, S. (2019a). Are we together or not? The temporal interplay of monitoring, physiological arousal and physiological synchrony during a collaborative exam. *Int. J. Comput.-Support. Collab. Learn.* 14, 467–490. doi: 10.1007/s11412-019-09311-4

Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X., and Siipo, A. (2019b). Going beyond what is visible: what multichannel data can reveal about interaction in the context of collaborative learning? *Comput. Hum. Behav.* 96, 235–245. doi: 10.1016/j.chb.2018.06.030

Martin, F., Sun, T., and Westine, C. D. (2020). A systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. *Comp. Educ.* 159, 1903–1929. doi: 10.1016/j.compedu.2020.104009

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., and Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion* 5, 175–190. doi: 10.1037/1528-3542.5.2.175

Mauss, I. B., and Robinson, M. D. (2009). Measures of emotion: a review. *Cognit. Emot.* 23, 209–237. doi: 10.1080/02699930802204677

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education* 57, 632–645.

Noordewier, M. K., and Breugelmans, S. M. (2013). On the valence of surprise. *Cognit. Emot.* 27, 1326–1334. doi: 10.1080/02699931.2013.777660

Obergriesser, S., and Stoeger, H. (2020). Students' emotions of enjoyment and boredom and their use of cognitive learning strategies – how do they affect one another? *Learn. Instr.* 66:101285. doi: 10.1016/j.learninstruc.2019.101285

Ocumpaugh, J., Baker, R. S., and Rodrigo, M. M. T. (2015). Baker Rodrigo Ocumpaugh monitoring protocol (BROMP) 2.0 technical and training manual, 1–6. doi: 10.1109/EUROCON.2015.7313780,

Ortony, A., Clore, G. L., and Collins, A. (2022). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., and Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *Br. J. Educ. Technol.* 50, 128–138. doi: 10.1111/bjet.12592

Peixoto, F., Mata, L., Monteiro, V., Sanches, C., and Pekrun, R. (2015). The achievement emotions questionnaire: validation for pre-adolescent students. *Eur. J. Dev. Psychol.* 12, 472–481. doi: 10.1080/17405629.2015.1040757

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341. doi: 10.1007/s10648-006-9029-9

Pekrun, R. (2017). Emotion and achievement during adolescence. *Child Dev. Perspect.* 11, 215–221. doi: 10.1111/cdep.12237

Pekrun, R. (2020). Commentary: self-report is indispensable to assess Students' learning. *Frontline Learn. Res.* 8, 185–193. doi: 10.14786/flr.v8i3.637

Pekrun, R. (2022). "Development of achievement emotions" in D. Dukes, A.C. Samson & E. A. Walle (Eds.), *The Oxford handbook of emotional development*, vol. 446. Oxford University Press.

Pekrun, R., Frenzel, A. C., Goetz, T., and Perry, R. P. (2007). The control-value theory of achievement emotions. An integrative approach to emotions in education. In P. A. Schutz & R. Pekrun (Eds.) *Emot. Educ.*, 13–36. Academic Press.

Pekrun, R., and Linnenbrink-Garcia, L. (2014). *International handbook of emotions in education*. London: Routledge.

Pekrun, R., and Stephens, E. J. (2010). Achievement emotions: a control-value approach. *Soc. Personal. Psychol. Compass* 4, 238–255. doi: 10.1111/j.1751-9004.2010.00259.x

Peterson, E. R., Brown, G. T. L., and Jun, M. C. (2015). Achievement emotions in higher education: a diary study exploring emotions across an assessment event. *Contemp. Educ. Psychol.* 42, 82–96. doi: 10.1016/j.cedpsych.2015.05.002

Pijeira Díaz, H. J. (2019). Electrodermal activity and sympathetic arousal during collaborative learning. [Doctoral dissertation, Univeristy of Oulu].

Pitt, E., and Norton, L. (2017). 'Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it. *Assess. Eval. High. Educ.* 42, 499–516. doi: 10.1080/02602938.2016.1142500

Read, J. C. (2008). Validating the fun toolkit: an instrument for measuring children's opinions of technology. *Cogn. Technol. Work* 10, 119–128. doi: 10.1007/s10111-007-0069-9

Rowe, A. D., Fitness, J., and Wood, L. N. (2014). The role and functionality of emotions in feedback at university: a qualitative study. *Aust. Educ. Res.* 41, 283–309. doi: 10.1007/s13384-013-0135-7

Russell, J. A. (1980). A Circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714

Tärning, B. (2018). Review of feedback in digital applications - does the feedback they provide support learning? *J. Inform. Technol. Educ. Res.* 17, 247–283. doi: 10.28945/4104

Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., and Price, M. J. (2021). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learn. Instr.* 72:101200. doi: 10.1016/j.learninstruc.2019.04.001

Thammasan, N., Stuldreher, I. V., Schreuders, E., Giletta, M., and Brouwer, A. M. (2020). A usability study of physiological measurement in school using wearable sensors. *Sensors* 20:5380. doi: 10.3390/s20185380

van Lehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ. Psychol.* 46, 197–221. doi: 10.1080/00461520.2011.611369

Van Wetering, M., Booij, E., and van Bruggen, W. (2020). Education in an artificially intelligent world. *Kennisnet Technol. Compass*, 2019–2020. Available at: https://www.kennisnet.nl/app/uploads/kennisnet/publicatie/Kennisnet-Technology-Compass-2019-2020.pdf

Vogl, E., and Pekrun, R. (2016). "Emotions that matter to achievement: student feelings about assessment" in *Handbook of human and social conditions in assessment* (London: Routledge)

Wu, R., and Yu, Z. (2022). Exploring the effects of achievement emotions on online learning outcomes: a systematic review. *Front. Psychol.* 13:931. doi: 10.3389/fpsyg.2022.977931

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 03 April 2024
DOI 10.3389/fpsyg.2024.1371014

OPEN ACCESS

EDITED BY
Javier Marín-Morales,
Universitat Politècnica de València, Spain

REVIEWED BY
Richard Gevirtz,
Alliant International University, United States
Lei Wang,
Northwestern Polytechnical University, China
Tian Liu,
Xi'an Jiaotong University, China

*CORRESPONDENCE
Shan Cheng
✉ chengshan2020@fmmu.edu.cn
Jianqi Wang
✉ wangjq@fmmu.edu.cn

RECEIVED 15 January 2024
ACCEPTED 20 March 2024
PUBLISHED 03 April 2024

CITATION
Cong L, Yu X, Huang M, Sun J, Lv H, Zhang T, Dang W, Teng C, Xiong K, Ma J, Hu W, Wang J and Cheng S (2024) Enhancing emotion regulation: investigating the efficacy of transcutaneous electrical acupoint stimulation at PC6 in reducing fear of heights. *Front. Psychol.* 15:1371014.
doi: 10.3389/fpsyg.2024.1371014

Enhancing emotion regulation: investigating the efficacy of transcutaneous electrical acupoint stimulation at PC6 in reducing fear of heights

Lin Cong[1,2], Xiao Yu[2], Meiqing Huang[1], Jicheng Sun[3], Hao Lv[2], Taihui Zhang[1], Weitao Dang[1], Chaolin Teng[1], Kaiwen Xiong[1], Jin Ma[1], Wendong Hu[1], Jianqi Wang[2]* and Shan Cheng[1]*

[1]School of Aerospace Medicine, Air Force Medical University, Xi'an, China, [2]School of Biomedical Engineering, Air Force Medical University, Xi'an, China, [3]Center for Military Medicine Innovation, Air Force Medical University, Xi'an, China

This study investigated the impact of transcutaneous electrical acupoint stimulation (TEAS) at Neiguan acupoint (PC6) on the physiological and behavioral responses of participants exposed in virtual height. 40 participants were included in the study and were randomly assigned to either a control group or an intervention group. Participants had an immersive experience with a VR interactive platform that provided somatosensory interaction in height stimulation scenes. Psychological scores, behavioral and cognitive performance, and physiological responses were recorded and analyzed. The results indicated that the intervention group had significantly lower fear scores compared to the control group. Analysis of heart rate variability revealed that the intervention group exhibited improved heart rate variability, indicating enhanced cardiovascular function and emotion regulation. The behavioral and cognitive results demonstrated that the intervention group exhibited higher left eye openness, faster reaction times, and greater movement distance, suggesting enhanced attentional focus, cognitive processing, and reduced avoidance behaviors. These findings suggest that TEAS at PC6 can effectively reduce fear and improve the regulation of physiological and behavioral responses to negative emotional stimuli.

KEYWORDS
transcutaneous electrical acupoint stimulation, fear of heights, PC6, emotion regulation, heart rate variability, virtual height stimulation

1 Introduction

Fear is an emotion closely linked to evolution as an adaptive response to environmental threats. When the fear responses are triggered excessively in non-fearful situations, they may develop fear-related disorders in individuals. Fear of heights is one of the most fundamental basic survival responses, often accompanied by physiological and behavioral changes (Liu et al., 2021). People may exhibit a mix of autonomic, behavioral, and cognitive self-preservation responses when they anticipate danger, including breathing faster, heart rate increases, freezing, and avoidance (Borkar and Fadok, 2021; Christianson, 2021). In some cases, irrational or excessive fear of heights can develop into acrophobia, which can significantly interfere with normal daily activities (Arroll et al., 2017). Acrophobia is estimated to affect around 3.1–6.4% of the general population (Lebeau et al., 2010). Additionally, visual height intolerance (vHI) is also a distressing susceptibility to height stimuli that affects 28% of the population, with about half of those who are susceptible reporting a negative impact on their quality of life (Huppert et al., 2013). From susceptibility to height intolerance to acrophobia, there exists a continuum of cognitive and physiological responses to height exposure (Brandt et al., 2015).

Many therapeutic techniques focus on reducing or eliminating of fear in both humans and animals. Pharmacological treatments, such as protein synthesis inhibitors, have been used in animal studies to alter learned fear, which may have long-term effects on the modulation of fear memories and defensive reactions (Nader et al., 2000; Mckenzie and Eichenbaum, 2011). Such interventions cannot be used safely in humans due to a number of serious side effects, including leukopenia, lipid abnormalities, sleepiness, and weight gain. Apart from pharmaceutical interventions, non-pharmacological strategies have the potential to facilitate to eliminate fear responses as well. Device-based methods, such as transcranial direct-current stimulation, transcranial magnetic stimulation, and transcutaneous vagus nerve stimulation, have been found to improve extinction and alleviate fear responses (Burger et al., 2016; Pennington and Fanselow, 2018; Clarke et al., 2020). However, it is important to note that the effectiveness of device-based techniques in reducing fear may vary. Some research has shown that the application of a-tDCS did not significantly reduce feelings of fear in patients or animals (Manteghi et al., 2017; Wout et al., 2017). These conflicting findings could be attributed to differences in brain targets and stimulation frequencies used in the studies.

Cognitive behavioral therapy (CBT) is a commonly used non-pharmacological strategy that aims to modify maladaptive emotional responses by changing an individual's thoughts, behaviors, or both (Kaczkurkin and Foa, 2015). Cognitive and behavioral flexibility allow us to adapt to different situations, shifting strategies as needed to meet changing environmental demands and also play a central role in evidence-based practice approaches for the treatment of fear and anxiety disorders (Powers et al., 2017). For example, attentional bias toward threat stimuli was found to improve in a study involving a single-session CBT for panic disorder (Kappelmann et al., 2020). About half of the variance in symptom change were explained by early reductions in attentional bias toward threat, which were found to predict better symptomatic improvement at a 1-month follow-up (Reinecke et al., 2013). Exposure therapy is one of the widely used CBT treatment method that has been found to be highly effective for specific phobias. Meta-analytical studies have shown that exposure therapy is more effective compared to no treatment, placebo treatment, and non-exposure-based active therapy conditions (Wolitzky-Taylor et al., 2008). These intervention strategies provide promising approaches for managing and reducing fear and anxiety. However, previous studies have focused on behavioral and the neural correlates of cognitive emotion regulation, and have not focused on investigating emotion regulatory strategies that directly involve the body, despite their effectiveness in clinical populations (Minewiser, 2017; Church et al., 2022; Menevşe and Yayla, 2023). Aside from the cognitive, behavioral and motivational concomitants, emotions have long been recognized as full-body events. Emotional Freedom Techniques (EFT) is a psychophysiological intervention that includes cognitive and somatic elements and it adds the novel ingredient of acupressure. Instead of using needles, practitioners stimulate acupuncture points by tapping on them with their fingertips (Stapleton et al., 2023). Extensive research on Clinical EFT has demonstrated its effectiveness in reducing symptoms of post-traumatic stress disorder (Clond, 2016; Sebastian and Nelms, 2017; Feinstein, 2022). Moreover, body tapping is also well suited for self-application in non-clinical settings because of its simplicity, such as sports performance, public speaking and university exams (Baker and Siegel, 2010; Church, 2010; Rahmi, 2013). The inclusion of body-based techniques in research and practice can provide a more comprehensive understanding and approach to emotion regulation.

Previous studies have primarily focused on emotion regulatory strategies, neglecting the necessary intervention duration. It may be due to the fact that reductions in fear and avoidance are easy to observe and measure, but they occur slowly. In many instances, clinical or experimental, the cognitive shifts are slow to develop, changing over weeks rather than minutes (Rachman, 2015). Several studies have found that EFT can lead to significant reductions in anxiety, even in a single-session application (Chatwin et al., 2016; Jasubhai and Mukundan, 2018; Dincer and Inangil, 2021). The visualization of body tapping as a means of emotional regulation has been shown to effectively alter immediate neural and behavioral responses to emotional stimulation by single-session (Wittfoth et al., 2020). Similarly, tapping acupoints has been shown to increase amygdala activation and decrease hippocampus activation in individuals with flight phobia (Wittfoth et al., 2022). It is suggested that EFT can be capable of regulating fear in the daily lives of healthy individuals, even with minimal time investment.

However, tapping acupuncture points with fingertips has limitations that restrict people's actions and is unsuitable for situations that require physical manipulation. Because manual stimulation of acupuncture points can produce endogenous opioids, increase the production of neurotransmitters and regulate cortisol (Menevşe and Yayla, 2023). Cortisol is the main stress hormone and modulates the autonomic nervous system, reduces heart rate, pain, and anxiety through these neurochemical changes (Lane, 2009; Napadow et al., 2009). Therefore, acupuncture may have the capacity to modulate emotional reactions. But the acupuncture is an invasive intervention technology, it has certain safety risks when used in some scenarios with poor sanitary conditions. One study showed that transcutaneous electrical acupoint stimulation (TEAS) can reduce post-operative stress response and improve heart rate variability, with effects no different from traditional acupuncture (Zhou et al., 2020). TEAS is an emerging therapeutic approach that combines the effects of transcutaneous electrical nerve stimulation (TENS) with acupuncture point stimulation (Szmit et al., 2023). Some studies have shown that acupuncture at PC6 can relieve and control palpitations, and its effect is closely related to autonomic nervous function (Liu et al., 2020; Ye et al., 2023). Based on these findings, we hypothesize that applying TEAS at Neiguan (PC6) can produce a similar emotion regulation effect. This approach utilizes a portable electrical stimulation device that can automatically stimulate acupoints (Air Force Medical University, Xi'an, China) to implement TEAS. This would enable a more convenient and comprehensive intervention, potentially applicable in multiple dangerous situations such as aerial work, driving or flying.

2 Materials and methods

2.1 Participants

Potential participants for the study were recruited via public advertisement at the university, and were screened for fear of heights using self-reported fear scores (Likert 10). Exclusion criteria for participation included vestibular and balance deficits, as well as other neurological or orthopedic disorders that affect postural control. Additionally, the subjects were required to have experience with virtual reality but not have been exposed to virtual heights previously. The study included 40 male participants (age: $M = 23.7$, $SD = 3.81$), who were recruited as volunteers with estimated scores between 6 and 9 (targeting a height-fearful but non-clinical population). One participant were examined but had to be excluded due to technical problems. Each participant had normal or corrected vision. Prior to participation, all individuals provided written informed consent and received 150 RMB as compensation at the end of the experiment. The study was approved by the Ethics Committee of the Air Force Medical University.

2.2 Virtual height stimulation

The virtual environment was created using the Unity3D engine and presented on an HTC Vive (HTC, Taiwan, China) with a 100° field of view, 1,080 × 1,200 pixels per eye, and a refresh rate of 90 Hz. Participants were exposed to the virtual scene via a head-mounted display and their head movements were tracked using an infrared positional tracking camera. Two HTC Vive base stations were used for 360° positional tracking, with the sensors positioned approximately 7 m apart from each other. The virtual scene depicted urban buildings and a continuous flow of vehicles (Figure 1A). After wearing the VR headset, participants stood on an elevated circular platform of the lookout tower. A transparent plate extended from the edge of the tower, allowing participants to observe the visual scene by moving their heads. The lookout platform is situated at a height of 250 m above the ground.

2.3 VR interactive platform

In order to realize the somatosensory interaction of VR in virtual height stimulation scenes, a carbon structural steel plate with a size of 4.8 m × 0.5 m is designed (Figure 1B). The plate has a modular design, allowing for quick and flexible deployment in various indoor and outdoor environments. The base is equipped with two cylinder spring shock absorbers, providing vertical freedom of movement for the plate. Since the two spring shock absorbers can work independently, they also provide the bridge floor with a certain degree of lateral freedom, allowing people to experience multi-angle stress shaking while walking. To facilitate VR device deployment, a special tracker installation slot is designed at the front end of the plate, enabling an immersive experience through synchronous combination of virtual and real elements. The corners around the plate are protected with flexible anti-collision cotton, and the ground is equipped with a safety cushion to ensure the safety of the subjects.

2.4 Psychological measures

Participants verbally reported their level of fear using a one-question 10-point Likert scale, ranging from 0 (no fear) to 9 (extreme fear). Additionally, the PANAS (Positive and Negative Affect Scale), consisting of 20 terms that describe different positive (e.g., active, inspired) and negative (e.g., scared, nervous) emotions, was

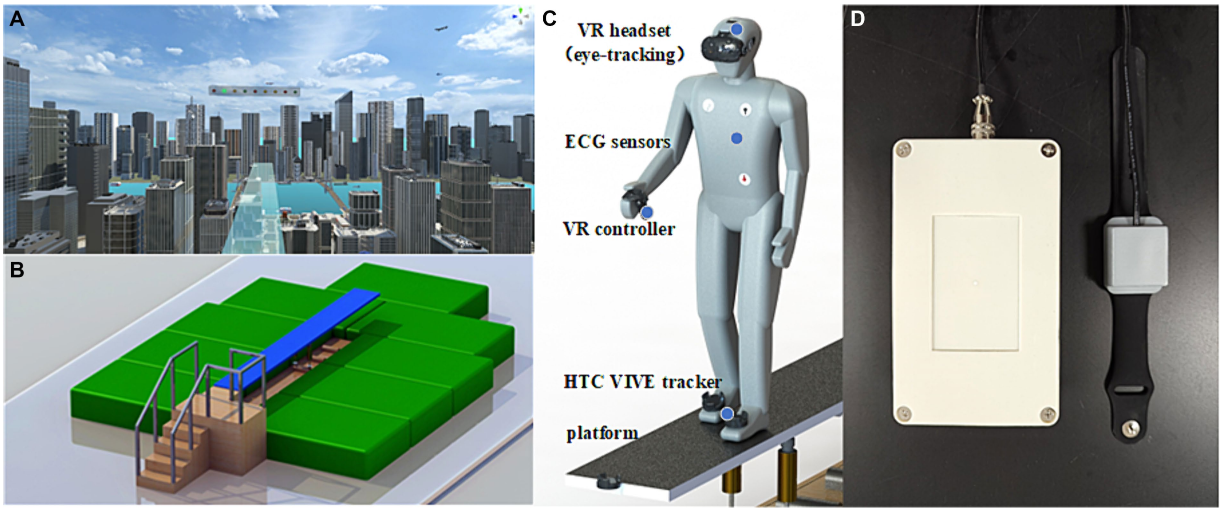


FIGURE 1
The schematic depiction of the experimental set up and the view of the diving board scene at 250 m above ground. **(A)** Participants were exposed to the virtual height scene using a HMD. **(B)** The VR interactive platform can provid somatosensory shaking through cylinder spring shock absorbers when participants walk on it. **(C)** Participants utilized VR controllers to complete cognitive tasks, and two HTC VIVE tracker were strapped to the instep of both feet to record movement distance. Additionally, three electrodes was used to monitor ECG by placing under the right collarbone, on the lower left costal arch, and on the lower left back. **(D)** The portable electrical stimulation device were used to stimulate PC6 during fear induction.

TABLE 1 Summary of the parameters of HRV in this study.

Method	Parameter	Unit	Description
Time-domain	MeanHR	bpm	The mean heart rate
	SDNN	ms	Standard deviation of RR intervals
	RMSSD	ms	Square root of the mean squared differences between successive RR intervals
	NN50	beats	Number of successive RR interval pairs that differ more than 50 ms
	pNN50	%	NN50 divided by the total number of RR intervals
	HRVTi		The integral of the RR interval histogram divided by the height of the histogram
	TINN	ms	Baseline width of the RR interval histogram
Frequency-domain	lnVLF		Natural logarithm of absolute powers in the very-low-frequency band
	lnLF		Natural logarithm of absolute powers in the low-frequency band
	lnHF		Natural logarithm of absolute powers in the high-frequency band
	LF/HF		Ratio between LF and HF band powers
Non-linear	SD1	ms	In Poincaré plot, the standard deviation perpendicular to the line-of-identity
	SD2	ms	In Poincaré plot, the standard deviation along the line-of-identity
	SD1/SD2		Ratio between SD1 and SD2
	ApEn		Approximate entrop
	SampEn		Sample entropy
	DFA,α1		In detrended fluctuation analysis, short term fluctuation slope
	DFA,α2		In detrended fluctuation analysis, long term fluctuation slope
	D_2		Correlation dimension

also employed to evaluate fear and ensure the reliability of the scale score. The scale has reported internal reliability coefficients of alpha = 0.88 and 0.87 for trait positive affect (PA) and negative affect (NA), respectively, based on a sample of 663 university students (Watson et al., 1988). Participants rated each adjective on a scale from 1 (very slightly) to 5 (extremely) to indicate their feelings during the experiment. In this study, we calculated the sum of scores for the negative terms scared, nervous and afraid in order to evaluate the participants' fear level.

2.5 Behavioral and cognitive measures

During the experiment, the participants' position within the virtual environment (VE) was continuously tracked, and their movement distance (MD) was recorded using two HTC VIVE trackers that were strapped to the instep of both feet. The accuracy of the position data from the VIVE trackers has been found to be acceptable when compared to an optoelectronic 3D motion-capturing system, with an average deviation of less than 1 cm in position and an average rotation shift of approximately 1.6° (van der Veen et al., 2019).

Eye-tracking data was also collected using the HTC Vive Pro Eye with a built-in eye tracker, having an accuracy estimation of 0.5°-1.1° and a sampling frequency of 120 Hz. Some studies have used different hardware and reported information provided by the manufacturer to indicate the capability and usability of eye-tracking devices (Wroblewski et al., 2014; Ogura et al., 2019). In this research, the Vive SRanipal SDK is used to access non-filtered and filtered eye-tracking data. The embedded head mounted display (HMD) calibration system is used to calibrate the eye-tracking data for each participant. The raw data provided by the eye tracker includes the origin of the gaze, gaze vectors, eye openness, pupil diameter, and data validity. Our study used the saccade amplitude (calculated based on the origin of the gaze), open openness, and pupil diameter of both eyes for analysis.

To evaluate cognitive performance, a nine-light task which consisted of nine position points in a row with the colors of red, yellow, and green, was used to evaluate cognitive performance (Yang et al., 2023). The participants were required to press the corresponding button on the VR controller when a light was lit in every 5 s. The light would turn off when the correct button was pressed or after 3 s. The participants' reaction time (RT) and accuracy rate (ACC) were automatically recorded during the experiment.

2.6 Physiological measures

Physiological indicators of fear or arousal during exposure to virtual heights were examined by monitoring participants' electrocardiogram (ECG) using a Bluetooth physiological monitor (Tianjin Puray Instruments Ltd., Tianjin, China). The ECG was obtained using three Ag/AgCl electrodes placed under the right collarbone, on the lower left costal arch, and on the lower left back. To analyze heart rate variability (HRV), the ECG data was processed using Kubios software (v.4.1.0, HRV analysis, University of Eastern Finland). HRV indices were calculated, including 7 time domains, 4 frequency domains, and 8 non-linear parameters, as shown in Table 1.

2.7 Experimental design and procedures

All participants were randomly divided into control group (20 subjects) and intervention group (20 subjects). We measured the physiological parameters of the two groups of subjects for 5 min prior to experiment, including heart rate, respiration, oxygen saturation, pulse, and skin temperature. The independent sample t test (data normally distributed) was used to analyze the physiological state of the two groups. There were no significant differences in any of the indices ($p > 0.05$), indicating that the physiological level of the two groups was consistent.

Before exposed in visual height, they were required to verbally report their level of fear, completed the PANAS and then equipped with VR headsets, physiological monitoring devices (Figure 1C), and a portable electrical stimulation (Figure 1D). Then, participants were given a message describing the concept of cognitive tasks in VR and practiced a 10-min nine-light task in a neutral virtual room to reduce the impact of task unfamiliarity. Following the practice session, participants were placed in the center of the VR tracking area and been exposed in the fear stimulation. They were asked to walk back and forth on the VR interactive platform with a fearful situation and complete the Nine-light task in 5 min. During fear induction, participants were given either intervention of TEAS or sham stimulation depending on their group. The control group also wore a portable electrical stimulation device during the experiment, but without receiving electrical stimulation. The stimulation methods used in this experiment were single pulse mode and continuous stimulation with a period of 5 s. After finishing the experiment, participants filled in the PANAS and verbally reported their state of fear again. Physiological, behavioral, and cognitive data were collected continuously during the experiment.

2.8 Data analysis

The normal distribution of the data was determined using the Shapiro–Wilk test and histogram. The independent sample t test were used for data sets with normal distribution, whereas the Mann–Whitney U test was used for data sets without normal distribution. Prior to conducting statistical analyses, the EEG signal were processed with a median filter and a 50 Hz notch filter to remove the baseline drift noise and power frequency noise, and then imported into Kubios HRV. The R-wave time instants are automatically detected by applying the built-in QRS detection algorithm. This in-house developed detection algorithm is based on the Pan–Tompkins algorithm (Pan and Tompkins, 1985). Kubios HRV detects artifacts with an automatic correction method which is more accurate and the method has been validated (Lipponen and Tarvainen, 2019). Artifacts are detected from a time series consisting of differences between successive RR intervals. During the analysis of HRV indices, two participants (one in the control group and another in the intervention group) had to be excluded due to loose electrodes resulting in poor ECG signal quality. Statistical analysis was performed using SPSS software, version 22.0, and a significance level of $p < 0.05$ was set.

3 Results

3.1 Subjective assessment results

The scale data of the two groups after the experiment were statistically analyzed. The results indicated a significant difference in the verbally reported fear scores ($Z = -2.309$, $p < 0.05$). Figure 2A displays a box plot illustrating the verbally reported fear scores, showing that the intervention group had lower scores compared to the control group. Additionally, as depicted in Figure 2B, the PANAS fear score, which includes the negative terms score of scared, nervous, and afraid, also exhibited a significant difference, with the intervention group demonstrating significantly lower scores than the control group ($Z = -2.265$, $p < 0.05$).

3.2 Heart rate variability

Table 2 presents the descriptive statistics and results of the participants' 19 HRV indices in both the control and intervention groups. The analysis of time domain measures revealed that the intervention group had higher SDNN values compared to the control group ($Z = -2.175$, $p < 0.05$). Similarly, the TINN was also significantly different ($Z = -2.584$, $p < 0.01$), with the intervention group exhibiting higher values than the control group. In terms of frequency domain indices, the lnHF showed a similar pattern to SDNN, with significantly

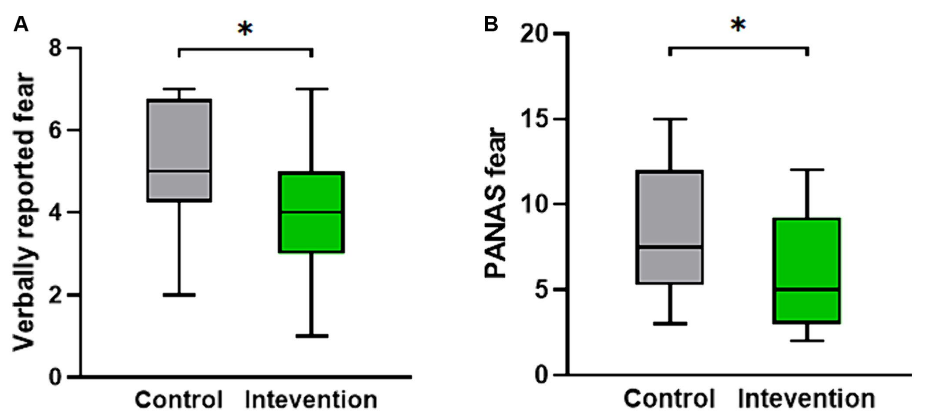


FIGURE 2
The distribution of psychological measures scores for both the intervention group and the control group. **(A)** The verbally reported fear score. **(B)** The PANAS fear score.*$p < 0.05$.

TABLE 2 Descriptive statistics and results of Mann–Whitney *U* tests of HRV indices in the control group and the intervention group.

Method	Parameter	Control Group (*N* = 19)					Intervention Group (*N* = 19)					*Z*
		M	SD	Q1	Med	Q3	M	SD	Q1	Med	Q3	
Time-domain	MeanHR	87.341	11.870	76.533	88.889	96.248	89.609	10.476	81.275	90.102	97.104	−0.598
	SDNN	35.306	9.700	28.247	31.593	38.842	43.727	14.603	33.667	42.046	50.092	−2.175*
	RMSSD	42.971	11.011	37.471	39.205	45.225	48.564	14.099	38.893	46.044	54.215	−1.445
	NN50	85.789	33.985	69.500	81.000	102.000	80.895	37.162	61.500	72.000	103.000	−0.701
	pNN50	21.576	9.785	15.502	18.790	26.554	24.380	10.014	17.568	20.189	29.914	−0.949
	HRVTi	8.969	1.950	7.699	8.820	9.644	9.582	2.260	7.681	9.091	10.846	−0.672
	TINN	185.895	50.329	151.000	165.000	223.000	243.263	80.588	192.500	226.000	272.500	−2.584**
Frequency-domain	lnVLF	4.267	0.789	3.576	4.437	4.820	4.611	0.585	4.177	4.438	5.033	−1.095
	lnLF	5.998	0.648	5.563	5.931	6.458	6.303	0.678	5.824	6.357	6.777	−1.387
	lnHF	5.527	0.703	5.027	5.592	6.070	6.058	0.819	5.459	5.964	6.606	−2.058*
	LF/HF	1.762	0.798	1.093	1.530	2.305	1.508	0.875	0.780	1.302	2.048	−1.182
Non-linear	SD1	30.424	7.799	26.529	27.752	32.027	34.402	9.988	27.548	32.599	38.388	−1.445
	SD2	39.288	12.216	30.333	34.777	44.592	51.115	18.993	36.973	50.227	56.040	−2.35*
	SD1/SD2	0.800	0.152	0.739	0.819	0.851	0.701	0.140	0.600	0.668	0.787	−1.971*
	ApEn	1.233	0.071	1.185	1.250	1.283	1.153	0.169	1.099	1.208	1.270	−1.299
	SampEn	2.010	0.193	1.920	2.057	2.146	1.898	0.291	1.786	1.919	2.112	−1.328
	DFA,α1	0.886	0.155	0.799	0.901	1.010	0.955	0.142	0.851	0.918	1.072	−1.182
	DFA,α2	0.421	0.132	0.347	0.421	0.478	0.407	0.114	0.336	0.389	0.463	−0.394
	D_2	1.933	1.424	0.727	1.611	2.667	2.698	1.492	1.437	2.595	3.992	−1.679

1st quartile (Q1), Median (Med), 3rd quartile (Q3). The significance levels (two-tailed) are marked as follows: $*p < 0.05$, $**p < 0.01$.

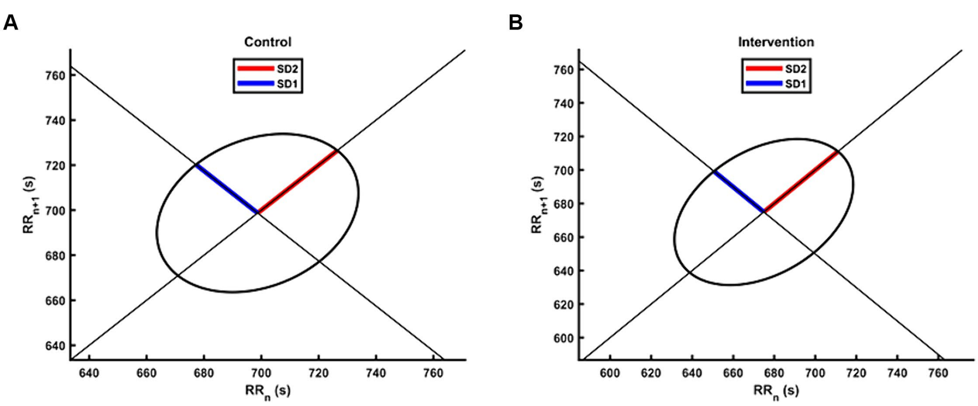


FIGURE 3
The Poincaré plot of HRV for the intervention group and the control group. The group-level scatters were not plotted due to the different number of RR intervals between participants during the 5 min experiment. **(A)** The Poincaré plot for the control group. **(B)** The Poincaré plot for the intervention group.

higher values observed in the intervention group ($Z = -2.058$, $p < 0.05$). In the non-linear measurements, the SD2 was significantly different ($Z = -2.35$, $p < 0.05$), with the intervention group demonstrating higher values than the control group. In contrast to SD2, the SD1/SD2 ratio in the intervention group was smaller than that in the control group ($Z = -1.971$, $p < 0.05$), resulting in a more distinct elliptical shape, as shown in Figure 3. There were no significant differences in other HRV indices ($p > 0.05$).

3.3 Behavioral and cognitive parameters

Table 3 shows the eye movement parameters in both the control and intervention groups. The control group exhibited lower left eye openness compared to the intervention group ($Z = -2.03$, $p < 0.05$) as shown in Figure 4A, but no significant difference was observed in right eye openness. There were no significant differences in saccade amplitude and pupil diameter indices.

TABLE 3 Descriptive statistics and results of Mann–Whitney *U* tests of eye movement indices in the control group and the intervention group.

Parameter	Control Group (N = 20)					Intervention Group (N = 20)					Z
	M	SD	Q1	Med	Q3	M	SD	Q1	Med	Q3	
Left eye saccade amplitude	49.645	6.539	45.328	49.371	51.219	45.425	5.502	42.711	45.221	49.168	−1.731
Right eye saccade amplitude	46.114	4.563	43.277	45.559	47.583	47.022	4.873	43.941	45.973	48.791	−0.73
Left eye pupil diameter	4.331	1.020	3.581	4.357	4.778	3.840	0.496	3.558	3.699	4.145	−1.542
Right eye pupil diameter	4.211	1.162	3.166	4.343	4.941	4.211	0.775	3.625	4.255	4.746	−0.135
Left eye openness	0.882	0.064	0.864	0.889	0.924	0.919	0.042	0.914	0.917	0.951	−2.03*
Right eye openness	0.898	0.056	0.849	0.913	0.942	0.873	0.082	0.840	0.887	0.942	−0.649
Average of both eyes saccade amplitude	47.879	4.495	44.293	47.443	52.130	46.224	4.588	42.873	46.185	49.205	−1.19
Average of both eyes pupil diameter	4.271	1.079	3.303	4.390	4.868	4.026	0.575	3.593	4.020	4.590	−0.433
Average of both eyes eye openness	0.905	0.051	0.870	0.918	0.945	0.879	0.062	0.845	0.890	0.928	−1.313

1st quartile (Q1), Median (Med), 3rd quartile (Q3). The significance levels (two-tailed) are marked as follows: *$p < 0.05$, **$p < 0.01$.

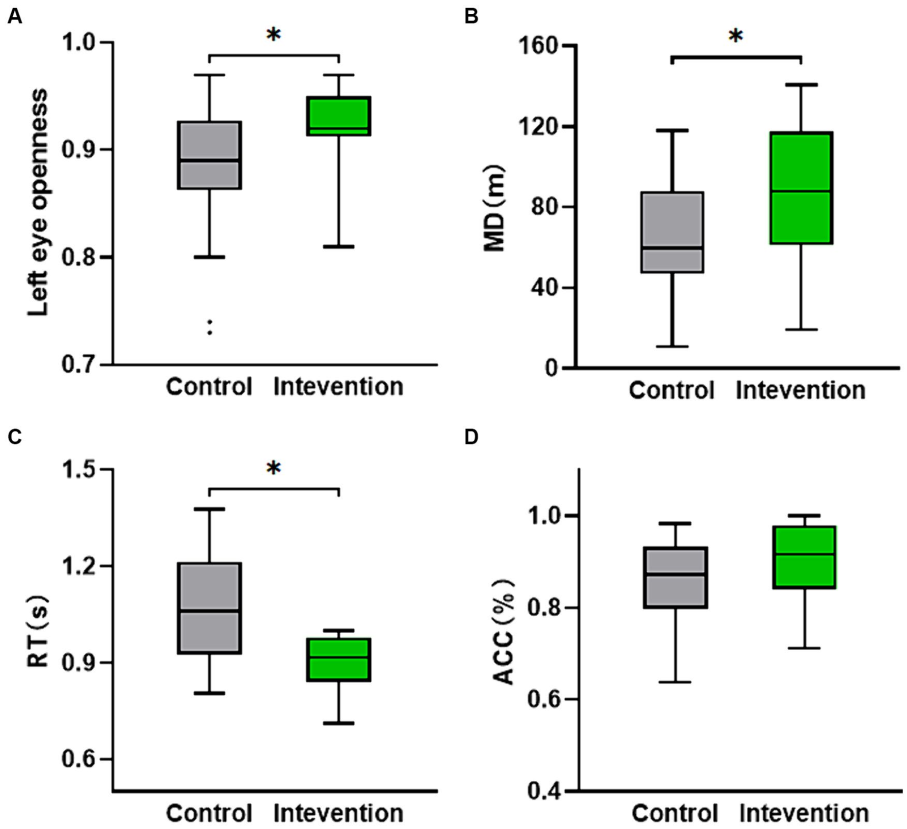


FIGURE 4
The distribution of behavioral and cognitive parameters values for both the intervention group and the control group. **(A)** Left eye openness. **(B)** Movement distance calculated from the participants' location while walking on the VR interactive platform. **(C,D)** The nine-light task performance (ACC and RT). *$p < 0.05$.

The independent samples *t*-test was conducted to analyze the performance of the nine-light task (ACC and RT) and MD. Figure 4B displayed a significant difference in MD ($t = -2.124$, $p < 0.05$), with the intervention group ($M = 84.783$, SD = 36.171) demonstrating a higher value compared to the control group ($M = 63.108$, SD = 27.823). Figure 4C revealed a significant difference in RT ($t = 2.323$, $p < 0.05$), with the intervention group ($M = 0.947$, SD = 0.150) exhibiting a lower value than the control group ($M = 1.064$, SD = 0.167). Although not statistically significant ($p > 0.05$), the ACC of the intervention group was still higher than that of the control group as depicted in Figure 4D.

4 Discussion

In this study, we investigated whether TEAS at PC6 could effectively reduce fear exposed to the virtual height stimulation by single-session. The results demonstrated that the intervention had significant effects on reducing fear, as evidenced by changes in physiological responses and behavioral and cognitive parameters. These findings provide evidence supporting the use of TEAS as a means of emotion regulation in response to negative emotional scenes in healthy participants.

The significant differences observed in verbally reported fear scores and PANAS fear scores between the intervention group and the control group suggest that intervention was effective. These findings align with previous research highlighting the anxiolytic effects of acupuncture (Smith et al., 2018; Fu et al., 2022). It has been well-established that fear and anxiety often coexist and share common physiological and cognitive processes (Etkin and Wager, 2007; Mcteague et al., 2010). Anxiety is characterized by apprehension, worry, and anticipation of future threats, and it can intensify the fear response and prolong emotional arousal. Individuals with anxiety disorders exhibit heightened fear responses and are more prone to developing phobias (Lissek et al., 2009, 2010). Additionally, our findings are consistent with previous research on cognitive-behavioral interventions for specific phobias (Wang et al., 2022). The intervention of TEAS likely affected maladaptive thoughts and beliefs associated with the stimuli, resulting in a reduction in fear and anxiety.

HRV reflects the adaptability and flexibility of the autonomic nervous system and is particularly associated with cardiac vagal tone, which is relevant for various psychophysiological phenomena, including self-regulation mechanisms linked to cognitive, affective, social, and health (Thayer et al., 2009; Mccraty and Shaffer, 2015). Previous research has examined the association between HRV and emotion regulation strategies, suggesting that changes in HRV can reflect the effects of emotional interventions to some extent (Geisler et al., 2013; Zaccaro et al., 2018). In the present study, the control group exhibited lower SDNN values compared to the intervention group, indicating that the intervention had a positive impact on HRV, as reflected by higher SDNN values in the intervention group. Research has shown that SDNN is associated with various physiological and psychological processes. Higher SDNN values have been linked to better cardiovascular health, improved stress resilience, and enhanced emotion regulation (Thayer and Lane, 2009; Laborde et al., 2017). Conversely, lower SDNN values have been associated with increased risk for cardiovascular diseases, poor mental health, and worse anxiety (Quintana and Heathers, 2014; Larsson et al., 2023). Therefore, the intervention may indicate a potential improvement in cardiovascular function and stress regulation. Some researchers suggest using HRV indices that clearly reflect identified physiological systems with a theoretical underpinning, such as RMSSD, peak-valley, and HF (Laborde et al., 2017). The HF band reflects parasympathetic activity and lower HF power is correlated with stress, panic, anxiety, or worry (Shaffer and Ginsberg, 2017). In this study, the intervention group exhibited higher lnHF compared to the control group, indicating that TEAS may have an effect on the parasympathetic activity of the participants. As research suggests, higher vagal tone is associated with better executive cognitive performance and improved emotional and health regulation (Thayer and Lane, 2009; Thayer et al., 2012). Higher vagal tone has also been linked to better emotion regulation abilities, including the ability to downregulate negative emotions and upregulate positive emotions (Kok et al., 2013). These findings support our result that TEAS can increase vagal tone and enhance emotion regulation abilities. A wider distribution with greater variability will result in a larger TINN value, indicating a more irregular heart rate pattern. Conversely, a narrower distribution with less variability will yield a more regular heart rate pattern. The significant difference observed in TINN between the two groups may indicate that the intervention can lead to a more regular heart rate pattern.

Non-linear analyses may be more appropriate and accurate for HRV analysis, as the autonomic nervous system exhibits complex and irregular fluctuations (Piskorski, 2005). SD1 is supposed to be more sensitive to quick and high frequent changes whereas SD2 is viewed as an indicator of long-term changes. The ratio of SD1/SD2, which measures the unpredictability of the RR time series, is used to assess autonomic balance. In our study, the intervention group had significantly higher SD2 compared to the control group, while the SD1/SD2 ratio was significantly lower. According to our result, the vagus and sympathetic balance was affected by the intervention, improving cardiac autonomic modulation and reducing stress state. One study demonstrated that SD2 decreased significantly during a stress session compared to a control session, which was attributed to university examinations (Melillo et al., 2011). This finding supports the conclusion that higher SD2 may represent a lower state of stress.

Behavioral parameters, such as eye movement parameters and task performance, were also assessed in this study. Eye movements and spatial attention are closely interconnected systems that often shift together in many circumstances (Hodgson, 2019). The intervention group exhibited higher left eye openness compared to the control group, indicating a potential improvement in attentional focus and cognitive processing. One possible explanation for the increased left eye openness in the intervention group is that participants needed to intermittently look at the VR displaying buttons while using the VR controller with their right hand for task button interaction. As a result, the left eye observed toward the VR controller generated a greater offset, leading to higher eye openness. People direct their eyes toward objects of interest with the aim of acquiring visual information. However, processing this information is constrained by capacity, requiring task-driven and salience-driven attentional mechanisms to select a few among the many available objects (Souto and Kerzel, 2021). Thus, our observations suggest that attention, driven by TEAS, is biased toward cognitive tasks rather than fear stimuli. Although there were no significant differences in saccade amplitude and pupil diameter, the intervention group exhibited faster RT in the nine-light task and greater MD compared to the control group. The RT is influenced by the stages of cognitive processing involved in a task. Cognitive processing involves several stages, including perception, attention, memory encoding, decision-making, and response execution (Mittelstädt and Miller, 2020). Each stage requires a certain amount of time, and the cumulative time taken by these stages determines the RT. Emotion regulation has been found to have an impact on cognitive performance. Research suggests that individuals who are better able to regulate their emotions tend to exhibit improved cognitive functioning, such as enhanced attention, memory, and task performance (Gross, 2015). The smaller RT indicated that the intervention had a positive effect on reducing participants' fear responses and improving cognitive performance. Furthermore, the greater movement distance observed in the intervention group suggests a reduction in avoidance behaviors and an improvement in coordination and execution of actions. Fear motivates different types of defensive behaviors and these defensive behaviors may in turn reduce, preserve, or amplify fear responding (Pittig et al., 2020). Based on the subjective scores and movement performance, we believe that participants in the intervention group experienced less fear, which led to lower avoidance behavior, manifested by an increase in walking distance.

5 Conclusion

Our study findings indicate that a single-session of TEAS at PC6 effectively reduces fear of heights. The results demonstrate that the intervention group had lower fear scores compared to the control group. Additionally, supporting evidence is derived from improvements in heart rate variability, increased left eye openness, faster reaction times, and greater movement distance observed in the intervention group. These findings suggest that TEAS has the potential to reduce fear, enhance cardiovascular function and emotion regulation, improve attentional focus and cognitive processing, and decrease avoidance behaviors. PC6 may be a suitable acupoint for regulating emotions in response to negative stimuli when using TEAS. However, further research is needed to fully understand the neural mechanisms underlying these effects.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Air Force Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LC: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. XY: Writing – review & editing, Writing – original draft, Methodology. MH: Writing – review & editing, Writing – original draft, Data curation. JS: Writing – review & editing, Writing – original draft, Methodology. HL: Writing – review & editing, Writing – original draft, Methodology. TZ: Writing – review & editing, Writing – original draft, Software, Formal analysis. WD: Writing – review & editing, Writing – original draft, Software, Formal analysis. CT: Writing – review & editing, Writing – original draft, Methodology, Formal analysis. KX: Writing – review & editing, Writing – original draft, Methodology. JM: Writing – review & editing, Writing – original draft, Validation, Formal analysis. WH: Writing – review & editing, Writing – original draft, Conceptualization. JW: Writing – review & editing, Writing – original draft, Conceptualization. SC: Writing – review & editing, Writing – original draft, Conceptualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (72101262) and Key R&D Plan of Shaanxi Province (2022SF-052 and 2022SF-428).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Arroll, B., Wallace, H. B., Mount, V., Humm, S. P., and Kingsford, D. W. (2017). A systematic review and meta-analysis of treatments for acrophobia. *Med. J. Aust.* 206, 263–267. doi: 10.5694/mja16.00540

Baker, A. H., and Siegel, L. S. (2010). Emotional freedom techniques (EFT) reduces intense fears: a partial replication and extension of Wells, Polglase, Andrews, Carrington, & Baker (2003). *Energy Psychol. Theory Res. Treat.* 2, 15–32. doi: 10.9769/EPJ.2010.2.2.AHB.LSS

Borkar, C. D., and Fadok, J. P. (2021). A novel Pavlovian fear conditioning paradigm to study freezing and flight behavior. *J. Vis. Exp.* 5:167. doi: 10.3791/61536

Brandt, T., Kugler, G., Schniepp, R., Wuehr, M., and Huppert, D. (2015). Acrophobia impairs visual exploration and balance during standing and walking. *Ann. N. Y. Acad. Sci.* 1343, 37–48. doi: 10.1111/nyas.12692

Burger, A. M., Verkuil, B., Van Diest, I., Van der Does, W., Thayer, J. F., and Brosschot, J. F. (2016). The effects of transcutaneous vagus nerve stimulation on conditioned fear extinction in humans. *Neurobiol. Learn. Mem.* 132, 49–56. doi: 10.1016/j.nlm.2016.05.007

Chatwin, H., Stapleton, P., Porter, B., Devine, S., and Sheldon, T. (2016). The effectiveness of cognitive behavioral therapy and emotional freedom techniques in reducing depression and anxiety among adults: a pilot study. *Integr. Med.* 15, 27–34.

Christianson, J. P. (2021). The head and the heart of fear. *Science* 374, 937–938. doi: 10.1126/science.abm6790

Church, D. (2010). The effect of EFT (emotional freedom techniques) on athletic performance: a randomized controlled blind trial. *Open Sports Sci. J.* 2, 94–99. doi: 10.2174/1875399X00902010094

Church, D., Stapleton, P., Vasudevan, A., and O'Keefe, T. (2022). Clinical EFT as an evidence-based practice for the treatment of psychological and physiological conditions: a systematic review. *Front. Psychol.* 13:951451. doi: 10.3389/fpsyg.2022.951451

Clarke, P. J. F., Van Bockstaele, B., Marinovic, W., Howell, J. A., Boyes, M. E., and Notebaert, L. (2020). The effects of left DLPFC tDCS on emotion regulation, biased attention, and emotional reactivity to negative content. *Cogn. Affect. Behav. Neurosci.* 20, 1323–1335. doi: 10.3758/s13415-020-00840-2

Clond, M. (2016). Emotional freedom techniques for anxiety. *J. Nerv. Ment. Dis.* 204, 388–395. doi: 10.1097/NMD.0000000000000483

Dincer, B., and Inangil, D. (2021). The effect of emotional freedom techniques on nurses' stress, anxiety, and burnout levels during the COVID-19 pandemic: a randomized controlled trial. *Explore* 17, 109–114. doi: 10.1016/j.explore.2020.11.012

Etkin, A., and Wager, T. D. (2007). Functional neuroimaging of anxiety: a Meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am. J. Psychiatry* 164, 1476–1488. doi: 10.1176/appi.ajp.2007.07030504

Feinstein, D. (2022). Uses of energy psychology following catastrophic events. *Front. Psychol.* 13:856209. doi: 10.3389/fpsyg.2022.856209

Fu, Q., Liu, M., Zhang, L., Yang, H., Zhang, L., Yang, S., et al. (2022). Head acupuncture plus Schuell's language rehabilitation for post-stroke aphasia: a systematic review and meta-analysis of 32 randomized controlled trials. *Chin. J. Integr. Med.* 28, 743–752. doi: 10.1007/s11655-022-3722-5

Geisler, F. C. M., Kubiak, T., Siewert, K., and Weber, H. (2013). Cardiac vagal tone is associated with social engagement and self-regulation. *Biol. Psychol.* 93, 279–286. doi: 10.1016/j.biopsycho.2013.02.013

Gross, J. J. (2015). Emotion regulation: current status and future prospects. *Psychol. Inq.* 26, 1–26. doi: 10.1080/1047840X.2014.940781

Hodgson, T. (2019). "*The relationship between spatial attention and eye movements*," Switzerland: Springer International Publishing AG, 255–278.

Huppert, D., Grill, E., and Brandt, T. (2013). Down on heights? One in three has visual height intolerance. *J. Neurol.* 260, 597–604. doi: 10.1007/s00415-012-6685-1

Jasubhai, D. S., and Mukundan, P. C. R. (2018). Cognitive Behavioural therapy and emotional freedom technique in reducing anxiety and depression in Indian adults. *Int. J. Emerg. Ment. Health* 20:403. doi: 10.4172/1522-4821.1000403

Kaczkurkin, A. N., and Foa, E. B. (2015). Cognitive-behavioral therapy for anxiety disorders: an update on the empirical evidence. *Dialogues Clin. Neurosci.* 17, 337–346. doi: 10.31887/DCNS.2015.17.3/akaczkurkin

Kappelmann, N., Suesse, M., Steudte-Schmiedgen, S., Kaldewaij, R., Browning, M., Michael, T., et al. (2020). D-cycloserine as adjunct to brief computerised CBT for spider fear: effects on fear, behaviour, and cognitive biases. *J. Behav. Ther. Exp. Psychiatry* 68:101546. doi: 10.1016/j.jbtep.2019.101546

Kok, B. E., Coffey, K. A., Cohn, M. A., Catalino, L. I., Vacharkulksemsuk, T., Algoe, S. B., et al. (2013). How positive emotions build physical health. *Psychol. Sci.* 24, 1123–1132. doi: 10.1177/0956797612470827

Laborde, S., Mosley, E., and Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research – recommendations for experiment planning, data analysis, and data reporting. *Front. Psychol.* 8:213. doi: 10.3389/fpsyg.2017.00213

Lane, J. (2009). The neurochemistry of counterconditioning: acupressure desensitization in psychotherapy. *Energy Psychol.* 1, 31–44.

Larsson, C. E., Cabassut, V., Peretout, P., Marliere, S., Vautrin, E., Piliero, N., et al. (2023). Assessment of the objective effect of virtual reality for preoperative anxiety in interventional cardiology. *Am. J. Cardiol.* 205, 207–213. doi: 10.1016/j.amjcard.2023.07.130

LeBeau, R. T., Glenn, D., Liao, B., Wittchen, H., Beesdo-Baum, K., Ollendick, T., et al. (2010). Specific phobia: a review of DSM-IV specific phobia and preliminary recommendations for DSM-V. *Depress. Anxiety* 27, 148–167. doi: 10.1002/da.20655

Lipponen, J. A., and Tarvainen, M. P. (2019). A robust algorithm for heart rate variability time series artefact correction using novel beat classification. *J. Med. Eng. Technol.* 43, 173–181. doi: 10.1080/03091902.2019.1640306

Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., et al. (2010). Overgeneralization of conditioned fear as a pathogenic marker of panic disorder. *Am. J. Psychiatry* 167, 47–55. doi: 10.1176/appi.ajp.2009.09030410

Lissek, S., Rabin, S. J., McDowell, D. J., Dvir, S., Bradford, D. E., Geraci, M., et al. (2009). Impaired discriminative fear-conditioning resulting from elevated fear responding to learned safety cues among individuals with panic disorder. *Behav. Res. Ther.* 47, 111–118. doi: 10.1016/j.brat.2008.10.017

Liu, P., Li, C., Lu, Y., and Liu, Q. (2020). A case report of acupuncture at Neiguan point (P6) for paroxysmal supraventricular tachycardia. *Complement. Med. Res.* 27, 364–368. doi: 10.1159/000506360

Liu, J., Lin, L., and Wang, D. V. (2021). Representation of fear of heights by basolateral amygdala neurons. *J. Neurosci.* 41, 1080–1091. doi: 10.1523/JNEUROSCI.0483-20.2020

Manteghi, F., Nasehi, M., and Zarrindast, M. (2017). Precondition of right frontal region with anodal tDCS can restore the fear memory impairment induced by ACPA in male mice. *EXCLI J.* 16, 1–13. doi: 10.17179/excli2016-693

Mccraty, R., and Shaffer, F. (2015). Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. *Glob. Adv. Health Med.* 4, 46–61. doi: 10.7453/gahmj.2014.073

McKenzie, S., and Eichenbaum, H. (2011). Consolidation and reconsolidation: two lives of memories? *Neuron* 71, 224–233. doi: 10.1016/j.neuron.2011.06.037

McTeague, L. M., Lang, P. J., Laplante, M., Cuthbert, B. N., Shumen, J. R., and Bradley, M. M. (2010). Aversive imagery in posttraumatic stress disorder: trauma recurrence, comorbidity, and physiological reactivity. *Biol. Psychiatry* 67, 346–356. doi: 10.1016/j.biopsych.2009.08.023

Melillo, P., Bracale, M., and Pecchia, L. (2011). Nonlinear heart rate variability features for real-life stress detection. Case study: students under stress due to university examination. *Biomed. Eng. Online* 10:96. doi: 10.1186/1475-925X-10-96

Menevşe, Ş., and Yayla, A. (2023). Effect of emotional freedom technique applied to patients before laparoscopic cholecystectomy on surgical fear and anxiety: a randomized controlled trial. *J. Perianesth. Nurs.* 39, 93–100. doi: 10.1016/j.jopan.2023.07.006

Minewiser, L. (2017). Six sessions of emotional freedom techniques remediate one Veteran's combat-related post-traumatic stress disorder. *Med. Acupunct.* 29, 249–253. doi: 10.1089/acu.2017.1216

Mittelstädt, V., and Miller, J. (2020). Beyond mean reaction times: combining distributional analyses with processing stage manipulations in the Simon task. *Cogn. Psychol.* 119:101275. doi: 10.1016/j.cogpsych.2020.101275

Nader, K., Schafe, G. E., and Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 406, 722–726. doi: 10.1038/35021052

Napadow, V., Dhond, R., Park, K., Kim, J., Makris, N., Kwong, K. K., et al. (2009). Time-variant fMRI activity in the brainstem and higher structures in response to acupuncture. *Neuroimage* 47, 289–301. doi: 10.1016/j.neuroimage.2009.03.060

Ogura, K., Sugano, M., Takabatake, S., Naitoh, Y., and Nakaoka, K. Data from: VR application for visual field measurement of unilateral spatial neglect patients using eye tracking. In The 7th IEEE International Conference on Healthcare Informatics. Xi'an, China: IEEE. (2019)

Pan, J., and Tompkins, W. J. (1985). A real-time QRS detection algorithm. *I.E.E.E. Trans. Biomed. Eng.* 32, 230–236. doi: 10.1109/TBME.1985.325532

Pennington, Z. T., and Fanselow, M. S. (2018). Indirect targeting of subsuperficial brain structures with transcranial magnetic stimulation reveals a promising way forward in the treatment of fear. *Biol. Psychiatry* 84, 80–81. doi: 10.1016/j.biopsych.2018.05.003

Piskorski, J. (2005). Filtering Poincare plots. *Comp. Meth. Sci. Tech.* 11, 39–48. doi: 10.12921/cmst.2005.11.01.39-48

Pittig, A., Wong, A. H. K., Glück, V. M., and Boschet, J. M. (2020). Avoidance and its bi-directional relationship with conditioned fear: mechanisms, moderators, and clinical implications. *Behav. Res. Ther.* 126:103550. doi: 10.1016/j.brat.2020.103550

Powers, M. B., de Kleine, R. A., and Smits, J. A. J. (2017). Core mechanisms of cognitive behavioral therapy for anxiety and depression: a review. *Psychiatr. Clin. North Am.* 40, 611–623. doi: 10.1016/j.psc.2017.08.010

Quintana, D. S., and Heathers, J. A. J. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. *Front. Psychol.* 5:805. doi: 10.3389/fpsyg.2014.00805

Rachman, S. (2015). The evolution of behaviour therapy and cognitive behaviour therapy. *Behav. Res. Ther.* 64, 1–8. doi: 10.1016/j.brat.2014.10.006

Rahmi, T. (2013). Efektivitas Emotional Freedom Technique dalam Mengatasi Trauma Gempa Ibu Rumah Tangga. *Pedagogi* 12, 107–114. doi: 10.24036/pedagogi.v12i2.2212

Reinecke, A., Waldenmaier, L., Cooper, M. J., and Harmer, C. J. (2013). Changes in automatic threat processing precede and predict clinical changes with exposure-based cognitive-behavior therapy for panic disorder. *Biol. Psychiatry* 73, 1064–1070. doi: 10.1016/j.biopsych.2013.02.005

Sebastian, B., and Nelms, J. (2017). The effectiveness of emotional freedom techniques in the treatment of posttraumatic stress disorder: a Meta-analysis. *Explore* 13, 16–25. doi: 10.1016/j.explore.2016.10.001

Shaffer, F., and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Front. Public Health* 5:258. doi: 10.3389/fpubh.2017.00258

Smith, C. A., Armour, M., Lee, M. S., Wang, L., and Hay, P. J. (2018). Acupuncture for depression. *Cochrane Database Syst. Rev.* 2018:CD004046. doi: 10.1002/14651858.CD004046.pub4

Souto, D., and Kerzel, D. (2021). Visual selective attention and the control of tracking eye movements: a critical review. *J. Neurophysiol.* 125, 1552–1576. doi: 10.1152/jn.00145.2019

Stapleton, P., Kip, K., Church, D., Toussaint, L., Footman, J., Ballantyne, P., et al. (2023). Emotional freedom techniques for treating post traumatic stress disorder: an updated systematic review and meta-analysis. *Front. Psychol.* 14:1195286. doi: 10.3389/fpsyg.2023.1195286

Szmit, M., Krajewski, R., Rudnicki, J., and Agrawal, S. (2023). Application and efficacy of transcutaneous electrical acupoint stimulation (TEAS) in clinical practice: a systematic review. *Adv. Clin. Exp. Med.* 32, 1063–1074. doi: 10.17219/acem/159703

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., and Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neurosci. Biobehav. Rev.* 36, 747–756. doi: 10.1016/j.neubiorev.2011.11.009

Thayer, J. F., Hansen, A. L., Saus-Rose, E., and Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: the Neurovisceral integration perspective on self-regulation, adaptation, and health. *Ann. Behav. Med.* 37, 141–153. doi: 10.1007/s12160-009-9101-z

Thayer, J. F., and Lane, R. D. (2009). Claude Bernard and the heart–brain connection: further elaboration of a model of neurovisceral integration. *Neurosci. Biobehav. Rev.* 33, 81–88. doi: 10.1016/j.neubiorev.2008.08.004

van der Veen, S., Bordeleau, M., Pidcoe, P., France, C., and Thomas, J. (2019). Agreement analysis between Vive and Vicon systems to monitor lumbar postural changes. *Sensors* 19:3632. doi: 10.3390/s19173632

Wang, H., Wright, B., Tindall, L., Cooper, C., Biggs, K., Lee, E., et al. (2022). Cost and effectiveness of one session treatment (OST) for children and young people with specific phobias compared to multi-session cognitive behavioural therapy (CBT): results from a randomised controlled trial. *BMC Psychiatry* 22:547. doi: 10.1186/s12888-022-04192-8

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54, 1063–1070. doi: 10.1037//0022-3514.54.6.1063

Wittfoth, D., Beise, J., Manuel, J., Bohne, M., and Wittfoth, M. (2022). Bifocal emotion regulation through acupoint tapping in fear of flying. *Neuroimage Clin.* 34:102996. doi: 10.1016/j.nicl.2022.102996

Wittfoth, D., Pfeiffer, A., Bohne, M., Lanfermann, H., and Wittfoth, M. (2020). Emotion regulation through bifocal processing of fear inducing and disgust inducing stimuli. *BMC Neurosci.* 21:47. doi: 10.1186/s12868-020-00597-x

Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., and Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: a meta-analysis. *Clin. Psychol. Rev.* 28, 1021–1037. doi: 10.1016/j.cpr.2008.02.007

Wout, M., Longo, S. M., Reddy, M. K., Philip, N. S., Bowker, M. T., and Greenberg, B. D. (2017). Transcranial direct current stimulation may modulate extinction memory in posttraumatic stress disorder. *Brain Behav.* 7:e00681. doi: 10.1002/brb3.681

Wroblewski, D., Francis, B. A., Sadun, A., Vakili, G., and Chopra, V. (2014). Testing of visual field with virtual reality goggles in manual and visual grasp modes. *Biomed. Res. Int.* 2014, 1–10. doi: 10.1155/2014/206082

Yang, J., Tang, M., Cong, L., Sun, J., Guo, D., Zhang, T., et al. (2023). Development and validation of an assessment index for quantifying cognitive task load in pilots under simulated flight conditions using heart rate variability and principal component analysis. *Ergonomics* 6, 1–11. doi: 10.1080/00140139.2023.2229075

Ye, Z., Zhu, L., Li, X., Gao, H., Wang, J., Wu, S., et al. (2023). PC6 electroacupuncture reduces stress-induced autonomic and neuroendocrine responses in rats. *Heliyon* 9:e15291. doi: 10.1016/j.heliyon.2023.e15291

Zaccaro, A., Piarulli, A., Laurino, M., Garbella, E., Menicucci, D., Neri, B., et al. (2018). How breath-control can change your life: a systematic review on psycho-physiological correlates of slow breathing. *Front. Hum. Neurosci.* 12:353. doi: 10.3389/fnhum.2018.00353

Zhou, W., Deng, Q., Jia, L., Zhao, H., Yang, M., Dou, G., et al. (2020). Acute effect of transcutaneous Electroacupuncture on Globus Pharyngeus: a randomized, single-blind, crossover trial. *Front. Med.* 7:179. doi: 10.3389/fmed.2020.00179

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 04 April 2024
DOI 10.3389/fpsyg.2024.1275142

OPEN ACCESS

EDITED BY
Federica Marcolin,
Polytechnic University of Turin, Italy

REVIEWED BY
Hiroki Tanaka,
Nara Institute of Science and Technology (NAIST), Japan
Jiahui Pan,
South China Normal University, China

*CORRESPONDENCE
Yanan Zhou
✉ zhouyanan@bfsu.edu.cn
Hebin Cheng
✉ chenghebin@sdmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 09 August 2023
ACCEPTED 20 March 2024
PUBLISHED 04 April 2024

CITATION
Wang D, Lian J, Cheng H and Zhou Y (2024) Music-evoked emotions classification using vision transformer in EEG signals. *Front. Psychol.* 15:1275142. doi: 10.3389/fpsyg.2024.1275142

Music-evoked emotions classification using vision transformer in EEG signals

Dong Wang[1,2†], Jian Lian[2†], Hebin Cheng[2*] and Yanan Zhou[3*]

[1]School of Information Science and Electrical Engineering, Shandong Jiaotong University, Jinan, China, [2]School of Intelligence Engineering, Shandong Management University, Jinan, China, [3]School of Arts, Beijing Foreign Studies University, Beijing, China

Introduction: The field of electroencephalogram (EEG)-based emotion identification has received significant attention and has been widely utilized in both human-computer interaction and therapeutic settings. The process of manually analyzing electroencephalogram signals is characterized by a significant investment of time and work. While machine learning methods have shown promising results in classifying emotions based on EEG data, the task of extracting distinct characteristics from these signals still poses a considerable difficulty.

Methods: In this study, we provide a unique deep learning model that incorporates an attention mechanism to effectively extract spatial and temporal information from emotion EEG recordings. The purpose of this model is to address the existing gap in the field. The implementation of emotion EEG classification involves the utilization of a global average pooling layer and a fully linked layer, which are employed to leverage the discernible characteristics. In order to assess the effectiveness of the suggested methodology, we initially gathered a dataset of EEG recordings related to music-induced emotions.

Experiments: Subsequently, we ran comparative tests between the state-of-the-art algorithms and the method given in this study, utilizing this proprietary dataset. Furthermore, a publicly accessible dataset was included in the subsequent comparative trials.

Discussion: The experimental findings provide evidence that the suggested methodology outperforms existing approaches in the categorization of emotion EEG signals, both in binary (positive and negative) and ternary (positive, negative, and neutral) scenarios.

KEYWORDS
music-evoked emotion, emotion classification, electroencephalographic, deep learning, transformer

1 Introduction

Emotion is intricately intertwined with all facets of the human experience and action. According to Jerritta et al. (2011), it has an impact on human attitudes and perceptions in both human-human contact and human-computer interaction. In the realm of artistic expression, music holds a paramount position as a means to convey and articulate human emotions. Music has been widely recognized as a means of evoking distinct emotive states, leading to its characterization as the language of emotions (Vuilleumier and Trost, 2015). In their investigations, Ekman (1999) and Gilda et al. (2017) introduced six distinct and quantifiable emotional states, namely happiness, sadness, anger, fear, surprise, and disgust, as the basis for implementing emotion identification.

10.3389/fpsyg.2024.1275142

Over time, other emotional states have been included in this collection, such as neutrality, arousal, and relaxation (Bong et al., 2012; Selvaraj et al., 2013; Goshvarpour et al., 2017; Minhad et al., 2017; Wei et al., 2018; Sheykhivand et al., 2020; Liu et al., 2022). In the context of machine learning, the establishment of distinct states for emotions serves as a significant framework for effectively addressing the challenge of emotion recognition. Numerous algorithms for music emotion identification based on machine learning have been proposed in the literature, with applications spanning composition and psychotherapy (Eerola and Vuoskoski, 2012; Cui et al., 2022).

Typically, a conventional music emotion identification system based on machine learning encompasses the subsequent stages:

- The collection of changes in emotions elicited by music is facilitated via the utilization of physiological information obtained by specialized sensors.
- The physiological samples that have been gathered are subjected to a processing procedure in order to remove any potential artifacts.
- The generation of representation pertaining to emotional states is thereafter accomplished by extracting features from the pre-processed data.
- By utilizing a classifier, it is possible to generate the corresponding category of music emotion for a given sample.

Numerous instruments utilized in the acquisition of physiological signals have been employed for the purpose of emotion recognition. Various physiological signals have been investigated for the purpose of emotion recognition. These include, body movement (Zhang et al., 2021), facial expression (Song, 2021), respiration (Siddiqui et al., 2021), galvanic skin response (Kipli et al., 2022), blood volume pulse (Semerci et al., 2022), skin temperature (Semerci et al., 2022), electromyography (Xu et al., 2023), photoplethysmographic (Cosoli et al., 2021), electrocardiogram (Hasnul et al., 2021), and EEG (Li et al., 2021). The non-invasive nature, affordability, and ability to capture data in real-time have contributed to the extensive utilization of EEG in the field of emotion identification (Alarcao and Fonseca, 2017), with a particular emphasis on music emotion categorization (Lin et al., 2006).

Several studies have introduced different approaches for emotion categorization utilizing EEG in the context of machine learning. For example, the study conducted by Sammler et al. (2007) examined the impact of valence on human emotions by analyzing EEG data and heart rate concurrently. The present study aimed to gather data on positive and negative emotions elicited by EEG signals during the auditory experience of consonant and discordant musical stimuli. Subsequently, the authors of the study (Koelstra et al., 2011) made available a publicly accessible dataset. The study conducted by Balasubramanian et al. (2018) examined the emotional reaction to various types of music using EEG data. The experimental findings have indicated that there is an increase in theta band activity in the frontal midline region when individuals are exposed to their preferred music. Conversely, the beta band would have an increase in activity when exposed to music that is perceived as undesirable. In their study, Ozel et al. (2019) introduced a methodology for emotion identification that involves the analysis of temporal-spectral EEG signals. Hou and Chen (2019) derived a set of 27-dimensional EEG characteristics to represent music-induced emotions, including calmness, pleasure, sadness, and rage. *Recently, Qiu et al. (2022) proposed an integrated framework of multi-modal EEG and functional near infrared spectroscopy to explore the influence of music on brain activity.*

In addition, the utilization of deep learning-based architectures in music emotion categorization has been widely adopted due to the shown effectiveness of deep learning in different domains such as machine vision and natural language processing. In their study, Han et al. (2022) conducted a comprehensive review of the existing literature pertaining to the assessment metrics, algorithms, datasets, and extracted features utilized in the analysis of EEG signals in the context of music emotion detection. In their publication, Nag et al. (2022) introduced the JUMusEmoDB dataset. The music emotion categorization challenge was addressed by the authors through the utilization of Convolutional Neural Network (CNN) based models, namely resnet50, mobilenet, squeezenet, and their own suggested ODE-Net. Eskine (2022) conducted a study examining the impact of music listening on creative cognition, a phenomenon that has been empirically demonstrated to enhance creative cognitive processes. The experimental findings provided evidence that cognitive function exhibited an increase inside the default mode. This was supported by the observed augmentation of spectral frequency power in the beta range throughout the entire brain, as well as in the theta range within the parietal region, and in the gamma range across the entire brain. In their study, Daly (2023) investigated the integration of functional magnetic resonance imaging (fMRI) and EEG techniques to develop an acoustic decoder for the purpose of classifying music emotions. The study employed an EEG-fMRI combined paradigm to capture neural responses during music listening among individuals. In this study, a deep learning model known as the long short-term memory (LSTM) was utilized to extract neural information from EEG signals during music listening. The objective was to rebuild the matching music clips based on this extracted information.

Both machine learning and deep learning techniques have demonstrated promising results in the categorization of music-evoked emotions. Nevertheless, there are a number of constraints associated with these approaches that must be addressed prior to their practical implementation in contexts such as medical diagnosis, namely in the realm of emotion identification. One aspect to consider is that the efficacy of machine learning techniques is heavily dependent on the selection of appropriate features. The task at hand continues to provide an unsolved problem as the extraction and selection of these characteristics from EEG data must be done in a manual manner. In addition, it should be noted that manually-designed features possess subjectivity and susceptibility to errors, perhaps rendering them unsuitable for the specific requirements of music emotion identification. In contrast, deep learning models like as CNNs have the ability to automatically

Abbreviations: EEG, Electroencephalographic; fMRI, functional magnetic resonance imaging; LSTM, long short term; CNN, convolutional neural network; ECG, electrocardiogram; EOG, electro-oculugram; GAP, global average pooling; FC, fully connected; GPU, graphical processing unit; TP, true positive; FN, false negative; TN, true negative; FP, false positive.

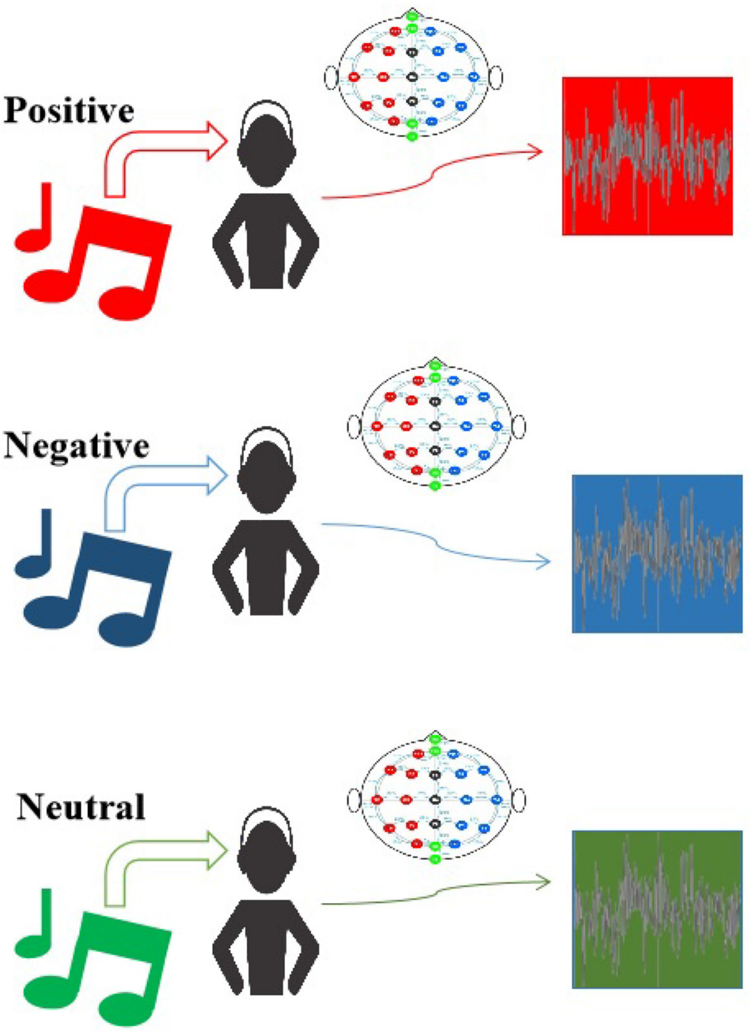


FIGURE 1
The data collecting process for classifying music-evoked emotions using an EEG equipment based on the 10–20 system (Homan et al., 1987).

extract internal representations from EEG inputs. Nevertheless, it is expected that the features derived from CNN models prioritize the consideration of the overall connection between distant EEG signals. This is due to the fact that CNN utilizes a local receptive field approach in the process of extracting features.

The present work introduces a transformer architecture for music-evoked emotion categorization, using a self-attention mechanism. This model incorporates the self-attention mechanism and positional embedding to describe the sequence of channels in EEG data, drawing inspiration from the vision transformer's work (Dosovitskiy et al., 2020). The suggested transformer model has the ability to extract both spatial representations, which correspond to self-attention modules, and temporal representations, which correspond to positional embedding. These representations are derived from multi-channel EEG data acquired from subjects who were listening to music. Furthermore, the transformer model that has been introduced has the capability to extract the relationships that exist among EEG signals across extended distances. In order to assess the efficacy of the suggested methodology, the experiments were conducted using both a publicly accessible dataset (Koelstra et al., 2011) and a privately held dataset. Furthermore, comparative tests were conducted to evaluate the performance of the proposed model in comparison to state-of-the-art algorithms. The experimental findings provide evidence that the suggested methodology exhibits superior performance compared to existing binary and ternary music emotion categorization algorithms. The suggested model has a positive conclusion, indicating its potential value as a tool for classifying music-evoked emotions.

The main contributions of this work can be summarized as follows:

- *This is an early application of the spatial-temporal transformer into the classification of music-evoked emotions.*
- *A novel dataset of music-evoked EEG signals was established.*
- *The proposed approach considers both the spatial connections among a set of EEG channels and the temporal sequence of each individual EEG signal.*

TABLE 1 The descriptions of the music excerpts used in this study.

ID	Type	Title	Singer	Durating (mm:ss)
1	Positive	Honey	Xinling Wang	03:33
2	Negative	Advanced animals	Wei Dou	04:38
3	Neutral	Reiki meditation	Reiki	06:03
4	Positive	Wu Ha	Weibo Pan	03:46
5	Negative	In case	SHIN	04:24
6	Neutral	Calm dreams	Sleep Tech	04:27
7	Positive	In Spring	Feng Wang	05:10
8	Negative	Cloudy day	Wenwei Mo	04:02
9	Neutral	Let the sun shine	Milk & Sugar	07:02
10	Positive	As broad as the sea and sky	Beyond	03:59
11	Negative	Negative	Black sun empire	05:44
12	Neutral	Illusionary daytime	Shirfine	04:10
13	Positive	Invisible wings	Shaohan Zhang	03:44
14	Negative	Unfortunately, its not you	Jingru Liang	04:45
15	Neutral	Song from a secret garden	Secret garden	03:33

- *The performance of our approach surpassed the state-of-the-art deep learning algorithms on both public and private datasets.*

The subsequent sections of this article are structured as follows: The methodology Section 2 contains information on the acquisition of EEG signals during music listening as well as the details of the presented deep learning model. Section 3 presents a detailed account of the experimental procedures conducted in this investigation, as well as a comprehensive comparison between the existing state-of-the-art methods and the technique proposed in the current study. This research concluded at Section 4.

2 Methodology

This section provides a comprehensive overview of the data gathering process employed in the present investigation. Furthermore, the subsequent sections of the article will present a comprehensive analysis of the suggested transformer model.

2.1 Dataset and pre-processing

The initial step in this study included the creation of a private dataset using multi-channel EEG, which involved the collection of three distinct music-evoked emotions: positive, negative, and neutral. The complete workflow is depicted in Figure 1.

During the course of data gathering, a total of 48 individuals were registered, including 24 females and 24 men. The age range of the participants was between 18 and 25 years, with an average age of 20.6. All individuals involved in the study were enrolled as students at the same institution's campus. Furthermore, it should be noted that the individuals exhibit robust physical and mental well-being. During the course of the project, the research team received advice and supervision from two psychology specialists, one female and one male, who possessed significant expertise in the field.

To ensure the consistency of the data gathering process, the following challenges were proactively addressed. Additionally, all participants were provided with instructions to thoroughly review the handbook and become acquainted with the workflow of EEG signal collecting. It should be noted that the manual has identified and emphasized the entries that are prone to errors, with the intention of facilitating the reader's attention toward the vital operations. Subsequently, the participants were requested to complete a questionnaire pertaining to their personal details. Subsequently, the participants were provided with instructions and guidance from the specialists in order to properly don the EEG electrode caps. Subsequently, the specialists would assess the adequacy of the EEG electrodes' contact and ensure that no detachment has occurred. Furthermore, the participants were instructed by the experts to initiate the signal gathering procedure by hitting the designated buttons. In addition, the EEG collection device utilized in the study was the Biosemi ActiveTwo system. The system employs the international 10–20 system, consisting of 32 channels, notably Fp1, AF3, F3, F7, FC5, FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC, Cz, C4, T8, Cp6, Cp2, P4, P8, PO4, and O2. Additionally, the sampling rate is set at 512Hz.

During the process of data collection, each participant was provided with instructions to listen to a total of 15 music clips. These clips were categorized into three distinct emotional categories, namely positive, negative, and neutral, with each category consisting of five clips. To note that the categories of these clips were determined by three psychological experts using a majority voting mechanism. The specifics about the music may be found in Table 1. The initial duration of the music clips varies among them. Nevertheless, the participant received a standardized 1-min audio clip for each piece of music. Each participant

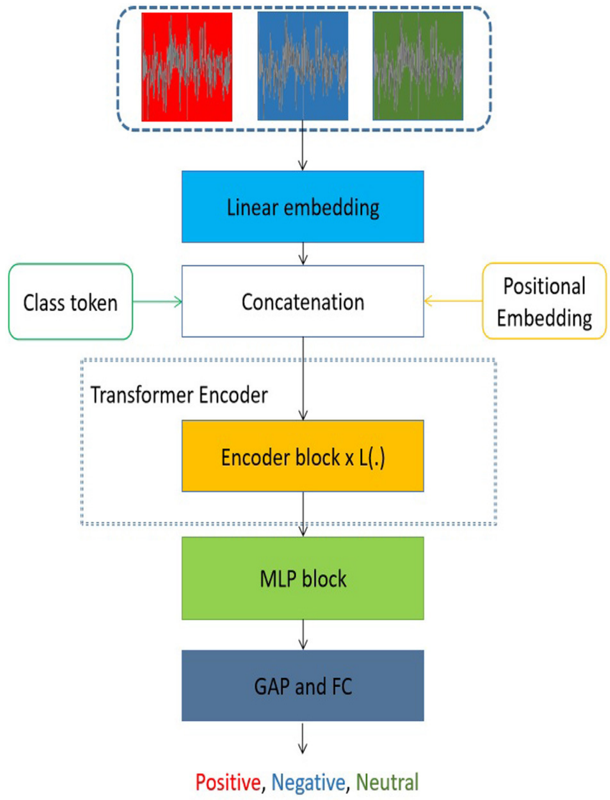


FIGURE 2
The architectural design of the proposed transformer model. The symbol L is used to describe the number of encoder blocks in the proposed model. The proposed model has several variants, with L(.) being either 12, 18, or 24. The abbreviation GAP is used to denote global average pooling, while FC represents fully-connected.

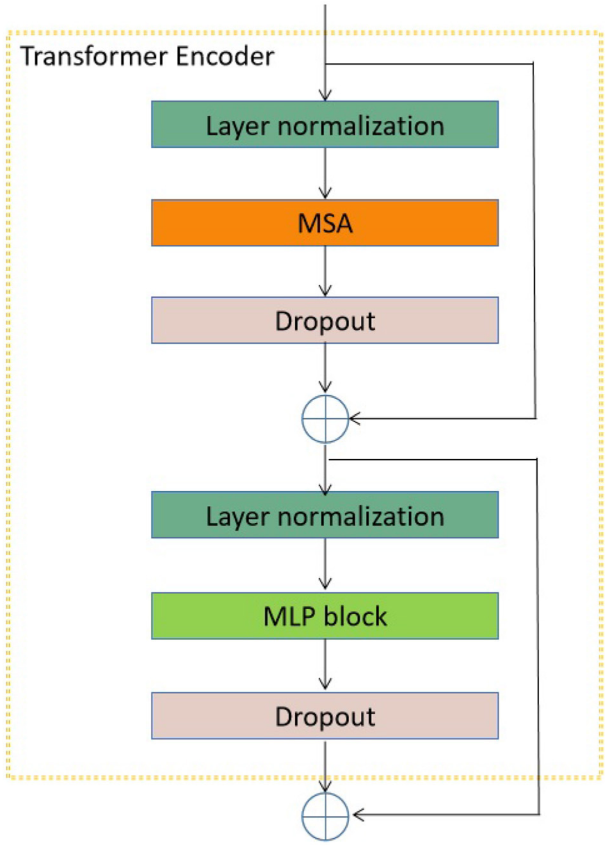


FIGURE 3
The encoder block utilized in the transformer model under consideration. The acronym MSA refers to multi-head self attention.

was instructed to listen to the music clips in a randomized sequence.

The subsequent section presents a comprehensive overview of the data collecting procedure involved in capturing EEG signals related to music-induced emotions.

(1) The participants were provided with instructions to achieve a state of calmness, following which the experts started the marking process to denote the commencement of each EEG recording. The duration of this process is expected to be 5 seconds.
(2) In each 75-second interval, the participants would undergo a 15-second pause to transition between music clips, followed by a 60-second period of actively listening to the music clip. Simultaneously, the experts would provide guidance to the participants on how to minimize superfluous bodily motions.
(3) Following the auditory experience, the individuals were directed by the experimental personnel to assign a value to the musical composition, with positive being denoted as +1, negative as –1, and neutral as 0. The duration of this procedure should not exceed 15 seconds, during which it is utilized for the purpose of transitioning the music.
(4) The participants proceeded with the auditory experience by sequentially engaging with the subsequent musical excerpt until the entirety of the 12 excerpts had been presented.

So as to guarantee the optimal state of the participants, the collection of music-evoked emotion EEG samples was limited to the time periods of 9 a.m. to 11 a.m. and 3 p.m. to 5 p.m. In order to mitigate interference from many sources such as heart rate, breathing, electrocardiogram (ECG), and electro-scalogram (EOG), the participants were given instructions to cover their eyes while the recording procedures were being conducted.

The dataset contains a total of 43,200 ($48 \times 15 \times 60 = 43,200$) seconds of EEG signals, with each second including 32 channels. Furthermore, the initial samples were partitioned into the epochs of 1 second duration, each consisting of $60,000$ data points. To note that there were still overlapping epochs in the samples since the trivial errors are difficult to avoid due to the human reaction times. Given the absence of any imbalance issue within the dataset, it can be observed that each category of music emotion EEG signals is comprised of an equal number of samples, specifically 20,000 epochs. Hence, in the context of binary classification, namely distinguishing between positive and negative classes, the proposed model was trained using a dataset including 40,000 epochs as input samples. In contrast, in the context of the ternary classification job, the entirety of the 60,000 epochs were utilized as the input. It should

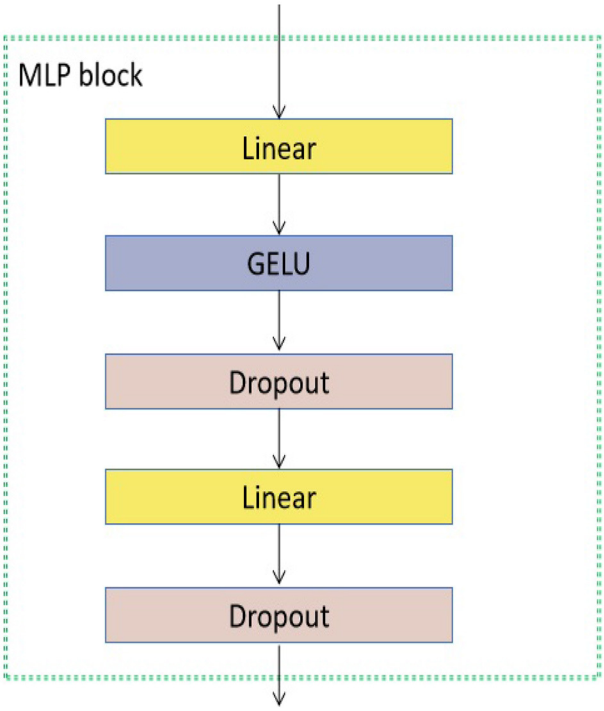


FIGURE 4
The MLP block used in the proposed transformer model. GELU denotes the activation function (Lee, 2023).

be noted that the presence of overlapping epochs has the potential to somewhat mitigate over-fitting.

In the pre-processing phase, the acquired EEG signals were subjected to a Notch filter (Serra et al., 2017) in order to remove the 50 Hz components originating from the power supply. Subsequently, a first-order low-pass filter with a frequency range of 0.5 to 45 Hz was utilized. Subsequently, the electroencephalography (EEG) data underwent a normalization process resulting in a range of values between 0 and 1.

2.2 The proposed transformer architecture

The transformer model presented in Figure 2 draws inspiration from the architecture of the vision transformer (Dosovitskiy et al., 2020). The suggested transformer model comprises three main components: (1) a linear embedding layer, (2) an encoder block, and (3) a multiple-layer perception (MLP) block. Initially, the linear embedding unit was utilized to turn a sequence of EEG data into a fixed-length input for the suggested transformer model. The flattened embedding includes the class token of the music emotion for each series of EEG data. In addition, the linear embedding is constructed by including the positional embedding, which encodes the sequential order of an individual EEG signal inside a sequence of EEG signals. It should be noted that every input sequence of EEG data pertains to the identical category of emotion elicited by music. Furthermore, the pivotal self-attention module (Fan et al., 2021; Liu et al., 2021; Wang et al., 2021), which aims to reveal the connections among distant EEG data, is located within the encoder block. In order to create a cohesive encoder module, it is necessary for the encoder block to be iteratively repeated. In addition to the self-attention layer included in each encoder block, there are many additional sorts of layers, namely layer normalization, dropout, and MLP block. The generation of representations for music emotion EEG signals may be achieved by the utilization of stacked transformer encoder blocks. Ultimately, the use of the MLP block was implemented to get the classification result by integrating a global average pooling (GAP) layer and a fully connected (FC) layer, commonly referred to as a linear layer. The transformer model under consideration has the potential to significantly expand the scope of receptive fields in comparison to designs based on CNNs. Additionally, the recovered representation from the multi-channel EEG data encompasses both local information pertaining to a series of signals and the global association between signals that are far apart.

TABLE 2 The proposed transformer model exhibits binary and ternary classification outcomes (average values and standard deviations).

Number of classes	Accuracy (%)	Sensitivity (%)	Specificity (%)
Binary	96.85 (1.73)	95.17 (1.68)	95.69 (2.01)
Ternary	95.74 (2.32)	94.32 (1.97)	95.25 (1.69)

In the proposed transformer model, the input sequences consist of individual EEG signals, each spanning a duration of 1 second and including 30 channels. Subsequently, the EEG signal sequence was flattened and transformed into a vector. In addition, it should be noted that the encoder block is iterated a varying number of times (12, 18, or 24) across different versions of the proposed transformer model. Furthermore, the structural composition of this encoder block is illustrated in Figure 3.

As seen in Figure 3, the encoder block has many components, namely layer normalization, MSA, dropout, and MLP block. The study did not include a comprehensive examination of the MSA unit due to its extensive coverage in existing studies (Vaswani et al., 2017; Dosovitskiy et al., 2020). The unit consisting of H heads was employed to assess the similarity between a query and its associated keys based on the assigned weight for each value (Vaswani et al., 2017). Furthermore, the utilization of the Layer normalizing module is employed to calculate the mean and variance required for normalizing from the entirety of the inputs to the neurons within a layer throughout a singular training instance (Ba et al., 2016). In this study, the dropout layer (Choe and Shim, 2019) is utilized as a regularization technique to mitigate the risk of overfitting. The architecture of the MLP block is seen in Figure 4.

The proposed technique allows for the formulation of the process of music emotion categorization in Equation 1–4:

$$z_0 = [x_{class}; x_p^1 E; x_p^2 E; \ldots; x_p^N] + E_{position}, \quad (1)$$

where the variable z_0 represents the output of the linear embedding layer. In this context, $N = 30$ represents the number of channels used as input. The variables x_{class} and $E_{position}$ refer to the class token and positional embedding, respectively.

$$z_l' = MSA(LN(z_{l-1})) + z_{l-1}, \tag{2}$$

$$z_l = MLP(LN(z_l')) + z_l', \tag{3}$$

$$y = LN(z_L^0), \tag{4}$$

where the layer normalization unit is denoted as $LN(.)$, where z_l represents the output of layer l, and y represents the output classification outcome.

3 Experimental results

3.1 Implementation details

The transformer model described in this study was constructed using the PyTorch framework (Paszke et al., 2019). The computational resources employed for the implementation were four NVidia RTX 3080 Graphical Processing Units (GPUs) with a total of 64 GB RAM. The best parameters of the proposed network were discovered using a trial and error technique. The learning rate is configured to be 0.004, accompanied by a weight decay of 0.05. Subsequently, a 10-fold cross-validation procedure was employed to assess the resilience of the suggested methodology. Initially, the input EEG data were partitioned into ten equitably sized groups. During each iteration, one out of the 10 groups was designated as the testing set, while the remaining nine groups were utilized as the

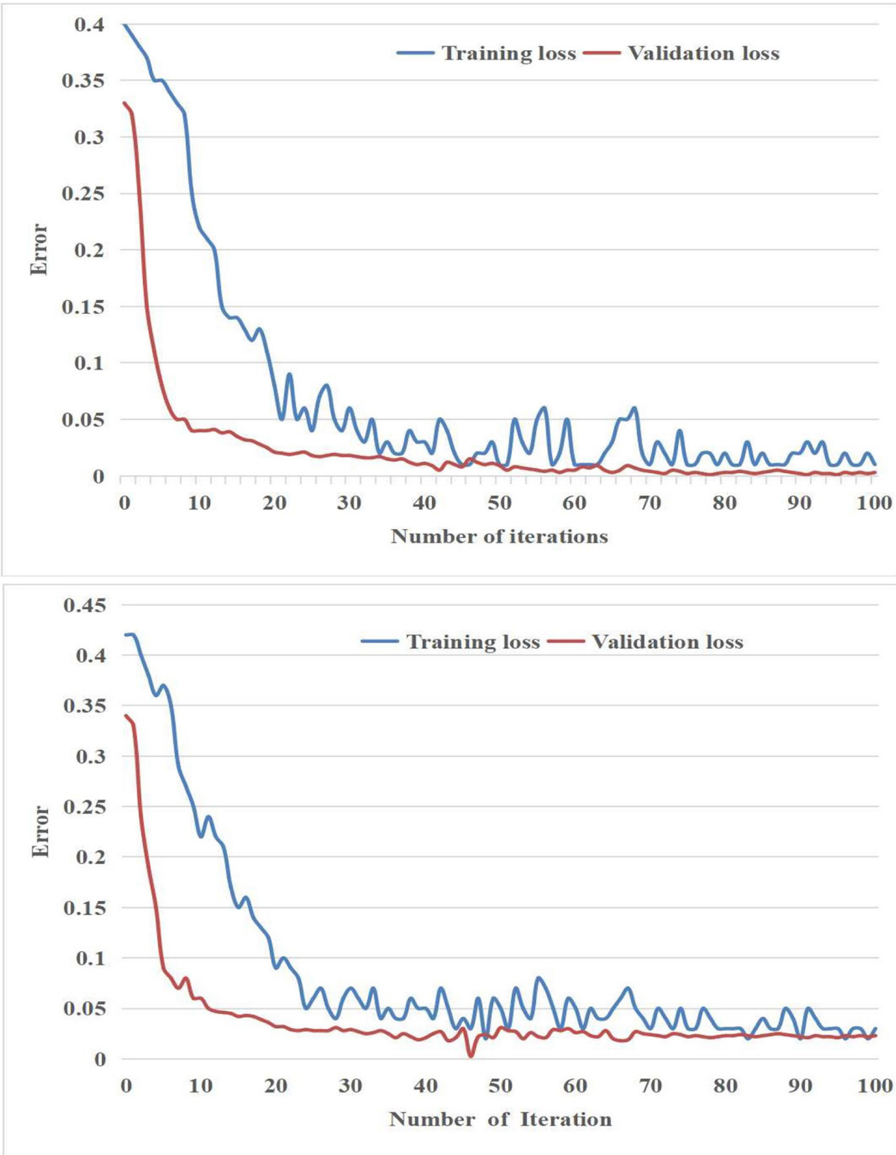


FIGURE 5
The suggested model's inaccuracy in **(Top)** binary classification and **(Bottom)** ternary classification.

TABLE 3 Binary classification comparison between the state-of-the-arts and ours.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
U-Net (Ronneberger et al., 2015)	88.56	88.71	89.05
Mask R-CNN (He et al., 2017)	87.43	86.39	86.56
ExtremeNet (Zhou et al., 2019)	89.49	89.87	88.51
TensorMask (Chen et al., 2019)	90.56	90.18	91.27
4D-CRNN (Shen et al., 2020)	92.57	92.32	93.08
FBCCNN (Pan and Zheng, 2021)	92.53	91.68	91.24
MTCNN (Rudakov, 2021)	93.02	93.55	94.17
SSGMC (Kan et al., 2022)	94.82	94.18	94.23
MViT (Fan et al., 2021)	90.42	91.39	90.72
PVT (Wang et al., 2021)	92.27	91.15	92.01
PiT (Heo et al., 2021)	93.53	92.85	93.78
Swin Transformer (Liu et al., 2021)	95.32	94.64	94.37
GPViT (Yang et al., 2022)	96.38	94.88	95.27
The proposed approach	96.85	95.17	95.69

TABLE 4 Ternary classification comparison between the state-of-the-arts and ours.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
U-Net (Ronneberger et al., 2015)	85.52	83.86	84.20
Mask R-CNN (He et al., 2017)	85.24	84.21	85.41
ExtremeNet (Zhou et al., 2019)	86.28	83.17	84.53
TensorMask (Chen et al., 2019)	88.32	86.51	87.02
4D-CRNN (Shen et al., 2020)	91.57	92.24	91.89
FBCCNN (Pan and Zheng, 2021)	91.27	91.38	92.24
MTCNN (Rudakov, 2021)	92.21	92.19	93.43
SSGMC (Kan et al., 2022)	92.18	91.57	94.28
MViT (Fan et al., 2021)	92.15	91.93	92.78
PVT (Wang et al., 2021)	91.23	90.46	91.37
PiT (Heo et al., 2021)	92.43	92.14	91.62
Swin transformer (Liu et al., 2021)	92.57	91.38	93.27
GPViT (Yang et al., 2022)	93.14	92.25	93.18
The proposed approach	95.74	94.32	95.25

training set. Hence, the mean result of 10 iterations was utilized as the ultimate output.

Furthermore, the assessment measures utilized in the experiments involved sensitivity, specificity, and accuracy. The mathematical formulation of these metrics is elucidated in in Equations 5–7.

$$Sensitivity = \frac{TP}{TP + FN}, \tag{5}$$

$$Specificity = \frac{TN}{TN + FP}, \tag{6}$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}, \tag{7}$$

where TP, FN, TN, and FP represent the terms true positive, false negative, true negative, and false positive, respectively.

3.2 Outcome of the proposed approach

Table 2 presents a summary of the average values and standard deviations (SD) obtained from the proposed method in the binary classification task, specifically in terms of average accuracy, sensitivity, and specificity. The average accuracy was found to be 96.85%, while the sensitivity and specificity were measured at 95.17% and 95.69% respectively. Furthermore, in the ternary categorization, the outcome rates were recorded as 95.74%, 94.32%, and 95.25%.

Furthermore, the loss curves of the suggested methodology throughout both the training and validation procedures were illustrated in Figure 5 It should be noted that the results presented in Figure 5 only include the initial 100 iterations of both the training and validation processes.

3.3 Comparison experiments between the state-of-the-arts and the proposed approach

To assess the efficacy of our suggested technique for music-evoked emotion categorization, we conducted comparative tests between our work and the state-of-the-art algorithms. Tables 2–4 present a comparative analysis of the current state-of-the-art deep learning models and our proposed approach. The proposed

TABLE 5 Comparison between the state-of-the-arts and ours on DEAP dataset (Koelstra et al., 2011).

Method	Detail	Accuracy	
		Valence	Arousal
3DCNN (Shawky et al., 2018)	CNN	88.52	89.36
CNN-LSTM (Yang et al., 2018)	LSTM	92.43	89.51
SAE-LSTM (Xing et al., 2019)	LSTM	86.32	81.27
Multi-column CNN (Yang et al., 2019)	CNN	93.81	94.15
4D-CRNN (Shen et al., 2020)	CRNN	95.34	93.62
FGCCNN (Pan and Zheng, 2021)	CNN	91.72	90.28
MTCNN (Rudakov, 2021)	CNN	95.34	95.49
GANSER (Zhang et al., 2022)	GAN	94.18	93.58
SSGMC (Kan et al., 2022)	Contrastive learning	96.12	94.62
The proposed approach	Transformer	97.41	97.02

methodology demonstrated superior performance compared to the current leading method. *To note that we did not take the traditional machine learning models (Qiu et al., 2022) into the comparison since they usually relied on manually-designed features.* The comparison experiments included the following models: U-Net (Ronneberger et al., 2015), Mask R-CNN (He et al., 2017), ExtremeNet (Zhou et al., 2019), TensorMask (Chen et al., 2019), 4D-CRNN (Shen et al., 2020), FBCCNN (Pan and Zheng, 2021), MTCNN (Rudakov, 2021), SSGMC (Kan et al., 2022) for the CNN-based models, and MViT (Fan et al., 2021), PVT (Wang et al., 2021), PiT (Heo et al., 2021), Swin Transformer (Liu et al., 2021), and GPViT (Yang et al., 2022) for the transformer-based models.

In order to conduct a comprehensive evaluation of the proposed approach, we proceeded to assess its performance alongside several state-of-the-art algorithms (Shawky et al., 2018; Yang et al., 2018; Xing et al., 2019; Yang et al., 2019; Shen et al., 2020; Pan and Zheng, 2021; Rudakov, 2021; Kan et al., 2022; Zhang et al., 2022) using the publicly accessible DEAP dataset (Koelstra et al., 2011). The results of this evaluation are presented in Table 5.

4 Discussion

Based on the empirical findings, it can be concluded that this approach exhibits greater efficacy compared to the existing state-of-the-art algorithms. It is worth mentioning that the comparative trials encompassed both CNN-based and transformer-based models. In contrast to CNN-based models, the suggested model has the capability to extract global connections between long-range multi-channels in EEG data, in addition to the local

TABLE 6 The impact of *H* and *L* on the performance of the proposed model in binary classification.

Model	Number of heads (H)	Number of layers (L)	Accuracy (%)
M_4_4	4	4	90.08
M_4_8	4	8	90.37
M_8_4	8	4	91.15
M_8_8	8	8	91.63
M_8_12	8	12	93.35
M_12_12	12	12	93.21
M_16_12	16	12	94.16
M_8_18	8	18	94.58
M_12_18	12	18	95.39
M_16_18	16	18	95.65
M_8_24	8	24	96.28
M_12_24	12	24	96.12
M_16_24	16	24	**96.53**

The bold value represents the best performance of accuracy with 16 heads and 24 layers.

information already present in the EEG signals. In contrast to transformer-based models (He et al., 2017; Chen et al., 2019; Zhou et al., 2019; Wu et al., 2020; Fan et al., 2021; Heo et al., 2021; Wang et al., 2021), the proposed approach has been specifically optimized to accommodate the unique characteristics of multi-channel EEG signals. For instance, the linear embedding layer of the proposed approach has been tailored to effectively align with the structural properties of multi-channel EEG signals. Furthermore, the outcomes shown in the ablation research also exhibited the efficacy of self-attention modules and encoder blocks.

4.1 Ablation study

As demonstrated in Table 6, the optimal configuration of the primary hyper-parameters was determined through comparison experiments. These experiments involved testing different combinations of the number of heads (H) in the MSA module and the number of transformer encoder layers (L) on a dataset that was manually collected and constituted 50% of the total dataset. The trials solely included binary music emotion categorization in order to streamline the ablation study procedure.

Therefore, the suggested model exhibits an ideal configuration while utilizing 16 heads ($H = 16$) and 24 layers ($L = 24$).

4.2 Limitations and future research

In addition, this study possesses certain limitations in addition to its contributions. The tests solely focused on the binary and ternary classification problems. In order to enhance the evaluation of the proposed approach, it is recommended to integrate the categorization of other types of emotions and employ a multi-label classification methodology. Meanwhile, this study adopted an offline learning strategy since the vision transformer-based models suffering from high resource occupancy. In addition, this study did not take cross-subject emotion recognition (He et al., 2021; Pan et al., 2023) into consideration, which may affect the applicability and universality of this study.

In subsequent investigations, further electroencephalography (EEG) data pertaining to the elicitation of emotions through music will be gathered. Furthermore, the suggested methodology holds potential for the identification of emotions across a wide range of applications.

5 Conclusion

The present work introduces a transformer model as a means of classifying music-evoked emotions. The model under consideration consists of three distinct phases, namely linear embedding, transformer encoder, and MLP layer. The purpose of the first phase is to generate flattened input features for the proposed model. These features are aimed to extract both local and global correlations between the multi-channel EEG data. Additionally, the MLP blocks aim to enhance the classification outcome. This study presents an initial implementation of a vision transformer-based model for the purpose of music emotion identification.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Shandong Management University's Human Research Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

DW: Writing – review & editing, Formal analysis, Validation. JL: Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. HC: Writing – original draft, Validation, Investigation, Formal analysis, Data curation. YZ: Writing – original draft, Validation, Investigation, Data curation.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The present study received support from the Natural Science Foundation of Shandong Province (Grant No. ZR2020MF133), the Key Laboratory of Public Safety Management Technology of Scientific Research and Innovation Platform in Shandong Universities during the 13th Five-Year Plan Period, the Collaborative Innovation Center of "Internet plus intelligent manufacturing" of Shandong Universities, and the Intelligent Manufacturing and Data Application Engineering Laboratory of Shandong Province.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alarcao, S. M., and Fonseca, M. J. (2017). Emotions recognition using EEG signals: a survey. *IEEE Trans. Affect. Comput.* 10, 374–393. doi: 10.1109/TAFFC.2017.2714671

Ba, J., Kiros, J. R., and Hinton, G. E. (2016). Layer normalization. *ArXiv, abs/1607.06450.*

Balasubramanian, G., Kanagasabai, A., Mohan, J., and Seshadri, N. G. (2018). Music induced emotion using wavelet packet decomposition—an EEG study. *Biomed. Sign. Proc. Control* 42, 115–128. doi: 10.1016/j.bspc.2018.01.015

Bong, S. Z., Murugappan, M., and Yaacob, S. B. (2012). "Analysis of electrocardiogram (ECG) signals for human emotional stress classification," in *IEEE International Conference on Robotics and Automation.* doi: 10.1007/978-3-642-35197-6_22

Chen, X., Girshick, R., He, K., and Dollár, P. (2019). "Tensormask: a foundation for dense object segmentation," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2061–2069. doi: 10.1109/ICCV.2019.00215

Choe, J., and Shim, H. (2019). "Attention-based dropout layer for weakly supervised object localization," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2214–2223. doi: 10.1109/CVPR.2019.00232

Cosoli, G., Poli, A., Scalise, L., and Spinsante, S. (2021). "Heart rate variability analysis with wearable devices: influence of artifact correction method on classification accuracy for emotion recognition," in *2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 1–6. doi: 10.1109/I2MTC50364.2021.9459828

Cui, X., Wu, Y., Wu, J., You, Z., Xiahou, J., and Ouyang, M. (2022). A review: Music-emotion recognition and analysis based on EEG signals. *Front. Neuroinf.* 16:997282. doi: 10.3389/fninf.2022.997282

Daly, I. (2023). Neural decoding of music from the EEG. *Sci. Rep.* 13:624. doi: 10.1038/s41598-022-27361-x

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020). An image is worth 16x16 words: transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929.*

Eerola, T., and Vuoskoski, J. K. (2012). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Percept.* 30, 307–340. doi: 10.1525/mp.2012.30.3.307

Ekman, P. (1999). *Basic Emotions, chapter 3*. London: John Wiley Sons, Ltd. 45–60. doi: 10.1002/0470013494.ch3

Eskine, K. (2022). Evaluating the three-network theory of creativity: Effects of music listening on resting state EEG. *Psychol. Music* 51, 730–749. doi: 10.1177/03057356221116141

Fan, H., Xiong, B., Mangalam, K., Li, Y., Yan, Z., Malik, J., et al. (2021). "Multiscale vision transformers," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6824–6835. doi: 10.1109/ICCV48922.2021.00675

Gilda, S., Zafar, H., Soni, C., and Waghurdekar, K. (2017). "Smart music player integrating facial emotion recognition and music mood recommendation," in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 154–158. doi: 10.1109/WiSPNET.2017.8299738

Goshvarpour, A., Abbasi, A., and Goshvarpour, A. (2017). An accurate emotion recognition system using ecg and gsr signals and matching pursuit method. *Biomed. J.* 40, 355–368. doi: 10.1016/j.bj.2017.11.001

Han, D., Kong, Y., Han, J., and Wang, G. (2022). A survey of music emotion recognition. *Front. Comput. Sci.* 16:166335. doi: 10.1007/s11704-021-0569-4

Hasnul, M. A., Aziz, N. A. A., Alelyani, S., Mohana, M., and Aziz, A. A. (2021). Electrocardiogram-based emotion recognition systems and their applications in healthcare–a review. *Sensors* 21:5015. doi: 10.3390/s21155015

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, 2961–2969. doi: 10.1109/ICCV.2017.322

He, Z., Zhong, Y., and Pan, J. (2021). An adversarial discriminative temporal convolutional network for EEG-based cross-domain emotion recognition. *Comput. Biol. Med.* 141:105048. doi: 10.1016/j.compbiomed.2021.105048

Heo, B., Yun, S., Han, D., Chun, S., Choe, J., and Oh, S. J. (2021). "Rethinking spatial dimensions of vision transformers," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 11936–11945. doi: 10.1109/ICCV48922.2021.01172

Homan, R., Herman, J. H., and Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electroencephalogr Clin. Neurophysiol.* 66, 376–382. doi: 10.1016/0013-4694(87)90206-9

Hou, Y., and Chen, S. (2019). Distinguishing different emotions evoked by music via electroencephalographic signals. *Comput. Intell. Neurosci.* 2019:3191903. doi: 10.1155/2019/3191903

Jerritta, S., Murugappan, M., Nagarajan, R., and Wan, K. (2011). "Physiological signals based human emotion recognition: a review," in *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*, 410–415. doi: 10.1109/CSPA.2011.5759912

Kan, H., Yu, J., Huang, J., Liu, Z., and Zhou, H. (2022). Self-supervised group meiosis contrastive learning for EEG-based emotion recognition. *Appl. Intel.* 53, 27207–27225. doi: 10.1007/s10489-023-04971-0

Kipli, K., Latip, A. A. A., Lias, K. B., Bateni, N., Yusoff, S. M., Suud, J. B., et al. (2022). "Evaluation of galvanic skin response (GSR) signals features for emotion recognition," in *International Conference on Applied Intelligence and Informatics* (Cham: Springer Nature Switzerland), 260–274. doi: 10.1007/978-3-031-24801-6_19

Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., et al. (2011). Deap: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15

Lee, M. (2023). Gelu activation function in deep learning: a comprehensive mathematical analysis and performance. *ArXiv, abs/2305.12073*.

Li, Y., Zheng, W., Zong, Y., Cui, Z., Zhang, T., and Zhou, X. (2021). A bi-hemisphere domain adversarial neural network model for EEG emotion recognition. *IEEE Trans. Affect. Comput.* 12, 494–504. doi: 10.1109/TAFFC.2018.2885474

Lin, W.-C., Chiu, H.-W., and Hsu, C.-Y. (2006). "Discovering EEG signals response to musical signal stimuli by time-frequency analysis and independent component analysis," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (IEEE), 2765–2768.

Liu, J., Sun, L., Huang, M., Xu, Y., and Li, R. (2022). Enhancing emotion recognition using region-specific electroencephalogram data and dynamic functional connectivity. *Front. Neurosci.* 16:884475. doi: 10.3389/fnins.2022.884475

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). "Swin transformer: hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 10012–10022. doi: 10.1109/ICCV48922.2021.00986

Minhad, K. N., Ali, S. H. M., and Reaz, M. B. I. (2017). A design framework for human emotion recognition using electrocardiogram and skin conductance response signals. *J. Eng. Sci. Technol.* 12, 3102–3119. doi: 10.1587/transinf.2017EDP7067

Nag, S., Basu, M., Sanyal, S., Banerjee, A., and Ghosh, D. (2022). On the application of deep learning and multifractal techniques to classify emotions and instruments using indian classical music. *Physica A.* 597:127261. doi: 10.1016/j.physa.2022.127261

Ozel, P., Akan, A., and Yilmaz, B. (2019). Synchrosqueezing transform based feature extraction from EEG signals for emotional state prediction. *Biomed. Sig. Proc. Control* 52, 152–161. doi: 10.1016/j.bspc.2019.04.023

Pan, B., and Zheng, W. (2021). Emotion recognition based on EEG using generative adversarial nets and convolutional neural network. *Comput. Mathem. Methods Med.* 2021:2520394. doi: 10.1155/2021/2520394

Pan, J., Liang, R., He, Z., Li, J., Liang, Y., Zhou, X., et al. (2023). St-scgnn: a spatio-temporal self-constructing graph neural network for cross-subject EEG-based emotion recognition and consciousness detection. *IEEE J. Biomed. Health Inf.* 28, 777–788. doi: 10.1109/JBHI.2023.3335854

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). "Pytorch: an imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems,* 32.

Qiu, L., Zhong, Y., Xie, Q., He, Z., Wang, X., Chen, Y., et al. (2022). Multi-modal integration of EEG-fnirs for characterization of brain activity evoked by preferred music. *Front. Neurorob.* 16:823435. doi: 10.3389/fnbot.2022.823435

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18* (Springer), 234–241. doi: 10.1007/978-3-319-24574-4_28

Rudakov, E. (2021). "Multi-task CNN model for emotion recognition from EEG brain maps," in *2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART), New York, NY USA* (IEEE). doi: 10.1109/BioSMART54244.2021.9677807

Sammler, D., Grigutsch, M., Fritz, T., and Koelsch, S. (2007). Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* 44, 293–304. doi: 10.1111/j.1469-8986.2007.00497.x

Selvaraj, J., Murugappan, M., Wan, K., and Yaacob, S. B. (2013). Classification of emotional states from electrocardiogram signals: a non-linear approach based on hurst. *BioMed. Eng. OnLine* 12:44. doi: 10.1186/1475-925X-12-44

Semerci, Y. C., Akgün, G., Toprak, E., and Barkana, D. E. (2022). "A comparative analysis of deep learning methods for emotion recognition using physiological signals for robot-based intervention studies," in *2022 Medical Technologies Congress (TIPTEKNO)*, 1–4. doi: 10.1109/TIPTEKNO56568.2022.9960200

Serra, H., Oliveira, J. P., and Paulino, N. F. (2017). "A 50 hz sc notch filter for iot applications," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–4. doi: 10.1109/ISCAS.2017.8050904

Shawky, E., El-Khoribi, R., Shoman, M., and Wahby Shalaby, M. (2018). EEG-based emotion recognition using 3D convolutional neural networks. *Int. J. Adv. Comput. Sci. Applic.* 9:843. doi: 10.14569/IJACSA.2018.090843

Shen, F., Dai, G., Lin, G., Zhang, J., Kong, W., and Zeng, H. (2020). EEG-based emotion recognition using 4D convolutional recurrent neural network. *Cogn. Neurodyna.* 14, 1–14. doi: 10.1007/s11571-020-09634-1

Sheykhivand, S., Mousavi, Z., Rezaii, T. Y., and Farzamnia, A. (2020). Recognizing emotions evoked by music using CNN-LSTM networks on EEG signals. *IEEE Access* 8, 139332–139345. doi: 10.1109/ACCESS.2020.3011882

Siddiqui, H. U. R., Shahzad, H. F., Saleem, A. A., Khakwani, A. B. K., Rustam, F., Lee, E., et al. (2021). Respiration based non-invasive approach for emotion recognition using impulse radio ultra wide band radar and machine learning. *Sensors (Basel, Switzerland)* 21:8336. doi: 10.3390/s21248336

Song, Z. (2021). Facial expression emotion recognition model integrating philosophy and machine learning theory. *Front. Psychol.* 12:759485. doi: 10.3389/fpsyg.2021.759485

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). "Attention is all you need," in *Advances in Neural Information Processing Systems,* 30.

Vuilleumier, P., and Trost, W. (2015). Music and emotions: from enchantment to entrainment. *Ann. NY Acad. Sci.* 1337, 212–222. doi: 10.1111/nyas.12676

Wang, W., Xie, E., Li, X., Fan, D.-P., Song, K., Liang, D., et al. (2021). "Pyramid vision transformer: a versatile backbone for dense prediction without convolutions," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 568–578. doi: 10.1109/ICCV48922.2021.00061

Wei, W., Jia, Q., Feng, Y., and Chen, G. (2018). Emotion recognition based on weighted fusion strategy of multichannel physiological signals. *Comput. Intell. Neurosci.* 2018:5296523. doi: 10.1155/2018/5296523

Wu, B., Xu, C., Dai, X., Wan, A., Zhang, P., Yan, Z., et al. (2020). Visual transformers: Token-based image representation and processing for computer vision. *arXiv preprint arXiv:2006.03677*.

Xing, X., Li, Z., Xu, T., Shu, L., and Xu, X. (2019). SAE+LSTM: a new framework for emotion recognition from multi-channel EEG. *Front. Neurorobot.* 13:37. doi: 10.3389/fnbot.2019.00037

Xu, M., Cheng, J., Li, C., Liu, Y., and Chen, X. (2023). Spatio-temporal deep forest for emotion recognition based on facial electromyography signals. *Comput. Biol. Med.* 156:106689. doi: 10.1016/j.compbiomed.2023.106689

Yang, C., Xu, J., Mello, S. D., Crowley, E. J., and Wang, X. (2022). Gpvit: a high resolution non-hierarchical vision transformer with group propagation. *ArXiv, abs/2212.06795*.

Yang, H., Han, J., and Min, K. (2019). A multi-column cnn model for emotion recognition from EEG signals. *Sensors* 19:4736. doi: 10.3390/s19214736

Yang, Y., Wu, Q., Qiu, M., Wang, Y., and Chen, X. (2018). "Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–7. doi: 10.1109/IJCNN.2018.8489331

Zhang, H., Yi, P., Liu, R., and Zhou, D. (2021). "Emotion recognition from body movements with as-LSTM," in *2021 IEEE 7th International Conference on Virtual Reality (ICVR)*, 26–32. doi: 10.1109/ICVR51878.2021.9483833

Zhang, Z., Zhong, S. H., and Liu, Y. (2022). Ganser: a self-supervised data augmentation framework for EEG-based emotion recognition. *IEEE Trans. Affect. Comput*. 14, 2048–2063. doi: 10.1109/TAFFC.2022.3170369

Zhou, X., Zhuo, J., and Krahenbuhl, P. (2019). "Bottom-up object detection by grouping extreme and center points," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 850–859. doi: 10.1109/CVPR.2019.00094

frontiers | Frontiers in Psychology

TYPE Correction
PUBLISHED 23 May 2024
DOI 10.3389/fpsyg.2024.1422498

Check for updates

OPEN ACCESS

APPROVED BY
Frontiers Editorial Office,
Frontiers Media SA, Switzerland

*CORRESPONDENCE
Yanan Zhou
zhouyanan@bfsu.edu.cn
Hebin Cheng
chenghebin@sdmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 24 April 2024
ACCEPTED 13 May 2024
PUBLISHED 23 May 2024

CITATION
Wang D, Lian J, Cheng H and Zhou Y (2024) Corrigendum: Music-evoked emotions classification using vision transformer in EEG signals. *Front. Psychol.* 15:1422498. doi: 10.3389/fpsyg.2024.1422498

Corrigendum: Music-evoked emotions classification using vision transformer in EEG signals

Dong Wang[1†], Jian Lian[1†], Hebin Cheng[1*] and Yanan Zhou[2*]

[1]School of Intelligence Engineering, Shandong Management University, Jinan, China, [2]School of Arts, Beijing Foreign Studies University, Beijing, China

KEYWORDS

music-evoked emotion, emotion classification, electroencephalographic, deep learning, transformer

A corrigendum on
Music-evoked emotions classification using vision transformer in EEG signals

by Wang, D., Lian, J., Cheng, H., and Zhou, Y. (2024). *Front. Psychol*. 15:1275142. doi: 10.3389/fpsyg.2024.1275142

In the published article, there was an error. It was not clearly denoted that authors Dong Wang and Jian Lian contributed equally to this work. The correct information is shown below.

Dong Wang[1†], Jian Lian[1†], Hebin Cheng[1*] and Yanan Zhou[2*]

†These authors have contributed equally to this work

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

TYPE Correction
PUBLISHED 12 August 2024
DOI 10.3389/fpsyg.2024.1466709

Check for updates

OPEN ACCESS

APPROVED BY
Frontiers Editorial Office,
Frontiers Media SA, Switzerland

*CORRESPONDENCE
Yanan Zhou
✉ zhouyanan@bfsu.edu.cn
Hebin Cheng
✉ chenghebin@sdmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 18 July 2024
ACCEPTED 26 July 2024
PUBLISHED 12 August 2024

CITATION
Wang D, Lian J, Cheng H and Zhou Y (2024) Corrigendum: Music-evoked emotions classification using vision transformer in EEG signals. *Front. Psychol.* 15:1466709.
doi: 10.3389/fpsyg.2024.1466709

Corrigendum: Music-evoked emotions classification using vision transformer in EEG signals

Dong Wang[1,2†], Jian Lian[2†], Hebin Cheng[2*] and Yanan Zhou[3*]

[1]School of Information Science and Electrical Engineering, Shandong Jiaotong University, Jinan, China, [2]School of Intelligence Engineering, Shandong Management University, Jinan, China, [3]School of Arts, Beijing Foreign Studies University, Beijing, China

KEYWORDS

music-evoked emotion, emotion classification, electroencephalographic, deep learning, transformer

A corrigendum on
Music-evoked emotions classification using vision transformer in EEG signals

by Wang, D., Lian, J., Cheng, H., and Zhou, Y. (2024). *Front. Psychol*. 15:1275142.
doi: 10.3389/fpsyg.2024.1275142

In the published article, there were errors in the affiliations for author Dong Wang. As well as 'School of Intelligence Engineering, Shandong Management University, Jinan, China', this author should also be affiliated to "School of Information Science and Electrical Engineering, Shandong Jiaotong University, Jinan, China". The corrected author affiliations are shown below.

Dong Wang[1,2†], Jian Lian[2†], Hebin Cheng[2*] and Yanan Zhou[3*]

[1]School of Information Science and Electrical Engineering, Shandong Jiaotong University, Jinan, China

[2]School of Intelligence Engineering, Shandong Management University, Jinan, China

[3]School of Arts, Beijing Foreign Studies University, Beijing, China

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 22 October 2024
DOI 10.3389/fpsyg.2024.1400223

Check for updates

OPEN ACCESS

EDITED BY
Ivonne Castiblanco Jimenez,
Polytechnic University of Turin, Italy

REVIEWED BY
Robert C. A. Bendall,
University of Salford, United Kingdom
Alexandru I. Tiba,
University of Oradea, Romania

*CORRESPONDENCE
Isabell Int-Veen
✉ isabell.int-veen@med.uni-tuebingen.de

†These authors have contributed equally to this work and share last authorship

RECEIVED 13 March 2024
ACCEPTED 01 October 2024
PUBLISHED 22 October 2024

CITATION
Int-Veen I, Volz M, Kroczek A, Fallgatter AJ, Ehlis A-C, Rubel JA and Rosenbaum D (2024) Emotion regulation use in daily-life and its association with success of emotion-regulation, self-efficacy, stress, and state rumination.
Front. Psychol. 15:1400223.
doi: 10.3389/fpsyg.2024.1400223

Emotion regulation use in daily-life and its association with success of emotion-regulation, self-efficacy, stress, and state rumination

Isabell Int-Veen[1]*, Magdalena Volz[2], Agnes Kroczek[1], Andreas J. Fallgatter[1,3,4], Ann-Christine Ehlis[1,3,4], Julian A. Rubel[2†] and David Rosenbaum[1†]

[1]Department of Psychiatry and Psychotherapy, Tübingen Center for Mental Health (TüCMH), University Hospital of Tübingen, Tübingen, Germany, [2]Psychotherapy Research Unit, Department of Psychology, Osnabrueck University, Osnabrueck, Germany, [3]German Center for Mental Health (DZPG), Tübingen, Germany, [4]LEAD Graduate School and Research Network, University of Tübingen, Tübingen, Germany

Introduction: Investigations on emotion regulation strategies (ERS) primarily focus on the influence of instructed emotion regulation (ER) on outcomes. However, recent work has shown that selection of ERS is dependent on, e.g., situational demands and personal resources.

Methods: In this current investigation, we used an online diary to investigate ERS used by free choice and their association with ER-success, stress and rumination. We identified four factors of ERS: cognitive perspective change, cognitive-behavioral problem-solving, suppression-distraction and body-social ERS. Associations of ERS with stress, state-rumination and ER-success were investigated using multilevel-mixed-models, allowing to separate within- and between-subject effects.

Results: Our results show that, on a within-subject level, all adaptive ERS were positively associated with ER-success, while maladaptive ERS as well as higher stress and state rumination were negatively associated with ER-success. On the other hand, only within-subject cognitive ERS were associated with higher self-efficacy. Maladaptive ERS-use was consequently positively associated with stress and state rumination. Surprisingly, only cognitive perspective change ERS were negatively associated with state rumination. Cognitive-behavioral problem-solving was positively associated with stress and success of emotion regulation.

Discussion: We interpret these results in the light of situational constraints of ERS-use and the importance of the assessment of these in future studies.

KEYWORDS
ecological momentary assessment, rumination, stress, emotion regulation, daily diary

Introduction

As soon as someone experiences emotions, one has to deal with them in one way or the other. This is also referred to as emotion regulation (ER) and has been intensely studied for the last few decades (Gross, 2014). The process model of emotion regulation (Gross, 1998, p. 275) defines ER as "[…] the processes by which individuals influence which emotions they have, when they have them, and how they experience and express these emotions. Emotion regulatory processes may be automatic or controlled, conscious or unconscious, and may have their effects at one or more points in the emotion generative process." More specifically, there are five stages where the emotion-generative process might be altered by ER: the situation selection (e.g., avoidance), the situation modification (e.g., behavioral problem-solving), attentional deployment (e.g., distraction, rumination), cognitive change (e.g., acceptance, reappraisal) and response modulation (e.g., suppression). Initially, research was primarily focused on the use of single ER-strategies (ERS) trying to identify them and differentiate their consequences and efficacy by experimental emotion-induction (Webb et al., 2012). However, recent studies found that often multiple ERS are used concurrently or sequentially, which is also referred to as emotion-polyregulation (Aldao and Nolen-Hoeksema, 2013; Brans et al., 2013; Ford et al., 2019). Ford et al. (2019) have only recently integrated polyregulation into the process model of ER by Gross (2015).

It is important to study temporal dynamics of the involved processes in order to capture them in an ecologically valid way. In naturalistic observation studies using for instance Ecological Momentary Assessment (EMA) (Koval et al., 2020; for a comprehensive overview see Koval and Kalokerinos, 2024) and experimental studies, research has shown a crucial impact of contextual factors on ER (Kobylińska and Kusev, 2019). For example, the type and intensity of emotion (Aldao, 2013; Barrett et al., 2001; Dixon-Gordon et al., 2015), motivation (Tamir et al., 2020), timing of implementation (Diedrich et al., 2016), and perceived controllability of the situation (Troy et al., 2013) influence not only the ER-success but also ERS-choice (Koval et al., 2015; Matthews et al., 2021; Sheppes et al., 2014). Interestingly, ER-success and ERS-choice are differently influenced by psychopathology: A recent study found that strategy selection rather than the implementation of those strategies is associated with mental disorders (Houben et al., 2023).

Several strategies have been identified that are typically used in case of higher stressor intensity or more emotionally loaded situations reflected by higher negative affect (putatively maladaptive ERS, e.g., distraction, rumination) and conversely others in case of lower stressor-intensity or lower levels of negative affect (putatively adaptive ERS, e.g., acceptance, reappraisal), which is probably a consequence of decreased cognitive resources and higher cognitive load (Aldao, 2013; Barrett et al., 2001; Blanke et al., 2022; Broderick, 2005; Dixon-Gordon et al., 2015; Sheppes, 2020; Sheppes et al., 2011, 2014). Further, ERS have been found to vary in their temporal deployment, with suppression and rumination occurring more at the beginning, and reappraisal and distraction occurring more toward the end of a negative emotional episode (Kalokerinos et al., 2017).

Of particular interest for this investigation is the association of ER with rumination. Originally, rumination was seen as a cognitive vulnerability to develop depressive disorders (Nolen-Hoeksema and Morrow, 1991). It is regarded as an abstract negative thinking-style where thoughts revolve around the past while having no goal-orientation (Teismann et al., 2012). In the context of Gross' model (2015), rumination is categorized as a form of attentional deployment where there is repetitively a passive, self-immersed focus on the emotional features and consequences of a situation. Ruminative processes and other forms of repetitive negative thought have been found not only in depression but also other psychopathologies such as anxiety, eating and substance-related disorders (Arditte et al., 2016; McLaughlin and Nolen-Hoeksema, 2011; Svaldi et al., 2012). Rumination is also observed in healthy individuals, for instance in response to stressful life events (Moberly and Watkins, 2006; Robinson and Alloy, 2003; Ruscio et al., 2015). Ruminative thinking has been found to be highly persistent by predicting future rumination, increased negative affect and depressive symptoms (Bean et al., 2020; Boemo et al., 2022; Connolly and Alloy, 2017; Kircanski et al., 2018; Moberly and Watkins, 2008; Rosenbaum et al., 2021, 2022; Rosenbaum et al., 2018a, 2018b; Ruscio et al., 2015). Interestingly, in an EMA-study by Koval et al. (2012) rumination and emotional inertia both independently predicted depression severity in healthy undergraduates as well as depressed patients.

The scope of this study was to investigate different patterns of ERS used in response to daily life stress and their association with stress-reactive rumination and ER-success and perceived self-efficacy. Ten predefined ERS were assessed in a large community sample. As the pattern of ERS use showed high interrelationships, we performed an exploratory multilevel factor analysis in order to reduce data complexity. We aimed to find patterns of which ERS are used most commonly at the same time window and how those groups were associated with ER-success (direct effect measure), self-efficacy, reduction in stress (indirect effect measure) and reduction in momentary rumination (indirect effect measure). We set up the following hypotheses: (I) We expected to find decreased ER-success in case of higher stress and higher state-rumination and higher success in case of putatively adaptive (and lower in case of maladaptive) ERS. (II) We expected to find decreased self-perceived self-efficacy in case of higher stress and higher state-rumination and higher self-efficacy in case of putatively adaptive (and lower in case of maladaptive) ERS. (III) Concerning stress, we expected to find higher stress in the evening in case of higher stress in the morning and higher rumination (Rosenbaum et al., 2022). Further, we expected to find adaptive ERS to be associated with reduced stress and rumination. (IV) Analogously, state rumination should be elevated in case of higher previous state rumination and higher stress. Likewise, we expected adaptive ERS to be negatively associated with rumination.

Methods

Sample

A total of 627 participants aged 18 years or older and fluent in German were recruited via flyers, emails and social media and completed the online assessment of demographic data. A total of 532 participants set up the online diary after giving written informed consent. All procedures were approved by the ethics committee at the University Hospital and University of Tübingen and in line with the Declaration of Helsinki in its latest version. After preprocessing, the final sample consisted of 144 participants with two data entries per day without day-night-shifts, ≥6 complete per-day data entries and no indication of careless responding (see Supplementary material S1).

Procedure

After participants received information regarding the study procedure and provided informed consent, demographic data and baseline questionnaires including a questionnaire assessing habitual ER

(FEEL-E; Grob and Horowitz, 2014), as well as the Perseverative Thinking Questionnaire (PTQ; Ehring et al., 2011), Ruminative Response Scale (RRS; Nolen-Hoeksema and Morrow, 1991), Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003) and Becks Depression Inventory II (BDI-II; Hautzinger et al., 2009) were assessed. Then, participants received instructions regarding the online-diary-setup, installing it on their own smartphones using the PsyAssessor researcher-edition V2, 2019 (Machine Learning Solutions, Luxembourg). Online diary entries were assessed for 14 consecutive days where participants received emails instructing them to enter data, once at midday and once in the evening, ~5 h apart. The exact times could be freely chosen and adapted anytime. In case no data was entered within 30 min, the corresponding data point was defined as missing. At the end of the study, participants received 15€ or course credit in case they completed >50% of all data entries (see Figure 1).

Questionnaires

FEEL-E

In order to assess habitual emotion regulation strategy use, participants completed the German "Fragebogen zur Erhebung der Emotionsregulation bei Erwachsenen" (FEEL-E; Grob and Horowitz, 2014). This questionnaire assesses how the corresponding person deals with different emotions (anxiety, sadness, and anger) by rating the degree to which they use the corresponding strategy in case of dealing with the specific emotion whereby two items correspond to the same strategy. Six adaptive strategies (problem-oriented, acceptance, cognitive problem solving, reevaluation, positive mood, forgetting) and six maladaptive strategies (withdrawal, self-devaluation, giving up, rumination, negative thinking, allocating blame) are rated on 5-point-Likert-Scale (1 = almost never, 5 = almost always). Like this, the FEEL-E allows to investigate emotion regulation in an emotion-specific as well as nonspecific, general way. Internal consistencies for the subscales has been proven to be high (adaptive strategies: $\alpha = 0.91$, maladaptive strategies: $\alpha = 0.88$). Further, test–retest-reliability after 8 months is relatively high: $r_{tt} = 0.79$ for both subscales.

PTQ

Trait rumination was assessed using the Perseverative Thinking Questionnaire comprising 15 items (Ehring et al., 2011): "The item pool comprised three items for each of the assumed process characteristics of repetitive negative thinking: (1a) repetitive (e.g., "The same thoughts keep going through my mind again and again"), (1b) intrusive (e.g., "Thoughts come to my mind without me wanting them to"), (1c) difficult to disengage from (e.g., "I cannot stop dwelling on them"), (2) unproductive (e.g., "I keep asking myself questions without finding an answer"), (3) capturing mental capacity (e.g., "My thought prevent me from focusing on other things")." Participants rate each item on a scale ranging from "0 = never" to "4 = almost always." Using a non-clinical and clinical sample, the PTQ has shown to have a high internal consistency (Cronbach's $\alpha > 0.94$; Ehring et al., 2011) and acceptable reliability ($r_{tt} = 0.69$; Ehring et al., 2011).

RRS

In order to assess inter-individual levels of trait rumination, the self-report Ruminative Response Scale (RRS; Nolen-Hoeksema and Morrow, 1991), a subscale of the Response Style Questionnaire (RSQ; Nolen-Hoeksema and Morrow, 1991), was used. The RRS consists of a total of 22 items which are rated on 4-point-Likert-Scales ranging from 1 = "almost never" to 4 = "almost always" and resulting in a total score ranging between 22 and 88. A high internal consistency has been observed in several studies and samples (Cronbach's $\alpha > 0.88$; Just and

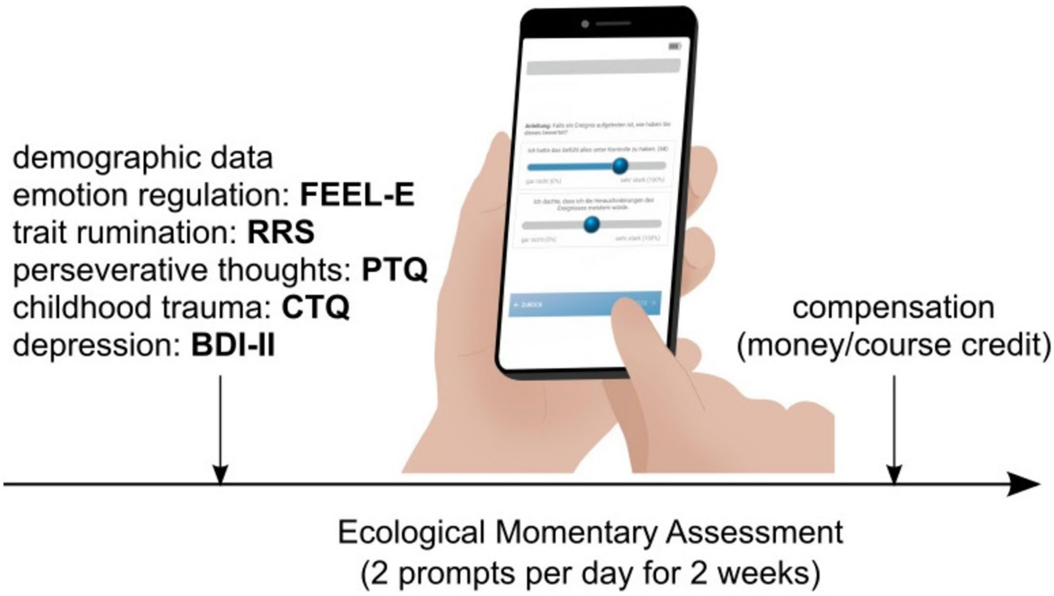


FIGURE 1
Overview over the time course of the study. FEEL-E, questionnaire assessing habitual emotion regulation; PTQ, Perseverative thinking questionnaire; RRS, Ruminative response scale; CTQ, Childhood trauma questionnaire; BDI-II, Beck depression inventory-II.

Alloy, 1997; Kasch et al., 2001; Moberly and Watkins, 2008; Nolen-Hoeksema and Morrow, 1991) including studies using the German version of the RRS (Cronbach's $\alpha = 0.89–0.92$; Wahl et al., 2011). Test–retest reliability, however, has been proven to fluctuate across different time spans as well as clinical and non-clinical samples: In case of non-clinical samples, test–retest reliability typically ranges between $r_{tt} = 0.80$ over 6 months (Nolen-Hoeksema et al., 1994) and $r_{tt} = 0.67$ over one year (Nolen-Hoeksema et al., 1999). In clinical samples, test–retest scores ranged between $r_{tt} = 0.36$ over 6 months (Kasch et al., 2001), and $r_{tt} = 0.47$ over one year (Just and Alloy, 1997).

CTQ

We used the Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003) to assess self-reported adverse childhood experiences at the age of 0–17 years. The 28-item short version of the original 70-item CTQ (Bernstein et al., 1994) comprises 25 clinical items and three validity items to screen denial (e.g., "I had the best family in the world"). The clinical items belong to five empirically derived subscales distinguishing emotional abuse, physical abuse, sexual abuse, emotional neglect and physical neglect which are summed up to the total score (range: 25–125). All statements are rated on 5-point-Likert-Scales ranging from 1 = "never true" to 5 = "very often true." Overall, the short version of the CTQ has demonstrated satisfactorily fulfilled quality criteria. Using multiple clinical and non-clinical samples, Bernstein et al. (2003) found a consistent five-factor structure and good evidence of criterion-related validity which was evaluated by therapists' maltreatment ratings using a structured interview, information provided by the patient as well as data of child protective investigations. Using a community sample, acceptable internal consistency of CTQ total scores (Cronbach's $\alpha = 0.91$) and subscale scores were observed (ranging from Cronbach's $\alpha = 0.58$ for physical neglect to Cronbach's $\alpha = 0.94$ for sexual abuse; Scher et al., 2001). Using the original version of the CTQ, high test–retest reliability values were found ranging between $r_{tt} = 0.79–0.81$ (Bernstein et al., 1994; Bernstein and Fink, 1988). Similar results were replicated by various other researchers (Burns et al., 2010, 2012; Huh et al., 2017), and most importantly also using a German translation in clinical as well as non-clinical samples (Bader et al., 2009; Klinitzke et al., 2012; Wingenfeld et al., 2010).

BDI-II

To assess the severity of depression symptoms, we utilized the Beck Depression Inventory II, a self-report questionnaire initially developed by Beck et al. (1961) and translated into German by Hautzinger et al. (1994). Following updates in diagnostic manuals, a revised version was created by Beck et al. (1966) and translated again by Hautzinger et al. (2009). The questionnaire evaluates the presence of 21 symptoms over the past two weeks, with symptom severity quantified as a total score ranging from 0 to 63. Additionally, cut-off scores are provided to aid in interpreting these total scores. Wang and Gorenstein (2013) reported high internal consistency (Cronbach's α around 0.9) and high test–retest reliability (mean interval of 2 weeks; r_{tt} around 0.7–0.9) across various populations and languages. Notably for our study, the German version has demonstrated good ability to differentiate between depressed patients and healthy controls (Kühner et al., 2007) and is regarded as an effective screening tool for Major Depressive Disorder (Kumar et al., 2002).

Online diary

Firstly, participants stated whether something pleasant/unpleasant had happened during the past 5 h (yes/no) and could enter a free text. Using two items which were to be averaged in one factor representing self-efficacy, participants rated their agreement to the following statements on a slider (0–100%): "I felt like I was in control of everything." and "I thought I would be able to overcome the challenges of the event." Next, stress was assessed using a slider (0–100%) and current mood using a circumplex (arousal: aroused/relaxed, valence: positive/negative). We then assessed state rumination using three modified items of the RRS and three modified items of the Perseverative Cognitions Questionnaire (Szkodny and Newman, 2019) where participants rated their agreement using 5-point-Likert-Scales. We used this scale in another study where internal reliability was proven to be high ($\alpha = 0.91$; Rosenbaum et al., 2022), which was replicated in the current sample ($\alpha = 0.89$). Next, the need for ER was assessed (yes/no). Lastly participants answered whether they had applied the corresponding ERS: cognitive problem-solving, problem-oriented action/behavioral activation, reappraisal, self-compassion, acceptance, mindfulness, suppression, distraction, social support and body-based regulation. These ERS were extracted from a prior study investigating ERS-use in an open answer format qualitatively (Rosenbaum et al., 2022). To avoid ambiguities in ERS-definitions, we provided examples (see Supplementary material S2). Then, four items assessed cognitive processes and ER-success was rated (0–100%): "How successful have you been in regulating your emotions in the last 5 h?." We chose to include this direct measure of ER-success as indirect measures such as stress and state rumination might be entangled with potential situational constraints and the regulation process itself (see Figure 2). Each diary entry took ~2 min (for all items see Supplementary material S3).

Data preprocessing

To be able to fit cross-lagged mixed models, we excluded data entries with missing subsequent assessments and created time-lag variables for state rumination, stress and ERS-use. Participants could choose the times of data assessment freely, which is why we further removed cases where data entries were < 3 or > 18 h apart or with an alternation of day/night. We identified careless responses using the Anomaly-Case-Index-List, which reflects the unusualness of a record with respect to the group deviation it belongs to. Cases with an index >2 were removed. We further excluded participants in case there were < 6 day-complete data-entries. Lastly, in order to differentiate within- and between-person-effects of stress and state rumination, we used person-mean-centering (Enders and Tofighi, 2007; Falkenström et al., 2017; Hoffman and Stawski, 2009) while all level-2-variables were grand-mean-centered.

Data analysis

Data analysis was done using SPSS (IBM Corp, 2019), Mplus version 8.9 (Muthén and Muthén, 2017), RStudio Version 1.4.1717 (RStudio Team, 2021) and R Version 4.0.3 (R Core Team, 2020) using the packages rmcorr (Bakdash and Marusich, 2017) for

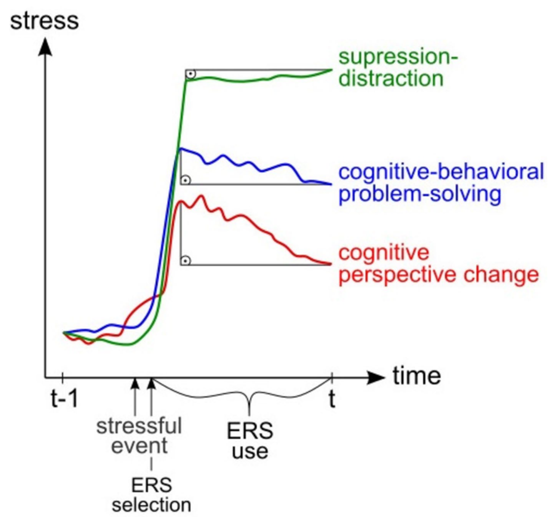


FIGURE 2
Schematic illustration of the assumed association of stress (y-axis), ER-success (displayed as rectangles) and stress-dependent ERS-use. As ERS are most probably selected in terms of situational constraints (here, different ERS are used in the case of different events), direct measures of ER-success are needed. Here, cognitive-behavioral problem-solving could be positively associated with stress in comparison to other ERS at (t) despite the undetected decline in stress after stressor-onset.

repeated-measures correlations, lme4 for mixed models (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017). As ERS-use was highly interrelated, we performed an exploratory multilevel factor analysis in Mplus in order to identify different factors. We then fitted autoregressive mixed models using maximum-likelihood-estimation in R. Please note that according to our hypotheses, we only included data entries where participants reported to have used at least one ERS. This study was not preregistered and no *a-priori* power analysis was conducted.

In order to investigate ER-success, we fitted a model with previous ER-success (t-1), current state rumination and stress as within- (WP) and between (BP)-participants factors (all variables with reference to the last 5 h). We further included the mean ERS-use of the different factors (t) (WP and BP). In a more complex model, we also added self-efficacy as WP- and BP-effects. The same models were fitted for self-efficacy instead of ER-success as dependent variable. To investigate predictors of stress at t, the basic model included previous stress (t-1) and current state rumination (t) (WP and BP) and the mean ERS-use of the different factors (t) (WP and BP). In the more complex models, we further added self-efficacy and ER-success separately as well as both at the same time. In all models, intercepts and auto-regressive effects were included as random effects allowing for heterogeneity between participants. Lastly, we fitted the analogue models for state rumination. Please note that for all dependent variables, we evaluated whether the interaction effects of stress and/or rumination explained significantly more variance compared to the most complex models without interaction effects using a Likelihood-Ratio-Test. This was never the case, which is why we report the less complex model here and the interaction-models in Supplementary material S4. In the article itself, we report the predictors of the models, as well as AIC and BIC. The correlation matrix of the predictors of the most complex model can be found in Supplementary material S5 for clarity. None of the predictors correlated above 0.49 with each other.

Results

Participants

On average, participants made a total of 8.51 subsequent data entries (SD = 2.25, range: 6–15; Please note that one participant entered data on 15 subsequent days) and were 24.47 years old (SD = 6.56). 90.28% of the sample were female, 9.03% male and 0.69% non-binary. 11.11% of the sample was currently/within the last 2 months in psychotherapeutic treatment (81.25% of those in treatment received (cognitive) behavioral psychotherapy, 13.50% received psychodynamic psychotherapy), while 22.92% of the total sample was diagnosed with at least one psychiatric disorder (most frequent primary diagnoses were F32 (depressive episode), $n = 13$; F50 (eating disorders), $n = 7$; F40 (phobic anxiety disorders), $n = 6$; for details see Supplementary material S6). 93.06% of the sample had never been in psychological in-patient care (0.69% once, 2.78% twice and 3.47% three or more times). The mean RRS was 31.31 (SD = 11.32), the mean BDI-II was 14.95 (SD = 10.55). Please note that this BDI-II-score is most probably biased due to the high number of females investigated (Roelofs et al., 2013).

This finding is in line with previous research investigating the BDI-II and or RRS in community samples (Economou et al., 2024; Faro and Pereira, 2020; Gomes-Oliveira et al., 2012; Roelofs et al., 2013).

Extraction of ERS-factors

After examining the data, it became clear that most participants reported the use of several ERS at a given time point. Namely, in 18.76% of all data entries, five different ERS were used concurrently, in 18.60% three, in 17.54% four, followed by six (15.82%), two (10.03%) and seven ERS (8.97%). Only in 3.59% of all data entries only one ERS was used, followed by eight concurrently used ERS in 3.43% and nine in 3.26% of data entries. Never were all 10 ERS used at the same time point. In 54.89% of all data entries, cognitive problem-solving was used, in 42.17% acceptance, followed by behavioral problem-solving, which was used in 35.73% of all cases, self-compassion (35.32%), distraction (34.75%) and reframing (30.59%), mindfulness (27.57%), suppression (27.49%), seeking social support (24.55%) and body-based regulation, which was used in 21.78% of all cases. Therefore, we decided to reduce complexity by performing an exploratory multilevel factor analysis with Oblimin rotation in Mplus. Please note that for this analysis, only participants with WP-variation within each ERS were included ($n = 38$). According to the Scree-Plot and the Kaiser-Guttman-criterion this resulted in four factors: Factor 1 consisted of self-compassion, acceptance, reframing and mindfulness, factor 2 included cognitive and behavioral problem-solving, factor 3 included suppression and distraction and factor 4 included body-based and social regulation. We named the four factors as follows: cognitive perspective change, cognitive-behavioral problem-solving, suppression-distraction, body-social (see Supplementary material S7). The Mplus output is to be found in Supplementary material S8. We also performed a PCA using SPSS

(IBM Corp, 2019, Version 28.0) where all data entries were included to identify different "clusters." The authors are aware that PCA does not take hierarchical data into account. Nevertheless, the results replicate the findings of the Mplus analysis where only a smaller subsample could be considered. The analysis is to be found in Supplementary material S9. Note that in the free text additional ERS have been reported that are not analyzed in this work. We decided to do so as the free text format is very unspecific (e.g., with respect to adaptive/maladaptive facets of strategies like eating) and the total frequencies of those strategies were low. However, for future investigations we would like to report those free text categories that are not captured by our setup so far. The following ERS have been reported: Drinking alcohol, smoking (related, but not reported, would be consuming drugs), eating sweets, eating and vomiting, praying, sleeping/taking naps, shopping, having sex, self-harm and risk behavior, writing diaries/self-expression/art (see Supplementary material S10).

Descriptive and correlational analysis of concurrent ERS-use

Still, ERS-factors were frequently used in combination with each other (see Supplementary materials S11, S12). Cognitive perspective change was descriptively used most frequently compared to all other factors (in a total of 1,040 data entries, see Supplementary material S11) whereas body-social ERS were used least frequently (in 819 data entries, see Supplementary material S11). While cognitive perspective change was most often used with two other factors (in 41.63% of all cases when it was used, see Supplementary material S12), namely cognitive behavioral problem-solving and suppression-distraction (see Supplementary material S10), all of the other factors were most often used with all three other factors (see Supplementary material S12). This resulted in medium positive correlations on a between-subject level among cognitive perspective change with cognitive behavioral problem-solving, $r(142) = 0.444$, $p < 0.001$, and body-social ERS, $r(142) = 0.313$, $p < 0.001$, as well as of body-social with cognitive behavioral problem-solving, $r(142) = 0.246$, $p < 0.01$ (see Supplementary material S13). On a within-person level, however, we found descriptively lower but also negative correlations (see Supplementary material S14): Cognitive perspective change was negatively associated with cognitive behavioral problem-solving, $r(1081) = -0.069$, $p < 0.05$, and with suppression-distraction, $r(1081) = -0.098$, $p < 0.01$, but positively associated with body-social ERS, $r(1081) = 0.086$, $p < 0.01$. Body-social ERS were further also negatively associated with within-person use of cognitive behavioral problem-solving, $r(1081) = -0.128$, $p < 0.001$, as well as with suppression-distraction, $r(1081) = -0.061$, $p < 0.05$ (see Supplementary material S14).

Direct ER-success

Fitting mixed models predicting ER-success (t), higher state rumination on a WP- (t) and BP-level as well as stress (t) on a WP-level were negatively associated with ER-success. Higher state rumination and stress were associated with decreased ER-success. In case of the ERS-use factors, we observed a significant association of WP-effects for all factors: increased ER-success in case of applying cognitive perspective change, cognitive-behavioral problem-solving and body-social ERS and decreased ER-success in case of suppression-distraction. Furthermore, we observed significant BP-effects for cognitive perspective change and the use of body-social ERS indicating increased ER-success in case of individuals using cognitive perspective change and body-social ERS more often compared to others. When self-efficacy was added to the previous model, BP-stress now yielded significance and self-efficacy on a WP- and BP-level (see Table 1).

Self-efficacy

We found higher previous self-efficacy (t-1) to be significantly associated with higher self-efficacy at t. Further, we observed lower previous WP-state rumination to yield a significant predictor whereas stress did not yield significance. Concerning ERS-factors we found higher WP-use of cognitive perspective change and cognitive-behavioral problem-solving to be associated with higher concurrent self-efficacy. When adding ER-success, we found WP-stress to now yield significance and the effect of WP-cognitive perspective change to diminish while ER-success was on a WP- and BP-level significantly associated with higher self-efficacy (see Table 2).

Stress

Fitting our models for stress, state rumination on a BP- and WP-level was positively associated with stress. On a BP-level, an increased use of cognitive perspective change was associated with lower stress while an increased use of cognitive-behavioral problem-solving was associated with significantly higher stress. On a WP-level, cognitive-behavioral problem-solving and suppression-distraction was associated with higher stress. Adding self-efficacy did not yield significance while adding ER-success was then significantly associated with higher stress on a WP- and BP-level. Further, the BP-effect of cognitive perspective change and the effect of WP-suppression-distraction no longer yielded significance. When we added ER-success and self-efficacy, WP-self-efficacy was associated with higher stress while ER-success remained negatively associated. Lastly, WP-suppression-distraction now no longer yielded significance (see Table 3).

State rumination

For state rumination, current stress (WP and BP) was significantly associated with higher rumination. Furthermore, we found cognitive perspective change to be negatively associated with rumination on a WP-level. The use of suppression-distraction was positively associated with increased rumination on a WP- and BP-level. When adding self-efficacy to the model, we found a significant negative association on a WP-level while all of the other predictors remained the same. Adding ER-success also yielded significance on a WP- and BP-level. Further, we found cognitive-behavioral problem-solving to be positively associated with state rumination while WP-perspective change now no longer yielded significance. Adding self-efficacy and ER-success at the same time resulted in WP-self-efficacy, WP- and BP-ER-success yielding significance while WP-perspective change no longer yielded significance and WP-cognitive-behavioral problem-solving was now positively associated with state rumination (see Table 4).

10.3389/fpsyg.2024.1400223

TABLE 1 Mixed models investigating predictors of ER-success_t.

	Model_1	Model_2
(Intercept)	63.32***	57.82***
	(1.23)	(2.11)
SuccessEmotionRegulation_t-1	0.04**	0.03
	(0.02)	(0.02)
WP_StateRum_t	−6.15***	−5.49***
	(0.84)	(0.84)
BP_StateRum_t	−6.05***	−5.37**
	(1.75)	(1.71)
WP_Stress_t	−0.23***	−0.23***
	(0.02)	(0.02)
BP_Stress_t	−0.11	−0.14*
	(0.07)	(0.07)
WP_PerspectiveChange_t	13.78***	12.32***
	(2.10)	(2.10)
BP_PerspectiveChange_t	9.78*	6.98
	(4.41)	(4.43)
WP_CognitiveBehavioralProblemSolving_t	8.09***	7.59***
	(1.57)	(1.55)
BP_CognitiveBehavioralProblemSolving_t	3.26	3.28
	(3.92)	(3.85)
WP_SuppressionDistraction_t	−6.71***	−6.76***
	(1.65)	(1.63)
BP_SuppressionDistraction_t	−5.67	−5.10
	(4.19)	(4.11)
WP_BodySocial_t	5.83***	5.44***
	(1.56)	(1.54)
BP_BodySocial_t	11.50**	11.07**
	(4.37)	(4.30)
WP Self-efficacy_t		0.09***
		(0.02)
BP_Self-efficacy_t		0.14***
		(0.04)
AIC	10,534.71	10,504.79
BIC	10,626.71	10,607.02
Num. obs.	1,226	1,226
Num. subjects	144	144
Var: subjects (Intercept)	65.01	69.03
Var: subjects Stress	0.00	0.00
Cov: subjects (Intercept) Stress	0.25	0.15
Var: Residual	262.87	256.40

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

TABLE 2 Mixed models investigating predictors of self-efficacy_t.

	Model_1	Model_2
(Intercept)	34.65***	6.86
	(1.85)	(8.60)
Self-efficacy_t-1	0.25***	0.24***
	(0.03)	(0.03)
WP_StateRum_t	−6.94***	−5.54***
	(1.30)	(1.31)
BP_StateRum_t	−3.61	−0.96
	(2.89)	(2.91)
WP_Stress_t	0.03	0.08*
	(0.04)	(0.04)
BP_Stress_t	0.06	0.12
	(0.12)	(0.12)
WP_PerspectiveChange_t	16.17***	12.97***
	(3.25)	(3.27)
BP_PerspectiveChange_t	14.30	10.48
	(7.47)	(7.47)
WP_CognitiveBehavioralProblemSolving_t	5.55*	3.67
	(2.42)	(2.42)
BP_CognitiveBehavioralProblemSolving_t	4.23	1.68
	(6.59)	(6.51)
WP_SuppressionDistraction_t	0.25	1.80
	(2.55)	(2.54)
BP_SuppressionDistraction_t	−1.67	0.16
	(7.01)	(6.92)
WP_BodySocial_t	4.00	2.67
	(2.41)	(2.40)
BP_BodySocial_t	5.47	0.59
	(7.37)	(7.43)
WP_SuccessEmotionRegulation _t		0.23***
		(0.05)
BP_SuccessEmotionRegulation_t		0.43***
		(0.13)
AIC	1,1579.27	1,1554.66
BIC	1,1671.28	1,1656.89
Num. obs.	1,226	1,226
Num. subjects	144	144
Var: subjects (Intercept)	282.90	303.35
Var: subjects Stress_t-1	0.02	0.02
Cov: subjects (Intercept) Stress_t-1	−0.81	−1.20
Var: Residual	619.74	602.42

***$p < 0.001$, *$p < 0.05$.

Discussion

The aim of the current online diary study was to investigate the use of emotion regulation strategies (ERS) in everyday life and their associations with perceived ER-success, self-efficacy, stress and state rumination. For this, after an assessment of baseline data, participants completed an online diary for 2 weeks. After exploration of the data and the observation of clustered use of multiple ERS at the same

TABLE 3 Mixed models investigating predictors of stress_t.

	Model_1	Model_2	Model_3	Model_4
(Intercept)	35.28***	35.34***	36.03***	36.09***
	(1.53)	(1.76)	(1.55)	(1.82)
Stress_t-1	0.24***	0.24***	0.22***	0.22***
	(0.03)	(0.03)	(0.03)	(0.03)
WP_StateRum_t	12.26***	12.55***	9.43***	9.89***
	(0.97)	(0.98)	(0.99)	(0.99)
BP_StateRum_t	8.19***	7.82***	6.71***	5.75**
	(1.63)	(1.70)	(1.70)	(1.75)
WP_PerspectiveChange_t	−1.43	−1.60	2.96	2.37
	(2.61)	(2.61)	(2.56)	(2.54)
BP_PerspectiveChange_t	−10.30*	−11.39*	−8.55	−9.46
	(4.56)	(4.93)	(4.65)	(4.94)
WP_CognitiveBehavioralProblemSolving_t	11.18***	11.17***	12.93***	12.80***
	(1.91)	(1.90)	(1.85)	(1.83)
BP_CognitiveBehavioralProblemSolving_t	16.63***	15.94***	17.39***	16.81***
	(3.85)	(4.08)	(3.83)	(4.03)
WP_SuppressionDistraction_t	6.04**	5.80**	3.44	3.08
	(2.03)	(2.01)	(1.98)	(1.96)
BP_SuppressionDistraction_t	7.71	8.54	7.17	7.69
	(4.27)	(4.47)	(4.26)	(4.43)
WP_BodySocial_t	−2.93	−3.06	−0.84	−1.05
	(1.93)	(1.92)	(1.88)	(1.86)
BP_BodySocial_t	−0.74	−1.12	1.54	1.82
	(4.61)	(4.89)	(4.69)	(4.95)
WP_Self-efficacy_t		0.02		0.05*
		(0.02)		(0.02)
BP_Self-efficacy_t		0.05		0.08
		(0.04)		(0.05)
WP_SuccessEmotionRegulation_t			−0.32***	−0.32***
			(0.04)	(0.03)
BP_SuccessEmotionRegulation_t			−0.18*	−0.27**
			(0.08)	(0.09)
AIC	1,0967.92	1,0985.49	1,0896.75	1,0911.38
BIC	1,1049.71	1,1077.50	1,0988.76	1,1013.61
Num. obs.	1,226	1,226	1,226	1,226
Num. subjects	144	144	144	144
Var: subjects (Intercept)	160.12	266.16	176.63	308.02
Var: subjects Stress_t-1	0.04	0.07	0.04	0.06
Cov: subjects (Intercept) Stress_t-1	−1.92	−3.51	−1.97	−3.86
Var: Residual	387.38	375.32	360.59	347.34

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

measurement time points, we firstly performed an exploratory multilevel factor analysis and a principal component analysis and identified four factors: cognitive perspective change, cognitive-behavioral problem-solving, suppression-distraction, body-social ERS. Although those factors already reduced data complexity, they were still rarely used alone. This underlines the importance of investigating emotion-polyregulation, which assumes that multiple ERS are implemented simultaneously/sequentially (Ford et al., 2019). Going further, this questions ERS-differentiation: our results show that ERS which were originally defined as altering different stages of the emotion-generative

TABLE 4 Mixed Models investigating predictors of state rumination_t.

	Model_1	Model_2	Model_3	Model_4
(Intercept)	1.29***	1.31***	1.28***	1.32***
	(0.05)	(0.05)	(0.06)	(0.05)
StateRum_t-1	0.38***	0.37***	0.37***	0.36***
	(0.03)	(0.03)	(0.03)	(0.03)
WP_Stress_t	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
BP_Stress_t	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
WP_PerspectiveChange_t	−0.24**	−0.18*	−0.13	−0.10
	(0.07)	(0.07)	(0.07)	(0.07)
BP_PerspectiveChange_t	−0.12	−0.09	−0.01	−0.03
	(0.16)	(0.16)	(0.17)	(0.16)
WP_CognitiveBehavioralProblemSolving_t	0.07	0.08	0.11*	0.13*
	(0.05)	(0.05)	(0.05)	(0.05)
BP_CognitiveBehavioralProblemSolving_t	−0.26	−0.26	−0.24	−0.20
	(0.14)	(0.14)	(0.15)	(0.14)
WP_SuppressionDistraction_t	0.35***	0.35***	0.29***	0.29***
	(0.06)	(0.05)	(0.05)	(0.05)
BP_SuppressionDistraction_t	0.49***	0.49***	0.38*	0.42**
	(0.14)	(0.14)	(0.15)	(0.14)
WP_BodySocial_t	−0.01	0.00	0.03	0.04
	(0.05)	(0.05)	(0.05)	(0.05)
BP_BodySocial_t	0.15	0.16	0.27	0.25
	(0.16)	(0.16)	(0.17)	(0.15)
WP_Self-efficacy_t		−0.00***		−0.00***
		(0.00)		(0.00)
BP_Self-efficacy_t		−0.00		−0.00
		(0.00)		(0.00)
WP_SuccessEmotionRegulation_t			−0.01***	−0.01***
			(0.00)	(0.00)
BP_SuccessEmotionRegulation_t			−0.01**	−0.01**
			(0.00)	(0.00)
AIC	2,332.97	2,335.14	2,304.66	2,312.19
BIC	2,414.75	2,427.15	2,396.67	2,414.42
Num. obs.	1,226	1,226	1,226	1,226
Num. subjects	144	144	144	144
Var: subjects (Intercept)	0.13	0.14	0.22	0.13
Var: subjects Stress_t-1	0.00	0.00	0.00	0.00
Cov: subjects (Intercept) Stress_t-1	−0.00	−0.00	−0.00	−0.00
Var: Residual	0.30	0.29	0.28	0.28

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

process (Gross, 1998), differentiated in the literature but all associated with cognitive perspective change (self-compassion, acceptance, reframing and mindfulness), loaded on one factor. Furthermore, the use of these ERS is almost always accompanied by the use of cognitive-behavioral ERS (cognitive and behavioral problem-solving). This gives rise to the idea that all of them might be facets of cognitive and behavioral ER which aims for cognitive-behavioral flexibility, for instance by cognitive perspective change. We further fitted mixed models with time-lagged variables investigating ER-success, self-efficacy, stress and state rumination dependent on ERS-factors.

Our results show that increased ER-success was significantly associated with lower stress, lower rumination, an increased use of cognitive perspective change, cognitive-behavioral problem-solving and body-social ERS. These results support our initial hypothesis (I). Interestingly, people using suppression-distraction habitually more often do not perceive their ER as less successful, however, when people use it more often than their person-mean, ER-success is perceived as less successful.

When we fitted the same models for self-efficacy as dependent variable, higher WP-rumination was associated with lower self-efficacy, which was in line with our hypothesis, whereas, contrary to our expectation, stress did not have a corresponding impact (II). We only found WP-ERS associations, namely cognitive perspective change and cognitive-behavioral problem-solving being associated with increased self-efficacy. The latter did no longer yield significance when ER-success was included. This finding suggests that the effect of ERS on self-efficacy is moderated by how successful the ER is perceived. However, the concrete causal relationship between ER-success and self-efficacy is unclear and difficult to investigate. Please note that BP-effects of ER-success and self-efficacy only correlate moderately, $r(142) = 0.328$, $p < 0.001$; however, this might also be influenced by a different item format (Likert-scale vs. slider). Disentangling these complex associations and potentially causal mechanisms will be an interesting endeavor for future research.

We further found a significant positive association of current state rumination with stress. Also, an increased WP-use of suppression-distraction was associated with higher stress. Further, individuals who use cognitive perspective change ERS habitually more often had lower stress ratings. These findings are in line with our hypotheses (III) and easily integrated into our previous findings showing an increased ER-success for cognitive perspective change as compared to the use of suppression-distraction. However, the effect of cognitive perspective change and WP-suppression-distraction no longer yielded significance when ER-success was added to the model. Higher ER-success was then associated with lower stress indicating that ER-success is more relevant than the respective strategy that was applied. However, it is important to keep in mind that there was quite some overlap between the strategies and their success. That is, cognitive perspective change was associated with more ER-success and suppression-distraction with lower success. However, in order to understand which strategies are perceived as more helpful for whom and in which situations, it still is important to get a better understanding of the effects of the different regulation strategies. When self-efficacy was added to the model, it was positively associated with stress. Please note that this seemed to be an artifact due to multicollinearity as this effect diminished when ER-success was not added to the model. In this context it is important to note that we did not randomize the order of items but it remained the same throughout the assessment. This might have an impact on how the items are answered that cannot be controlled in our analysis. Studies that randomize the order of items would not need to identify the potential effects that the order has on the data or responses to the items, but rather control for them independently.

Interestingly, despite the fact that cognitive-behavioral problem-solving was positively associated with ER-success on the WP level, we found an increased use of cognitive-behavioral problem-solving (WP and BP) to be associated with increased stress. This was originally not hypothesized but could be explained by situational demands that make the ERS-use more likely under high/low stress. In case of higher stress, cognitive-behavioral problem-solving might be used more frequently, which is why it is associated with increased stress but also increased ER-success. The stress-dependent use of ERS, namely cognitive perspective change and cognitive-behavioral problem-solving, might be explained by situational constraints such as limited cognitive resources and different needs for resources (Aldao, 2013; Barrett et al., 2001; Blanke et al., 2022; Broderick, 2005; Dixon-Gordon et al., 2015; Sheppes, 2020; Sheppes et al., 2011, 2014). As previously noted, our design does not allow us to determine whether ERS are used in a stress-dependent manner or if stress influences their use.

Please note that according to our hypotheses, we only used data entries where participants reported to have used at least one ERS when fitting our models. Additionally, we excluded participants with fewer than six consecutive data entries from the analysis to avoid distortions from less reliable subjects and to estimate parameters based on a sufficiently large data set. However, these criteria were not pre-registered. Future studies should pre-register their analyses to ensure unbiased exclusion of data points.

One question arising from these results is the definition of ER-efficacy and the appropriate measurement. This could be either done using a direct measure of ratings of ER-success or the actual difference in stress. The measure used might impact the results concerning ERS-efficiency; however, as ERS are most probably selected by situational constraints, direct measures of ER-success provide important additional information.

Investigating state rumination, we found higher stress to be associated with higher rumination on a BP- and WP-level, which supported our hypothesis (IV). Furthermore, we found a WP-increased use of cognitive perspective change to be associated with decreased state rumination. This effect, however, also diminished when ER-success was added to the model which was negatively associated with state rumination. This further resulted in WP-cognitive-behavioral problem-solving to yield a significant positive association, which also supports the idea of a substantial amount of shared variance in the model. State-of-the-art psychotherapeutic interventions already propose stressor-intensity-dependent ERS-implementation and aim to tackle ruminative processes (Goldberg et al., 2019; Hayes et al., 2007; Mennin and Fresco, 2013; Segal et al., 2008; Watkins, 2018) as they have shown to play an important role in various psychopathologies (Aldao et al., 2010). Our findings underscore the importance of rumination in everyday-life also in healthy individuals. As habitual rumination is associated with increased risk for elevated stress and negative affect, which is associated with increases in state rumination, this results in a self-sustaining loop of ruminative inertia (Bean et al., 2020; Connolly and Alloy, 2017; Kircanski et al., 2018; Moberly and Watkins, 2008; Rosenbaum et al., 2022; Ruscio et al., 2015).

The most profound limitation of the current study is the non-experimental design and correlative data which does not allow the investigation of causal mechanisms. As participants could choose ERS freely, this includes complex temporal dynamic processes of ERS selection and implementation and the reciprocal impact of situational constraints which are not controlled for in our set-up. Future studies might investigate these associations using an experimental EMA-design where participants are instructed to use ERS dependent on their current stress. Like this, it is possible to directly evaluate the situation-specificity of ER because with our data, no systematic variation of stressor-intensity has been implemented. As experimental laboratory studies with emotion-inductions often differ in terms of various crucial aspects compared to ecologically valid Ecological Momentary Assessment (EMA) data (e.g., the controllability and importance of negative events), EMA is promising for investigating these associations and it would be beneficial to combine the best of both study-designs.

A very promising and only recently developed approach is the situated assessment method (SAM^2) (Dutriaux et al., 2023). It can be utilized to identify individual differences in habitual behavior and combined the evaluations of target behaviors and their situational influences to develop a comprehensive profile of an individual's behavior across various situations.

Another important point to note is the very unequal sex distribution which might potentially have caused biases or make our findings less generalizable to men. First evidence regarding this comes from studies investigating coping styles and everyday stressors finding higher stress and more emotion-focused coping styles in women (Almeida and Kessler, 1998; Matud, 2004). On the other hand, however, effect sizes are low (Matud, 2004) and other studies find no gender differences (Porter et al., 2000).

Nevertheless, our results underpin the idea of situation-specific ERS-use. The present study advances the insufficiently explored interplay between emotion regulation, stress, and state rumination while emphasizing the importance of cognitive emotion regulation strategies in overcoming momentary rumination. Additionally, the findings challenge the previously distinct categorizations of emotion regulation strategies in the literature. Future research should abstain from the artificial investigation of the use of single ERS and instead focus on the use of multiple ERS in order to improve psychological theories and treatments. The categorization of cognitive ERS should be reconsidered as they might be aspects of one construct.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee at the University Hospital and University of Tübingen. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

II-V: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. MV: Writing – review & editing, Writing – original draft, Investigation, Data curation. AK: Writing – review & editing, Writing – original draft. AF: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition. A-CE: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition. JR: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. DR: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We acknowledge support from the Open Access Publication Fund of the University of Tübingen.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1400223/full#supplementary-material

References

Aldao, A. (2013). The future of emotion regulation research: capturing context. *Perspect. Psychol. Sci.* 8, 155–172. doi: 10.1177/1745691612459518

Aldao, A., and Nolen-Hoeksema, S. (2013). One versus many: capturing the use of multiple emotion regulation strategies in response to an emotion-eliciting stimulus. *Cognit. Emot.* 27, 753–760. doi: 10.1080/02699931.2012.739998

Aldao, A., Nolen-Hoeksema, S., and Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clin. Psychol. Rev.* 30, 217–237.

Almeida, D. M., and Kessler, R. C. (1998). Everyday stressors and gender differences in daily distress. *J. Pers. Soc. Psychol.* 75, 670–680. doi: 10.1037//0022-3514.75.3.670

Arditte, K. A., Shaw, A. M., and Timpano, K. R. (2016). Repetitive negative thinking: a transdiagnostic correlate of affective disorders. *J. Soc. Clin. Psychol.* 35, 181–201. doi: 10.1521/jscp.2016.35.3.181

Bader, K., Hänny, C., Schäfer, V., Neuckel, A., and Kuhl, C. (2009). Childhood trauma questionnaire – psychometrische Eigenschaften einer deutschsprachigen version. *Z. Klin. Psychol. Psychother.* 38, 223–230. doi: 10.1026/1616-3443.38.4.223

Bakdash, J. Z., and Marusich, L. R. (2017). Repeated measures correlation. *Front. Psychol.* 8:456. doi: 10.3389/fpsyg.2017.00456

Barrett, L. F., Gross, J., Christensen, T. C., and Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation. *Cognit. Emot.* 15, 713–724. doi: 10.1080/02699930143000239

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67:1. doi: 10.18637/jss.v067.i01

Bean, C. A. L., Heggeness, L. F., Kalmbach, D. A., and Ciesla, J. A. (2020). Ruminative inertia and its association with current severity and lifetime course of depression. *Clin. Psychol. Sci.* 8, 1007–1016. doi: 10.1177/2167702620949174

Beck, A. T., Steer, R. A., and Brown, G. K. (1966). *Beck depression inventory (BDI).* TX: Psychological Corporation San Antonio.

Beck, A. T., Ward, C., Mendelson, M., Mock, J., and Erbaugh, J. J. A. G. P. (1961). Beck depression inventory (BDI). *Arch. Gen. Psychiatry*, 4, 561–571.

Bernstein, D. P., and Fink, L. (1988). Childhood trauma questionnaire: A retrospective self-report. San Antonio, TX: The Psychological Corporation.

Bernstein, D. P., Fink, L., Handelsman, L., and Foote, J. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *Am. J. Psychiatry* 151, 1132–1136. doi: 10.1037/t02080-000

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., et al. (2003). Development and validation of a brief screening version of the childhood trauma questionnaire. *Child Abuse Negl.* 27, 169–190. doi: 10.1016/s0145-2134(02)00541-0

Blanke, E. S., Bellingtier, J. A., Riediger, M., and Brose, A. (2022). When and how to regulate: everyday emotion-regulation strategy use and stressor intensity. *Affect. Sci.* 3, 81–92. doi: 10.1007/s42761-021-00087-1

Boemo, T., Nieto, I., Vazquez, C., and Sanchez-Lopez, A. (2022). Relations between emotion regulation strategies and affect in daily life: a systematic review and meta-analysis of studies using ecological momentary assessments. *Neurosci. Biobehav. Rev.* 139:104747. doi: 10.1016/j.neubiorev.2022.104747

Brans, K., Koval, P., Verduyn, P., Lim, Y. L., and Kuppens, P. (2013). The regulation of negative and positive affect in daily life. *Emotion* 13, 926–939. doi: 10.1037/a0032400

Broderick, P. C. (2005). Mindfulness and coping with dysphoric mood: contrasts with rumination and distraction. *Cogn. Ther. Res.* 29, 501–510. doi: 10.1007/s10608-005-3888-0

Burns, E. E., Fischer, S., Jackson, J. L., and Harding, H. G. (2012). Deficits in emotion regulation mediate the relationship between childhood abuse and later eating disorder symptoms. *Child Abuse Negl.* 36, 32–39. doi: 10.1016/j.chiabu.2011.08.005

Burns, E. E., Jackson, J. L., and Harding, H. G. (2010). Child maltreatment, emotion regulation, and posttraumatic stress: the impact of emotional abuse. *J. Aggress. Maltreat. Trauma* 19, 801–819. doi: 10.1080/10926771.2010.522947

Connolly, S. L., and Alloy, L. B. (2017). Rumination interacts with life stress to predict depressive symptoms: an ecological momentary assessment study. *Behav. Res. Ther.* 97, 86–95. doi: 10.1016/j.brat.2017.07.006

Diedrich, A., Hofmann, S. G., Cuijpers, P., and Berking, M. (2016). Self-compassion enhances the efficacy of explicit cognitive reappraisal as an emotion regulation strategy in individuals with major depressive disorder. *Behav. Res. Ther.* 82, 1–10. doi: 10.1016/j.brat.2016.04.003

Dixon-Gordon, K. L., Aldao, A., and De Los Reyes, A. (2015). Emotion regulation in context: examining the spontaneous use of strategies across emotional intensity and type of emotion. *Personal. Individ. Differ.* 86, 271–276. doi: 10.1016/j.paid.2015.06.011

Dutriaux, L., Clark, N. E., Papies, E. K., Scheepers, C., and Barsalou, L. W. (2023). The situated assessment method (SAM2): establishing individual differences in habitual behavior. *PLoS One* 18:e0286954. doi: 10.1371/journal.pone.0286954

Economou, A., Konsolaki, E., Kalsi, I. A., and Psychountaki, M. (2024). The Beck depression inventory-II in community-dwelling middle-aged and older Greeks: factor structure and demographic associations. *SAGE Open* 14:21582440241252264. doi: 10.1177/21582440241252264

Ehring, T., Zetsche, U., Weidacker, K., Wahl, K., Schönfeld, S., and Ehlers, A. (2011). The perseverative thinking questionnaire (PTQ): validation of a content-independent measure of repetitive negative thinking. *J. Behav. Ther. Exp. Psychiatry* 42, 225–232. doi: 10.1016/j.jbtep.2010.12.003

Enders, C. K., and Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychol. Methods* 12, 121–138. doi: 10.1037/1082-989X.12.2.121

Falkenström, F., Finkel, S., Sandell, R., Rubel, J. A., and Holmqvist, R. (2017). Dynamic models of individual change in psychotherapy process research. *J. Consult. Clin. Psychol.* 85, 537–549. doi: 10.1037/ccp0000203

Faro, A., and Pereira, C. R. (2020). Factor structure and gender invariance of the Beck depression inventory–second edition (BDI-II) in a community-dwelling sample of adults. *Health Psychol. Behav. Med.* 8, 16–31. doi: 10.1080/21642850.2020.1715222

Ford, B. Q., Gross, J. J., and Gruber, J. (2019). Broadening our field of view: the role of emotion polyregulation. *Emot. Rev.* 11, 197–208. doi: 10.1177/1754073919850314

Goldberg, S. B., Tucker, R. P., Greene, P. A., Davidson, R. J., Kearney, D. J., and Simpson, T. L. (2019). Mindfulness-based cognitive therapy for the treatment of current depressive symptoms: a meta-analysis. *Cogn. Behav. Ther.* 48, 445–462. doi: 10.1080/16506073.2018.1556330

Gomes-Oliveira, M. H., Gorenstein, C., Lotufo Neto, F., Andrade, L. H., and Wang, Y. P. (2012). Validation of the Brazilian Portuguese version of the Beck depression inventory-II in a community sample. *Braz. J. Psychiatr.* 34, 389–394. doi: 10.1016/j.rbp.2012.03.005

Grob, A., and Horowitz, D. (2014). Fragebogen zur Erhebung der Emotionsregulation bei Erwachsenen. Bern: Hans Huber.

Gross, J. J. (1998). The emerging field of emotion regulation: an integrative review. *Rev. Gen. Psychol.* 2, 271–299. doi: 10.1037/1089-2680.2.3.271

Gross, J. J. (2014). "Emotion regulation: Conceptual and empirical foundations" in Handbook of emotion regulation. *2nd* ed (New York, NY: The Guilford Press), 3–20.

Gross, J. J. (2015). Emotion regulation: current status and future prospects. *Psychol. Inq.* 26, 1–26. doi: 10.1080/1047840X.2014.940781

Hautzinger, M., Bailer, M., and Keller, F. (1994). Beck- Depressions-Inventar (BDI): Bearbeitung der deutschen Ausgabe. Bern: Hans Huber.

Hautzinger, M., Keller, F., and Kühner, C. (2009). Beck - Depressions - Inventar (BDI-II) Deutsche Ausgabe, (2. Auflage) Pearson Assessment.

Hayes, S. C., Strosahl, K., and Wilson, K. G. (2007). Akzeptanz und Commitment-Therapie: ein erlebnisorientierter Ansatz zur Verhaltensänderung. *II. Auflage* Edn. München: CIP-Medien.

Hoffman, L., and Stawski, R. S. (2009). Persons as contexts: evaluating between-person and within-person effects in longitudinal analysis. *Res. Hum. Dev.* 6, 97–120. doi: 10.1080/15427600902911189

Houben, M., Kalokerinos, E. K., Koval, P., Erbas, Y., Mitchell, J., Pe, M., et al. (2023). The association between the selection and effectiveness of emotion-regulation strategies and psychopathological features: A daily life study. *Clin. Psychol. Sci.* 13:21677026231203664. doi: 10.1177/21677026231203662

Huh, H. J., Kim, K. H., Lee, H.-K., and Chae, J.-H. (2017). The relationship between childhood trauma and the severity of adulthood depression and anxiety symptoms in a clinical sample: the mediating role of cognitive emotion regulation strategies. *J. Affect. Disord.* 213, 44–50. doi: 10.1016/j.jad.2017.02.009

IBM Corp. (2019). *IBM SPSS Statistics for Windows (Version 27.0)* [Computer software]. IBM Corp.

Just, N., and Alloy, L. B. (1997). The response styles theory of depression: tests and an extension of the theory. *J. Abnorm. Psychol.* 106, 221–229. doi: 10.1037/0021-843X.106.2.221

Kalokerinos, E. K., Résibois, M., Verduyn, P., and Kuppens, P. (2017). The temporal deployment of emotion regulation strategies during negative emotional episodes. *Emotion* 17, 450–458. doi: 10.1037/emo0000248

Kasch, K. L., Klein, D. N., and Lara, M. E. (2001). A construct validation study of the response styles questionnaire rumination scale in participants with a recent-onset major depressive episode. *Psychol. Assess.* 13, 375–383. doi: 10.1037/1040-3590.13.3.375

Kircanski, K., Thompson, R. J., Sorenson, J., Sherdell, L., and Gotlib, I. H. (2018). The everyday dynamics of rumination and worry: precipitant events and affective consequences. *Cognit. Emot.* 32, 1424–1436. doi: 10.1080/02699931.2017.1278679

Klinitzke, G., Romppel, M., Häuser, W., Brähler, E., and Glaesmer, H. (2012). Die deutsche Version des Childhood Trauma Questionnaire (CTQ) – psychometrische Eigenschaften in einer bevölkerungsrepräsentativen Stichprobe. *PPmP* 62, 47–51. doi: 10.1055/s-0031-1295495

Kobylińska, D., and Kusev, P. (2019). Flexible emotion regulation: how situational demands and individual differences influence the effectiveness of regulatory strategies. *Front. Psychol.* 10:72. doi: 10.3389/fpsyg.2019.00072

Koval, P., Butler, E. A., Hollenstein, T., Lanteigne, D., and Kuppens, P. (2015). Emotion regulation and the temporal dynamics of emotions: effects of cognitive reappraisal and expressive suppression on emotional inertia. *Cognit. Emot.* 29, 831–851. doi: 10.1080/02699931.2014.948388

Koval, P., and Kalokerinos, E. K. (2024). Daily diaries and ecological momentary assessment. In *Handbook of emotion regulation*. eds. J. J. Gross and B. Q. Ford, 3rd ed. (The Guilford Press), pp. 31–39.

Koval, P., Kalokerinos, E. K., Verduyn, P., and Greiff, S. (2020). Introduction to the special issue: capturing the dynamics of emotion and emotion regulation in daily life with ambulatory assessment. *Eur. J. Psychol. Assess.* 36, 433–436. doi: 10.1027/1015-5759/a000599

Koval, P., Kuppens, P., Allen, N. B., and Sheeber, L. (2012). Getting stuck in depression: the roles of rumination and emotional inertia. *Cognit. Emot.* 26, 1412–1427. doi: 10.1080/02699931.2012.667392

Kühner, C., Bürger, C., Keller, F., and Hautzinger, M. (2007). Reliabilität und validität des revidierten beck-depressionsinventars (BDI-II): befunde aus deutschsprachigen stichproben. *Nervenarzt* 78, 651–656. doi: 10.1007/s00115-006-2098-7

Kumar, G., Steer, R. A., Teitelman, K. B., and Villacis, L. (2002). Effectiveness of Beck depression inventory–ii subscales in screening for major depressive disorders in adolescent psychiatric inpatients. *Assessment* 9, 164–170. doi: 10.1177/10791102009002007

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Matthews, M., Webb, T. L., Shafir, R., Snow, M., and Sheppes, G. (2021). Identifying the determinants of emotion regulation choice: a systematic review with meta-analysis. *Cognit. Emot.* 35, 1056–1084. doi: 10.1080/02699931.2021.1945538

Matud, M. P. (2004). Gender differences in stress and coping styles. *Personal. Individ. Differ.* 37, 1401–1415. doi: 10.1016/j.paid.2004.01.010

McLaughlin, K. A., and Nolen-Hoeksema, S. (2011). Rumination as a transdiagnostic factor in depression and anxiety. *Behav. Res. Ther.* 49, 186–193. doi: 10.1016/j.brat.2010.12.006

Mennin, D. S., and Fresco, D. M. (2013). Emotion regulation therapy. In *Handbook of emotion regulation*. eds. J. J. Gross and B. Q. Ford, 3rd ed. (The Guilford Press), pp. 383–390.

Moberly, N. J., and Watkins, E. R. (2006). Processing mode influences the relationship between trait rumination and emotional vulnerability. *Behav. Ther.* 37, 281–291. doi: 10.1016/j.beth.2006.02.003

Moberly, N. J., and Watkins, E. R. (2008). Ruminative self-focus, negative life events, and negative affect. *Behav. Res. Ther.* 46, 1034–1039. doi: 10.1016/j.brat.2008.06.004

Muthén, L. K., and Muthén, B. O. (2017). Mplus: Statistical analysis with latent variables: User's guide (Version 8). Los Angeles, CA: Muthén et Muthén.

Nolen-Hoeksema, S., Larson, J., and Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *J. Pers. Soc. Psychol.* 77, 1061–1072. doi: 10.1037/0022-3514.77.5.1061

Nolen-Hoeksema, S., and Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: the 1989 Loma Prieta earthquake. *J. Pers. Soc. Psychol.* 61, 115–121. doi: 10.1037/0022-3514.61.1.115

Nolen-Hoeksema, S., Parker, L. E., and Larson, J. (1994). Ruminative coping with depressed mood following loss. *J. Pers. Soc. Psychol.* 67, 92–104. doi: 10.1037/0022-3514.67.1.92

Porter, L. S., Marco, C. A., Schwartz, J. E., Neale, J. M., Shiffman, S., and Stone, A. A. (2000). Gender differences in coping: a comparison of trait and momentary assessments. *J. Soc. Clin. Psychol.* 19, 480–498. doi: 10.1521/jscp.2000.19.4.480

R Core Team (2020). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Robinson, M. S., and Alloy, L. B. (2003). Negative cognitive styles and stress-reactive rumination interact to predict depression: a prospective study. *Cogn. Ther. Res.* 27, 275–291. doi: 10.1023/A:1023914416469

Roelofs, J., van Breukelen, G., de Graaf, L. E., Beck, A. T., Arntz, A., and Huibers, M. J. H. (2013). Norms for the Beck depression inventory (BDI-II) in a large Dutch community sample. *J. Psychopathol. Behav. Assess.* 35, 93–98. doi: 10.1007/s10862-012-9309-2

Rosenbaum, D., Hilsendegen, P., Thomas, M., Haeussinger, F. B., Metzger, F. G., Nuerk, H.-C., et al. (2018a). Cortical hemodynamic changes during the Trier social stress test: an fNIRS study. *NeuroImage* 171, 107–115. doi: 10.1016/j.neuroimage.2017.12.061

Rosenbaum, D., Int-Veen, I., Laicher, H., Torka, F., Kroczek, A., Rubel, J., et al. (2021). Insights from a laboratory and naturalistic investigation on stress, rumination and frontal brain functioning in MDD: an fNIRS study. *Neurobiol. Stress* 15:100344. doi: 10.1016/j.ynstr.2021.100344

Rosenbaum, D., Int-Veen, I., Rubel, J., Laicher, H., Kroczek, A., Lawyer, G., et al. (2022). Associations of different emotion regulation strategies with coping-efficacy, rumination and stress. *Cognit. Ther. Res.* 46, 889–901. doi: 10.1007/s10608-022-10307-y

Rosenbaum, D., Thomas, M., Hilsendegen, P., Metzger, F. G., Haeussinger, F. B., Nuerk, H.-C., et al. (2018b). Stress-related dysfunction of the right inferior frontal cortex in high ruminators: an fNIRS study. *NeuroImage Clin.* 18, 510–517. doi: 10.1016/j.nicl.2018.02.022

RStudio Team (2021). RStudio: Integrated development environment for R. Boston, MA: RStudio, PBC.

Ruscio, A. M., Gentes, E. L., Jones, J. D., Hallion, L. S., Coleman, E. S., and Swendsen, J. (2015). Rumination predicts heightened responding to stressful life events in major depressive disorder and generalized anxiety disorder. *J. Abnorm. Psychol.* 124, 17–26. doi: 10.1037/abn0000025

Scher, C. D., Stein, M. B., Asmundson, G. J. G., McCreary, D. R., and Forde, D. R. (2001). The childhood trauma questionnaire in a community sample: psychometric properties and normative data. *J. Trauma. Stress.* 14, 843–857. doi: 10.1023/A:1013058625719

Segal, Z. V., Williams, J. M. G., and Teasdale, J. D. (2008). Die achtsamkeitsbasierte Kognitive Therapie der Depression: ein neuer Ansatz zur Rückfallprävention. Tübingen: DGVT Verl.

Sheppes, G. (2020). Transcending the "good & bad" and "here & now" in emotion regulation: costs and benefits of strategies across regulatory stages. *Adv. Exp. Soc. Psychol.* 61, 185–236. doi: 10.1016/bs.aesp.2019.09.003

Sheppes, G., Scheibe, S., Suri, G., and Gross, J. J. (2011). Emotion-regulation choice. *Psychol. Sci.* 22, 1391–1396. doi: 10.1177/0956797611418350

Sheppes, G., Scheibe, S., Suri, G., Radu, P., Blechert, J., and Gross, J. J. (2014). Emotion regulation choice: a conceptual framework and supporting evidence. *J. Exp. Psychol. Gen.* 143, 163–181. doi: 10.1037/a0030831

Svaldi, J., Griepenstroh, J., Tuschen-Caffier, B., and Ehring, T. (2012). Emotion regulation deficits in eating disorders: a marker of eating pathology or general psychopathology? *Psychiatry Res.* 197, 103–111. doi: 10.1016/j.psychres.2011.11.009

Szkodny, L. E., and Newman, M. G. (2019). Delineating characteristics of maladaptive repetitive thought: development and preliminary validation of the perseverative cognitions questionnaire. *Assessment* 26, 1084–1104. doi: 10.1177/1073191117698753

Tamir, M., Vishkin, A., and Gutentag, T. (2020). Emotion regulation is motivated. *Emotion* 20, 115–119. doi: 10.1037/emo0000635

Teismann, T., Hanning, S., Brachel, R.Von, and Willutzki, U. (2012). Kognitive Verhaltenstherapie depressiven Grübelns. Berlin: Springer-Verlag.

Troy, A. S., Shallcross, A. J., and Mauss, I. B. (2013). A person-by-situation approach to emotion regulation: cognitive reappraisal can either help or hurt, depending on the context. *Psychol. Sci.* 24, 2505–2514. doi: 10.1177/0956797613496434

Wahl, K., Ertle, A., Bohne, A., Zurowski, B., and Kordon, A. (2011). Relations between a ruminative thinking style and obsessive–compulsive symptoms in non-clinical samples. *Anxiety Stress Coping* 24, 217–225. doi: 10.1080/10615806.2010.482985

Wang, Y.-P., and Gorenstein, C. (2013). Psychometric properties of the Beck depression inventory-II: a comprehensive review. *Rev. Bras. Psiquiatr.* 35, 416–431. doi: 10.1590/1516-4446-2012-1048

Watkins, E. (2018). Rumination-focused cognitive-behavioral therapy for depression. London: Guilford.

Webb, T. L., Miles, E., and Sheeran, P. (2012). Dealing with feeling: a meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychol. Bull.* 138, 775–808. doi: 10.1037/a0027600

Wingenfeld, K., Spitzer, C., Mensebach, C., Grabe, H., Hill, A., Gast, U., et al. (2010). Die deutsche version des childhood trauma questionnaire (CTQ): erste befunde zu den psychometrischen kennwerten. *PPmP* 60, 442–450. doi: 10.1055/s-0030-1247564

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 13 December 2024
DOI 10.3389/fpsyg.2024.1440425

OPEN ACCESS

EDITED BY
Javier Marín-Morales,
Universitat Politècnica de València, Spain

REVIEWED BY
Xindong Ye,
Wenzhou University, China
Maria Eleonora Minissi,
Polytechnic University of Valencia, Spain

*CORRESPONDENCE
Sharanya Lal
✉ s.lal@utwente.nl

RECEIVED 29 May 2024
ACCEPTED 08 November 2024
PUBLISHED 13 December 2024

CITATION
Lal S, Eysink THS, Gijlers HA, Veldkamp BP, Steinrücke J and Verwey WB (2024) Explicit metrics for implicit emotions: investigating physiological and gaze indices of learner emotions.
Front. Psychol. 15:1440425.
doi: 10.3389/fpsyg.2024.1440425

Explicit metrics for implicit emotions: investigating physiological and gaze indices of learner emotions

Sharanya Lal*, Tessa H. S. Eysink, Hannie A. Gijlers, Bernard P. Veldkamp, Johannes Steinrücke and Willem B. Verwey

Depertent of Learning, Data-Analytics and Technology, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands

Learning experiences are intertwined with emotions, which in turn have a significant effect on learning outcomes. Therefore, digital learning environments can benefit from taking the emotional state of the learner into account. To do so, the first step is real-time emotion detection which is made possible by sensors that can continuously collect physiological and eye-tracking data. In this paper, we aimed to find features derived from skin conductance, skin temperature, and eye movements that could be used as indicators of learner emotions. Forty-four university students completed different math related tasks during which sensor data and self-reported data on the learner's emotional state were collected. Results indicate that skin conductance response peak count, tonic skin conductance, fixation count, duration and dispersion, saccade count, duration and amplitude, and blink count and duration may be used to distinguish between different emotions. These features may be used to make learning environments more emotionally aware.

KEYWORDS

emotion recognition, learner emotions, physiological signals, eye-tracking, sensors, wearables

1 Introduction

Learning experiences are greatly influenced by the emotions of a learner (Pekrun, 2006). Therefore, it is pertinent that educators, designers, and researchers consider a learner's emotions while creating learning systems that offer personalized support. For this to be possible, the first step is to be able to perceive a learner's emotions and ideally in a continuous and non-obtrusive manner. This is what we address in the present paper.

1.1 Emotions in learning environments

There is a large body of work on the interaction of learner emotion and learning. For example, Csikszentmihalyi's (1990) seminal work on 'flow' suggests that optimal learning experiences occur when individuals are in such a state of concentration that they lose track of time. This is often accompanied by deep enjoyment and happiness. An example is Pekrun et al.'s (2017) longitudinal study on the development of mathematical competencies of adolescents (grades 5 to 10) that found that enjoyment and pride positively predicted subsequent annual assessment scores; the converse was found to be true for anger, anxiety, shame, hopelessness and boredom. Boredom in particular has been found to be persistent in learning situations of durations ranging from 30 to 75 min and is associated with 'gaming the system' (in which students simply manipulate the system to succeed at a task instead of actually learning the content) (Baker et al., 2010) and poor learning (Baker et al., 2010; Craig et al., 2004). According to D'Mello and Graesser (2012), students in a state of flow experience confusion when faced with an obstacle to their goals, but their state of flow is restored if they are able to solve their problem. If they cannot, this confusion makes way for frustration and then boredom. So, while one cannot assert that optimal learning experiences consist solely of positive emotions (such as deep enjoyment and happiness), current research suggests that persistent negative emotions do affect learning negatively. Therefore, it is expected that learning systems that detect these emotions in order to adapt their support will provide optimal learning experiences.

1.2 Emotion detection

Research into emotions has traditionally used self-reported data (Wu et al., 2016). Apart from the obvious subjective nature of self-reports, this approach is also limited by the temporal mismatch between when an emotion is experienced and its corresponding data are collected (Yadegaridehkordi et al., 2019). Moreover, if one's aim is to develop a system that can detect and adapt to emotions, it is impractical to constantly interrupt the learning process to ask the learner for input. This calls for objective, time-specific and unobtrusive data collection. Wearable and portable sensor technology today makes this possible because emotions are accompanied by physiological and behavioral responses. For example, people can find themselves with sweaty palms or a racing heart when extremely anxious. Sometimes people find themselves "wide-eyed" with surprise or shaking with fear or anger. Physiological and movement sensors can detect these signals, and using the appropriate techniques, one may make inferences about the associated emotion.

1.3 Sensor data in educational and emotion research

In their review of physiology-based mobile educational systems, Hernández-Cuevas and Crawford (2021) found that one of the most used measures, after eye-tracking, was heart rate (i.e., the number of heart beats per minute). In line with this, Liu et al. (2022) systematically reviewed learning analytics based on wearable devices, examining 120 articles published between 2011 and 2021, and found that heart rate and skin conductance (i.e., the level of perspiration in response to an emotional stimulus) were two of the most widely used sensor data. Heart rate has been included in several studies such as ones that investigated measures of mental workload (Sharma et al., 2020), student interaction (Darnell and Krieg, 2019) and cognitive load (Larmuseau et al., 2020). Among other things, skin conductance has been studied to profile sympathetic arousal of students during a physics class (Pijeira-Díaz et al., 2018), identify momentary student engagement in an afterschool program (Lee et al., 2019) and measure mental workload (Sharma et al., 2020). In emotion research specifically, skin conductance and heart rate have been investigated in the context of fatigue and drowsiness (Adão Martins et al., 2021). In the field of emotions during learning, some studies have found that skin conductance reflected stress (Brouwer et al., 2017), emotional arousal (i.e., the strength of an emotional state) (Jindrová et al., 2020), anxiety (Harley et al., 2019; Meer et al., 2016) and shame (Harley et al., 2019). In more recent preliminary explorations of emotions during parent–child learning activities, Avelar et al. (2022) and Shaby and Bokhove (2023) found that skin conductance could be used to discern different emotions. In their meta-study on test-anxiety and measures of physiological arousal, Roos et al. (2021) found that both skin conductance and heart rate significantly increased with self-reported test anxiety. However, there is no clear consensus yet on how exactly these signals vary with different emotions or how much variance they can explain. For example, Van Bruinessen et al. (2016) found no significant relationship between self-reported anxiety and skin conductance. In another study, Ritz et al. (2005) investigated physiological response to the viewing of pictures from IAPS (International Affective Picture System) and found that while heart rate significantly increased for both negative and positive emotions, there were no considerable changes in skin conductance. Nevertheless, a review by Ba and Hu (2023) showed that skin conductance and heart rate were two of the most studied measures of autonomic nervous system activity associated with emotions, and available evidence suggests that there is value in further exploration.

However, Kreibig (2010) warned that progress in research was hindered by the "exclusive use of convenience measures such as HR [heart rate] and electrodermal activity, as sole indicators of the activation state of the organism" (p. 409) and that it is essential to select more measures to determine patterns. Skin temperature is one such measure (Adão Martins et al., 2021; Noroozi et al., 2020). In their study with female undergraduate students, Rimm-Kaufman and Kagan (1996) found that hand skin temperature increased while watching film clips designed to induce happy affect and reduced when asked threatening personal questions. In a similar vein, McFarland (1985) found that music that was perceived as inducing negative emotions stopped an increase and perpetuated a decrease in skin temperature; calming music had the opposite effect. On the other hand, Jang et al. (2015) found that skin temperature increased with boredom (a negative emotion). Lal et al.'s (2021) study physiological correlates of learner emotions during different programming tasks suggested the same. Meanwhile, Jang et al. (2019) investigated the reliability of skin temperature as a response to different emotions and found it to be an unreliable indicator. Mixed results from past studies warrant further investigation into skin temperature as an indicator of emotions.

A relatively new approach in emotion research is tracking eye movements (Lim et al., 2020). Eye-tracking for emotion detection has usually been used in combination with physiological signals and there is mounting evidence that this is indeed useful (Lim et al., 2020). An example of this is Aracena et al.'s (2015) use of neural networks on multimodal data that included blinks and saccades (quick eye movements between fixations) to differentiate between negative, neutral and positive emotions. Other examples of successful use of eye-tracking include the use of gaze features for emotion recognition in patients with mesial temporal lobe epilepsy (Gomez-Ibañez et al., 2014) and in individuals in the autism spectrum (Tsang, 2018). While a wide variety of features has been used in the past, studies do not concur on which are the most effective for emotion recognition (Lim et al., 2020), and hence this warrants further research.

1.4 Present study

As outlined earlier, despite extensive research on skin conductance and heart rate as indicators of emotions, past results are varied. On the other hand, skin temperature despite being an easily accessible physiological measure that could be used as an emotional indicator, has rarely been investigated with respect to learner emotions (Noroozi et al., 2020). Moreover, eye-tracking has only recently been included in emotion detection and here too, findings are inconclusive. More interestingly, Noroozi et al. (2020) found that out of the 207 publications included in their systematic review of multimodal metrics to capture the learning process, only 15 included the emotional aspect of learning. This imbalance in past literature underscores the need for further research that focuses on (under represented) indicators of emotions, specifically of learners. In the present study, we investigated skin conductance, heart rate, skin temperature, and eye movement metrics as indices of learners' emotions. To this end, we adopted a dimensional approach to emotions based on Russell's (1980) widely accepted circumplex model of emotions, which posits that emotions may be represented along two orthogonal dimensions, arousal and valence. Emotional arousal may be defined as the activation level or strength of an emotion, while emotional valence is its hedonic nature (Pekrun, 2006; Russell et al., 1989; Thayer, 1967; Thayer, 1978). Consequently, emotions fall into any one of the four quadrants determined by the axes arousal and valence – high arousal-negative valence (for example, frustration), high arousal-positive valence (for example, happiness), low arousal-positive valence (for example, calmness) and low arousal-negative valence (for example, boredom). For a visual representation of different emotions on a two-dimensional scale, see Figure 6 of Russell (1980). Our research was motivated by the need to develop emotionally aware systems that can encourage positive and reduce persistent negative emotions during learning. Therefore, in the present study we investigated relevant (and unobtrusive) indices of the four emotional quadrants. Learner emotions were indexed by self-reported arousal and valence. In line with the literature cited earlier, we investigated the prospects of using the following three measures of physiological arousal – skin conductance, heart rate and skin temperature, and features derived from all three events that take place during eye-movement (Hessels et al., 2018) – blinks, saccades and fixations. It is important to note that the selection of measures was based also on the possibility of using them in real-world classrooms. The study was quasi-experimental and involved data that was collected at regular intervals during multiple math-related tasks in a counterbalanced set-up.

2 Methods

2.1 Participants

Participants consisted of 44 (32 females and 12 males, 18–26 years old, M_{age} = 20.09 years, *sd* = 1.89, 40 right-handed and 2 left-handed) bachelor's students from the Faculty of Behavioural, Management and Social sciences at the University of Twente (the Netherlands). The sample consisted of persons of 12 nationalities, namely: Bulgarian (n = 1), Brazilian (n = 1), Chinese (n = 1), Croatian (n = 1), French (n = 1), German (n = 21), Greek-German (n = 1), Malaysian (n = 1), Dutch (n = 11), Dutch-German (n = 1), Dutch-Ukrainian (n = 1) and Romanian (n = 3). All participants had at least working knowledge of English. Participants were recruited through an online participant management system. Participation was voluntary in exchange for 2 study credits. All participants had provided informed consent to participate in the study, which included the collection of demographic, physiological, eye-tracking and self-reported data.

2.2 Materials

2.2.1 Tasks and baseline stimulus

Participants engaged with three different math related tasks. All tasks were designed such that they could be done with just one hand. This was to mitigate motion artefacts in signal data from the wearables used in the study. The first task henceforth called the *shape matching task*, utilized an elementary school level math simulation called "Shapes Matching: Scored" (Lindenmuth, n.d.; Figure 1). In each round, a two-dimensional shape was displayed at the top of the screen. Participants were required to scan the row of shapes below the presented one and click on the matching shape. Instructions included no indication of how long the task would last and all points scored were inconsequential. While there were indefinite rounds in the task, participants completed only as many as was possible in 12 min. The second task was a high school level *coordinate geometry puzzle* (Figure 2). Participants were required to move the cursor to move the point (marked by the orange arrow) along the coordinate plane in order to find "the rule that governs the shape of the point." The rule was that the point changed shape based on whether it lays inside, on, or outside a parabola. Instructions stated that the activity would only end when participants solved the puzzle (to win €20), or time ran out (with no indication of when exactly that would be). However in reality, this task too ended after 12 min. Participants thus had an indefinite number of attempts at solving the puzzle. Irrespective of what answer the participant gave, the researcher told them they were wrong, essentially making this a frustrating activity. The third task was online (Tetris, 1985). Tetris is a video game in which players attempt to complete rows of a grid by arranging differently shaped tetrominoes that fall onto the playing field. The game has been found to bring players into a positive emotion of effortless attention (Harmat et al., 2015). It is not only widely popular among gamers and recreational mathematics enthusiasts, but is also a widely studied game (Mayer, 2014), drawing interest from psychologists, mathematicians, and computer scientists (e.g., Chanel et al., 2008; Maier et al., 2019; Tsuruda, 2010). Tetris (and its variations) has been found to improve algorithmic thinking and spatial skills and is often used to teach geometrical concepts such as rotation, translation and reflection (Clements et al., 1997; Sims and Mayer, 2001; Williams and Bright, 1991). In an attempt to foster engagement in this task, participants were offered prize money based on their Tetris score. Participants won €5 for the first 20,000 points in Tetris. For every 10,000 points after that, they won €1. If they made it to the leader board, they won additional money – up to €5 – depending on their rank. (Twenty-three participants won on average €6.52 (*sd* = 3.46), with the highest prize money being €22). Participants knew they had 12 min to play and were free to restart the game any number of times within that period.

In an effort to bring participants to more or less the same starting point in terms of their emotional state, participants were instructed to

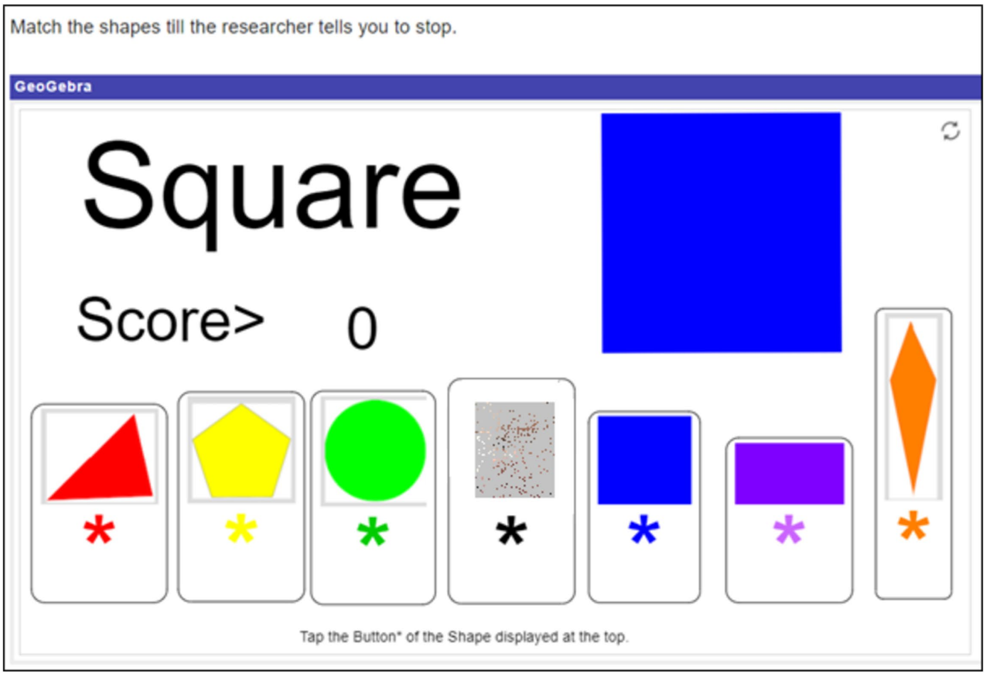


FIGURE 1
Created with GeoGebra®, by Lindenmuth https://www.geogebra.org/m/rvz58cma#material/eMCXcErd. In this instance, clicking on the blue square in the row would increase the score by 1 point.

"Sit still. Relax and clear your mind while watching this relaxing video" of an underwater scene with calm instrumental music (4 min). The decision to not use a math-related baseline was a conscious one – the baseline was intended to elicit minimal stimulation without priming participants to the math-related nature of the study. The use of this video is a deviation from the traditional 'resting baseline' in which participants usually do not engage in any activity and stay in a relaxed or 'resting' state. Research finds that engaging in an activity that requires minimal cognitive effort minimizes intrusive thoughts and produces results equal to or better than resting baseline conditions (Jennings et al., 1992). In fact, Piferi et al. (2000) found that watching a relaxing aquatic video produced lower (cardiovascular) baseline levels than traditional methods.

The tasks, baseline stimulus and self-reports (described below) were finalized after three rounds of iterative pilot testing involving a total of 13 pilot participants. Based on researcher observations and participant responses, we expected participants to report low arousal positive emotions while watching the baseline video, low arousal negative emotions over time while matching shapes, high arousal negative emotions over time doing the coordinate geometry puzzle, and high arousal positive emotions playing Tetris.

2.2.2 Measures and instrumentation

Self-reported measures were collected using the Affect Grid (Russell et al., 1989), a 1-item scale of emotions along the two dimensions, emotional arousal (Cronbach's $\alpha = 0.81$) and valence (Cronbach's $\alpha = 0.79$), both of which could range from values 1 to 9. The affect grid was selected because it is a quick capture tool and is appropriate for repeated measurements (Killgore, 1998). Respondents check a square on the 9×9 grid to indicate their emotion along two dimensions, valence (along the X-axis) and arousal (along the Y-axis). The mid-point of the grid is (5,5) which denotes neutral valence or arousal. High arousal-negative valence, high arousal-positive valence, low arousal-positive valence and low arousal-negative valence emotions are marked in the first, second, third and fourth quadrants, respectively. An open-ended fill-in-the-blank statement "I am feeling __" was used to verify that the participant had in fact thought through the filling of the affect grid. This self-report was administered every 4 min.

Four physiological measures were assessed. Skin conductance and heart rate were measured using the Shimmer3 GSR+, a biosensing unit with an average sampling frequency of 128 Hz. Skin conductance was measured between two stainless steel electrodes attached to the palmar region of two fingers. Heart rate was derived from a photoplethysmogram signal collected by a pulse probe clipped to the ear. Peripheral skin temperature was measured with the infrared thermopile sensor (sampling rate of 4 Hz) of the Empatica E4 biosensing wristband. Eye-tracking was done using Tobii Fusion Pro, a screen-based eye-tracker with sampling frequencies up to 250 Hz. Its camera was attached horizontally at the bottom of the computer screen the participants' tasks were displayed on.

Several supporting tools were used in the study. The Tobii Fusion Pro and the Shimmer3 GSR+ were configured using Tobii's eye-tracking manager and Shimmer's 'Consensys' software. These were synchronized on iMotions, an integrative software platform for biometric research. iMotions was used to time and present the tasks and instructions, and collect, visualize, and pre-process data from the two devices mentioned above. Calibration of the eye-tracker for each participant was also done on iMotions using a 9-point calibration slide; light calibration was done using a single grey screen. Skin temperature data were streamed to Empatica's cloud-based repository via an android application set up on a mobile phone which in turn was connected via Bluetooth to the Empatica E4. Other pieces of equipment used were a 24 inch AOC G2460PF computer monitor

10.3389/fpsyg.2024.1440425

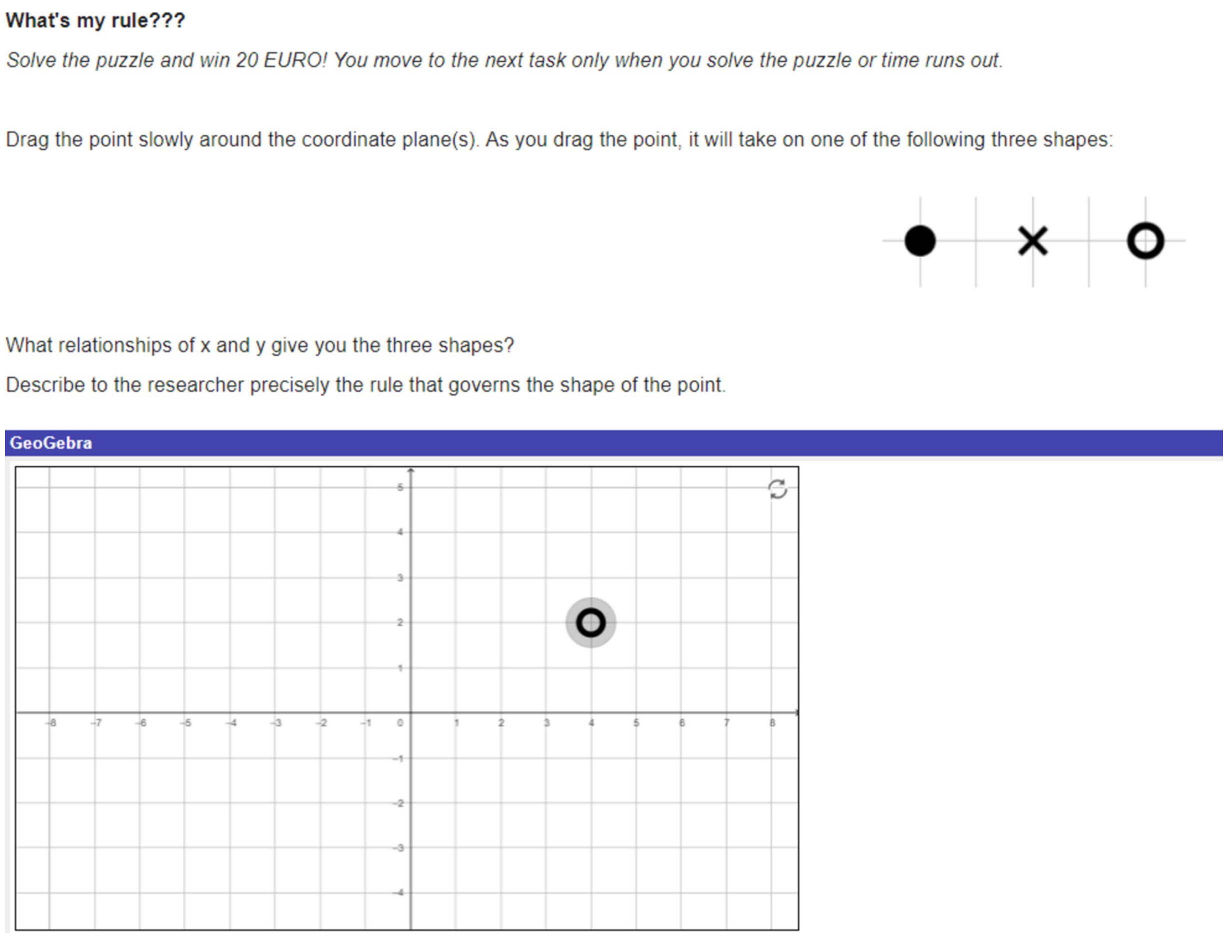


FIGURE 2
Created with GeoGebra®.

(refresh rate 144 Hz) used as the participant's primary screen, a Microsoft Surface Pro touchscreen tablet used to record participants' self-reports, a wired mouse for the use of the participants and two HP Elitebook laptops (64-bit operating system, Intel(R) Core(TM) i5-8250U processor with CPU @ 1.60GHz) – one that served as the researcher's primary device (used to initiate and monitor the study on iMotions) and the other for the researcher to note down observations. A Jellycomb 1920 × 1,080-pixel webcam was mounted on the top-centre of the participant's computer to collect face recordings. These were to be used to explain missing eye-tracking data if any. A room thermometer was used to record ambient temperature at the start of the experiment.

2.3 Procedure

Before the experiment, participants (wherever applicable) tied or pinned up long or loose strands of hair, removed makeup and accessories from their wrists and ears, and rolled up their sleeves. They were individually seated in a closed, well-lit and thermoregulated room (average ambient temperature 25 degrees Celsius). After completing an informed consent form and a demographics questionnaire, they received a general outline of the experimental set-up, procedure, tools, and expected code of conduct. The Affect Grid specifically was explained in detail – arousal was described as "how activated or aroused you feel" while pointing at the y-axis on the Affect Grid and valence was described as "how unpleasant or pleasant your emotion is" while pointing at the x-axis. This was followed by check-for-understanding (CFU) questions such as "You are running late for an exam and your bike has a flat tyre. Where would you mark an 'X' on the grid?" and a quick think-aloud of the reporting. For example, if a participant pointed to the first (high arousal-high valence) quadrant, the researcher would say something along the lines of "Yes, maybe because you feel anxious and anxiety is a negative emotion that is also activating" Other CFU questions/scenarios used were "You are thinking about a party you will attend this evening with your friends," "You just did yoga/meditation and are feeling relaxed" and "You are sitting in a very boring lecture and falling asleep." Next, participants were informed that all the instructions they would need would be on the screen and that the researcher would not help them with the tasks. They were also informed that if there was prize money attached to a task, this information would be in the task's instructions – textual information about prize money preceded the coordinate geometry puzzle and Tetris.

The Shimmer3 GSR+ unit was strapped tightly to the non-dominant forearm to mitigate motion artefacts. Its electrodes were attached to the third and fourth proximal phalanges of the participant's non-dominant hand. The ear clip was attached to the corresponding earlobe. Participants were also fitted with an Empatica

E4 on the same hand making sure that the thermopile sensor made complete contact with the dorsal side of the hand. Once the wearables were switched on and streaming data, participants waved their hand around a few times while the researcher ran a visual check on the signals. Thereafter, participants sat still for at least 10 min while signal readings were checked. This was done to ensure that the electrodes coupled with the participant's skin before the start of the experiment.

Participants were seated at a distance between 60 cm and 70 cm from the computer screen such that (a) they could see the reflection of their nose on the eye-tracker's frontal surface (an indication that the eye-tracker was at an optimum distance and height) and (b) iMotions' 'eye finder' widget indicated that both eyes were being detected. Participants placed their non-dominant hand on their lap and were discouraged from fidgeting or making big motions during the study.

The experiment was set up on iMotions meaning that all sensors were synchronized, and stimuli were timed and displayed on the platform. The study commenced when the quality of the eye-tracker calibration was deemed "excellent" by iMotions. The experiment started with a baseline reading (4 min) after which participants performed the three tasks. Each task lasted 12 min and participants made multiple attempts at the tasks during these periods. The order of the three tasks was counterbalanced across participants. Every four minutes, participants paused to complete the self-report on the touchscreen tablet. Thus, 13 instances of the self-report were collected – one for the baseline and three for each task. Participants had a 1-min 'cooling off' period between tasks while a screen with the instructions 'Sit still, relax and clear your mind' was displayed on their monitor. Participants were debriefed at the end of the experiment.

2.4 Signal processing

Most studies using skin conductance split the signal into the phasic component (i.e., the fast-moving signal that is an immediate response to stimuli) and the tonic component (i.e., the slow-moving signal) (Horvers et al., 2021). However, there is no consensus on which component to use (Horvers et al., 2021). Therefore, we investigated both. We used iMotions' R notebooks with their default parameters to process the raw skin conductance signal (measured in μS) (iMotions, 2022a). A time window of 4,000 ms was set as the threshold to determine gaps in the signal (due to signal drops) that would be linear interpolated. Missing data in gaps longer than the threshold were not interpolated and the resulting signal fragments were processed separately.

The phasic component was separated from the tonic component using a median filter over a time window of 8,000 ms. A lowpass Butterworth filter of 5 Hz was applied for noise filtration of the phasic signal. Skin conductance response (SCR) peaks were extracted from the phasic component using a 0.01 μS peak onset threshold, a 0 μs offset threshold and a 0.005 μS peak amplitude threshold. An onset is when the phasic signal surpasses a predetermined onset threshold, and an offset is when the signal drops below an offset threshold. A peak is the maximum value of a phasic signal within a time window determined by an onset-offset pair. Its amplitude is calculated as the difference between the value at the highest point and the value of the phasic signal at the onset. Each onset-offset pair defines a window, and the maximum value attained by the signal in this window is considered a peak. A value was marked as a peak if it was above the amplitude threshold of 0.005 μs in a window longer than 500 ms.

Eye-tracking data were processed using iMotions' R notebooks using their default parameters (iMotions, 2022b). Blinks were registered when data for both eyes were lost (an indication that eyes are closed) between 20 ms and 500 ms. Blinks were merged when the time between them was less than 70 ms. Fixation and saccade features were extracted using an I-VT (velocity-threshold identification) filter – if eyes moved slower than a velocity threshold of 30 degrees per second, the event was classified as a fixation and if they moved faster, a saccade was recorded. Fixation dispersion (i.e., the spread of a fixation's gaze points) was calculated by Imotions as the root mean square of the samples belonging to that fixation. Gaps in the signal shorter than 75 ms were interpolated.

Heart rate was calculated within iMotions. Linear interpolation was used to fill these gaps in the signal if the percentage of invalid data points was less than 10%. Since skin temperature was not collected in iMotions, it was synchronized (post-experiment) with the tasks using a Python script. Visual checks indicated no missing data.

For all participants, skin conductance and eye-tracking data with less than 20 dB signal-to-noise ratio (as indicated by iMotions) and all signals with greater than 10% missing data were excluded from the analysis. This resulted in excluding HR data of 10 participants from the analysis.

All in all, the following 13 data features were extracted – SCR peak count, average SCR peak amplitude, tonic skin conductance level, fixation count, duration and dispersion, saccade count, duration and amplitude, blink count and duration, heart rate, and skin temperature.

2.5 Analysis

To address issues of subjectivity and individual physiological variability, standardized values of physiological and gaze measures were used – mean values for the four-minute windows corresponding to each self-report (from the baseline and the three tasks) were calculated and finally z-scores *per participant* were computed. Self-reported emotions were labeled 1–4 based on the Affect Grid quadrant they fell in. For example, an 'X' on position (7,9) of the Affect Grid was labeled as 2 because it was in the second quadrant. Signals labeled with similar emotion labels (i.e., in the same quadrant) could come from different tasks, meaning that a label did not represent a task exclusively. Data points on the axes and therefore not in any quadrant [for example (5,7) or (8,5)] were excluded from further quantitative analyses. Finally, to determine if signals varied across emotional quadrants and if yes, what the pairwise differences were, Kruskal-Wallis tests with *post hoc* Dunn tests were performed. Since the baseline video was intended solely to provide a common point of departure for all participants and was therefore not included in the study's counterbalancing design, baseline readings were used for the computation of z-scores; however, they were excluded from the pairwise comparisons. This approach of using within-person z-scores provides a more reliable method of handling variability as compared to making calculations relative to a single baseline, because it acknowledges that: (a) a person's 'baseline' can shift due to various factors, making it difficult to capture a "true" baseline in a single measurement, (b) a single baseline measurement may not fully reflect an individual's actual physiological state and (c) subjective experiences of baseline tasks vary, and it is difficult to bring all participants to true 'baseline levels' with one standardized task.

3 Results

3.1 Self-reported emotional states across baseline and tasks

A significant within-subject difference in arousal and valence ratings across different points in time were observed, Pillai's Trace = 0.94, $F(18, 720) = 35.50$, $p < 0.001$. On average, participants recorded the following Affect Grid values: (a) during the baseline reading, low arousal (rating < 5) and positive valence (rating > 5), (b) during the shape-matching task, a steady decline in arousal and pleasure (i.e., valence), (c) on the coordinate geometry puzzle, high arousal (rating > 5) and negative valence (rating < 5), and (d) on Tetris, high arousal (rating > 5) positive valence (rating > 5) (see Figures 3, 4). According to the open-ended fill-in-the-blank statements, most (34.1%) participants felt "relaxed" after the baseline. Most (22.2%) high arousal-negative valence ratings were accompanied by the word "frustrated" or "confused" (19.8%), followed by "nervous" (6.2%) or "annoyed" (6.2%). The most commonly used words used to supplement high arousal-positive valence ratings were "excited" (14.2%) or "happy" (12.8%), followed by "good" (10.6%) or "focused" (9.9%). Low arousal-positive valence ratings were accompanied by words such as "relaxed" (32.2%), "bored" (15.3%), "calm" (8.5%) and "sleepy" (6.8%). Lastly, most participants reported low arousal-negative valence when they felt "bored" (27.6%), "annoyed" (10.5%), "tired" (7.9%) or "sleepy" (7.9%).

3.2 Physiological and gaze features across emotional quadrants and pairwise differences

Kruskal-Wallis tests indicated no significant differences in SCR peak amplitude [$\chi^2 = 6.57$, df = 3, $p = 0.09$], average heart rate [$\chi^2 = 5.58$, df = 3, $p = 0.13$] and skin temperature [$\chi^2 = 6.55$, df = 3, $p = 0.09$]. However, tonic skin conductance levels [$\chi^2 = 10.31$, df = 3, $p = 0.02$], SCR peak count [$\chi^2 = 24.99$, df = 3, $p < 0.001$], fixation count [$\chi^2 = 19.33$, df = 3, $p < 0.001$], fixation duration [$\chi^2 = 24.04$, df = 3, $p < 0.001$], fixation dispersion [$\chi^2 = 35.45$, df = 3, $p < 0.001$], saccade count [$\chi^2 = 29.33$, df = 3, $p < 0.001$], saccade duration [$\chi^2 = 33.36$, df = 3, $p < 0.001$], saccade amplitude [$\chi^2 = 26.92$, df = 3, $p < 0.001$], blink count [$\chi^2 = 31.25$, df = 3, $p < 0.001$] and blink duration [$\chi^2 = 42.34$, df = 3, $p < 0.001$] showed variations across quadrants. Table 1 shows the significant pairwise differences obtained from *post hoc* comparisons using Dunn's tests.

4 Discussion

Emotions are integral to the learning experience. In this study, we sought to determine physiological and gaze indices of learner emotions with the goal of informing the design of learning systems that can adapt to the emotions of learners. Our findings indicate that when students experience emotions such as relaxation, characterized by low arousal and positive valence, they exhibit more and shorter fixations along with fewer blinks than when they experience high arousal-negative valence emotions such as frustration. Furthermore,

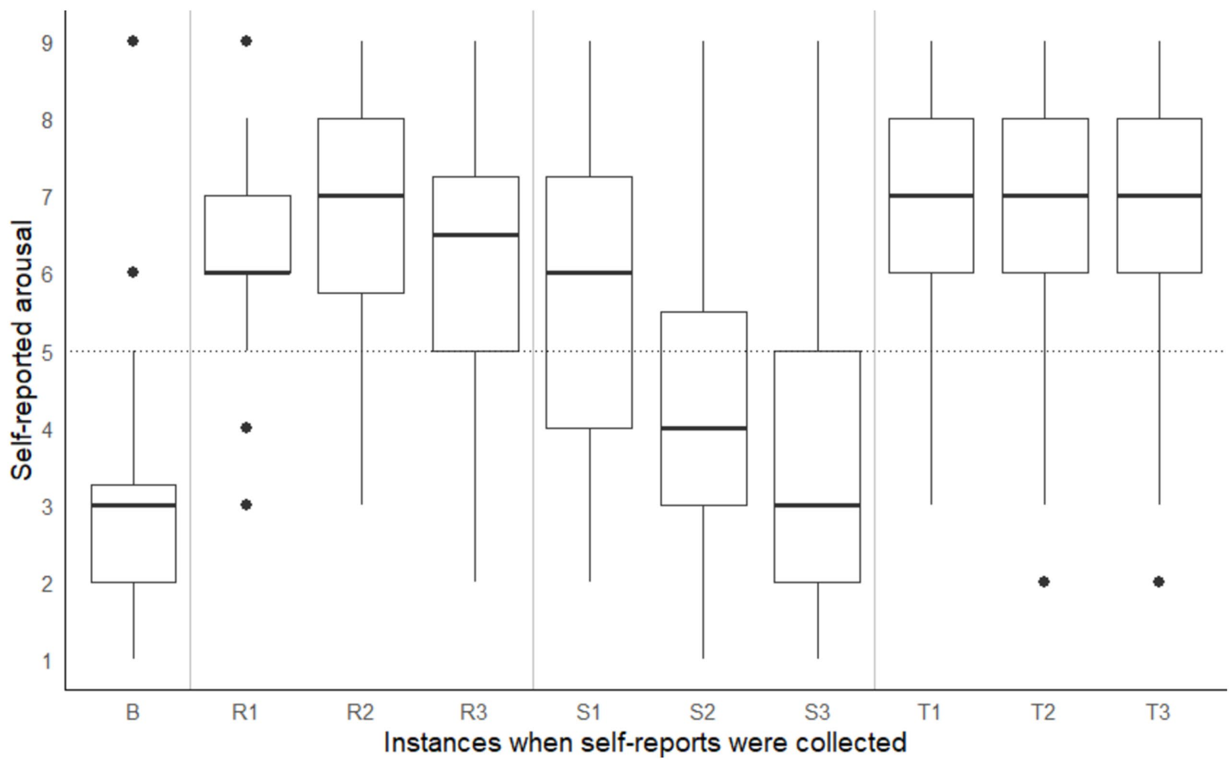


FIGURE 3
Self-reported arousal taken at baseline (B) and three points during the coordinate geometry puzzle (R1, R2, R3), shape-matching task (S1, S2, S3), and Tetris (T1, T2, T3). Note that task order was counterbalanced across participants.

10.3389/fpsyg.2024.1440425

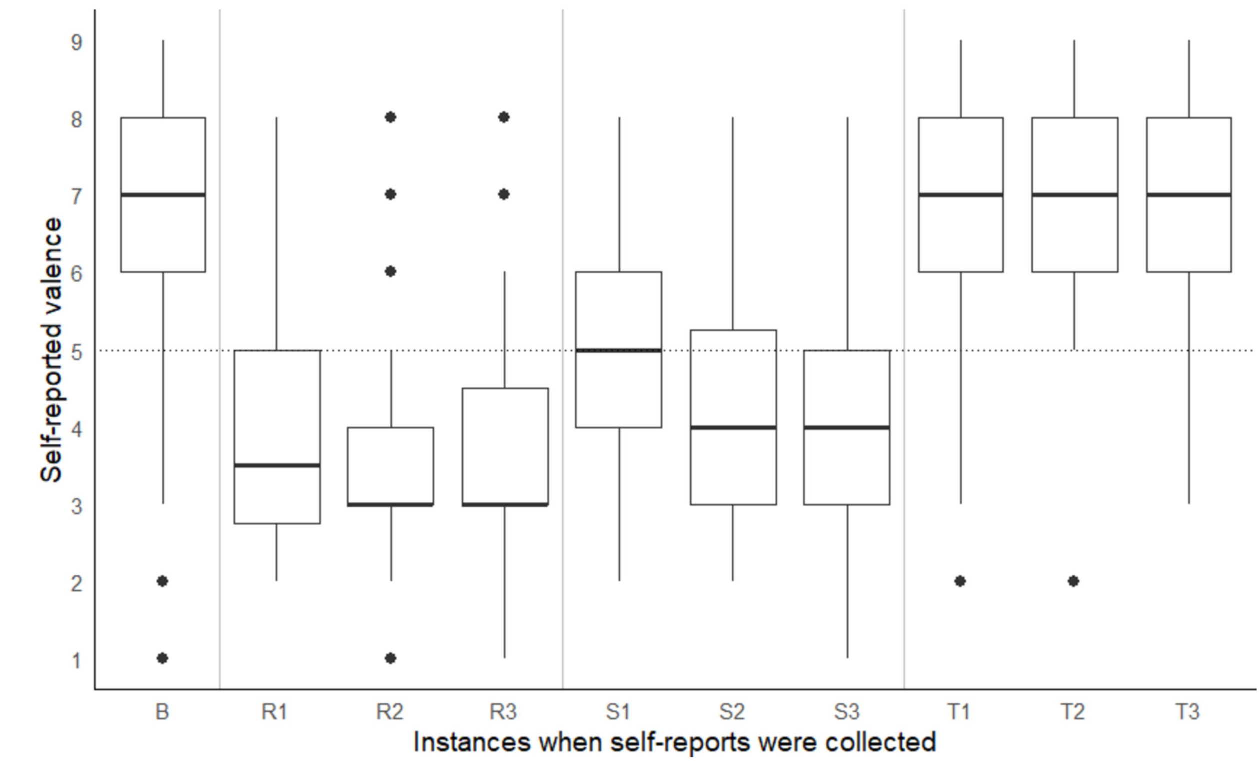


FIGURE 4
Self-reported valence taken at baseline (B) and three points during the coordinate geometry puzzle (R1, R2, R3), shape-matching task (S1, S2, S3) and Tetris (T1, T2, T3). Note that task order was counterbalanced across participants.

TABLE 1 Significant differences in features between quadrants of affect grid.

	Pairwise comparison of affect grid quadrants				
Feature	High arousal – Negative valence vs. High arousal – Positive valence	High arousal – Negative valence vs. Low arousal – Positive valence	High arousal – Negative valence vs. Low arousal – Negative valence	High arousal – Positive valence vs. Low arousal – Positive valence	High arousal – Positive valence vs. Low arousal – Negative valence
Tonic skin conductance			$p = 0.005$		
SCR peak count			$p = 0.002$		$p < 0.001$
Fixation count		$p = 0.004$	$p = 0.006$	$p = 0.006$	$p = 0.009$
Fixation duration		$p = 0.009$	$p = 0.007$	$p = 0.002$	$p < 0.001$
Fixation dispersion	$p < 0.001$				$p < 0.001$
Saccade count			$p = 0.01$	$p = 0.002$	$p < 0.001$
Saccade duration			$p < 0.001$	$p = 0.01$	$p < 0.001$
Saccade amplitude			$p < 0.001$	$p = 0.02$	$p < 0.001$
Blink count	$p < 0.001$	$p = 0.01$	$p = 0.002$		
Blink duration	$p < 0.001$				$p < 0.001$

Green: Group 1 > Group 2. For example, fixation duration in High Arousal – Negative Valence quadrant > Low Arousal – Positive Valence. Yellow: Group 1 < Group 2. For example, fixation count in High Arousal – Negative Valence quadrant < Low Arousal – Positive Valence.

low arousal-positive valence emotions differ from high arousal-positive emotions such as enjoyment in that they are accompanied by more and shorter fixations, fewer blinks, and more, slower and larger saccades. On the other hand, low arousal-negative emotions such as boredom differ from both high arousal-negative emotions and high arousal-positive emotions in that they are associated with fewer SCR peaks, more and shorter fixations, and more, slower and longer saccades. They are also characterized by longer blink durations and higher fixation dispersion than high arousal-positive emotions and fewer blink counts and lower tonic skin conductance than high arousal-negative emotions. Findings also indicate that high arousal-negative emotions are associated with higher fixation dispersion, blink count and blink duration as compared to high arousal-positive emotions. In fact, high arousal-negative emotions appear to be associated with the highest number of blinks. Heart rate and skin temperature were not found to be significant indicators of emotions.

High skin conductance observed during both positive and negative high-arousal emotions reaffirms the general understanding that this measure is a reliable index of emotional arousal (Boucsein, 2012). Meanwhile, examining the cognitive processes linked to each gaze measure and emotional state may provide possible explanations for the different eye movement patterns observed in this study. The high fixation counts observed during low-arousal emotions (e.g., boredom and relaxation) in this study reflect the findings of Foulsham et al. (2013) and Steindorf and Rummel (2019), that mindless reading can be associated with a high number of fixations. Earlier research has suggested that fixation count increases when "distractors are similar to targets" (Rayner, 1998), which may explain our findings, as boredom could lead to increased distraction, or in other words, increased attention to non-task-related elements. On the other hand, the high fixation durations observed during high-arousal emotions (such as frustration or excitement) in this study are likely due to increased visual attention and cognitive engagement, as these states often arise when individuals are task-oriented (Pekrun, 2006). Similarly, research on situational awareness—particularly in driving, where gaze dispersion is associated with heightened awareness of one's surroundings—suggests that more dispersed gaze indicates greater situational awareness (Liang et al., 2021). This correlation may account for the high fixation dispersion observed during high-arousal, negative-emotional states when students in this study may have engaged in more extensive visual searches to solve a problem. In contrast, past results from a SART (Sustained Attention to Response Task) indicate that an attentive state is accompanied by higher fixation dispersion as compared to a mind wandering state (i.e., when one's mind unconsciously wanders away from the task at hand) (Lee et al., 2021). Since mind wandering is a likely response to boredom (Randall et al., 2019), a low arousal-negative emotion, it may be possible to attribute this to the high fixation dispersion associated with this quadrant.

Additionally, several studies have found that eye movements during mind wandering are slower and less active than during attentive states (Faber et al., 2017; Uzzaman and Joordens, 2011). This is suggestive of slower and longer saccades associated with distracted visual scanning as compared to a focused visual pattern during task engagement. This is a possible explanation for the long saccade durations and large saccade amplitudes observed during low-arousal emotions (such as boredom or relaxation) in this study. Research on eye movements during stressful or anxiety-inducing situations, such as self-description in a foreign language, recalling a stressful event, viewing stress-or fear-inducing videos, and performing mental workload tasks (Giannakakis et al., 2017; Korda et al., 2021; Maffei and Angrilli, 2019) have shown a significant positive relationship between stress/anxiety and blink rate. The high blink rates observed during high-arousal, negative-valence emotions in this study are consistent with these findings. With respect to blink duration, research in vigilance and human factors suggests that longer blink durations during low-arousal, negative-valence states may be indicative of fatigue or drowsiness (Schleicher et al., 2008; Stern et al., 1996). This could explain this study's findings of high blink duration during low arousal-negative emotions. Unfortunately, we do not have a possible explanation for the high saccade count during low arousal-negative emotions and large blink durations during high arousal-negative emotions.

In their review of eye-tracking metrics related to emotional and cognitive processes, Skaramagkas et al. (2023) highlighted the complex, non-linear relationship between gaze measures and emotions, a finding that aligns with the results of this study. Insights from this study draw attention to the importance of integrating multimodal data for emotion detection. Overall, our findings highlight the potential of physiological and gaze measures to distinguish between different learner emotions, thus paving the way for potential intervention moments when a learner moves from one emotional state to another.

Results notwithstanding, it is important to note the limitations of this study. Firstly, in this study, we were unable to differentiate between low arousal-negative emotions and low arousal-positive emotions. A possible explanation is the clustering of self-reported valence near-neutral in the third and fourth quadrants. Of the 98 ratings in these quadrants, 45 had valence values between 4 and 6, making it hard to distinguish between the groups of emotions. This is further reflected in the overlap of emotion labels, as the words "bored" and "sleepy" are associated with both quadrants. Secondly, despite all three tasks requiring some extent of visual scanning of a scene and visual selection of an object of interest, it is possible that (other) variability in the task demands influenced eye movement. This is potentially a confounding variable in our study and future studies can benefit from attempting to mitigate this by ensuring comparability of the visual demands of their tasks. It is also worth emphasizing that this study was conducted in a controlled environment to ensure the relevance of the signals, focusing on their applicability, rather than the sensors themselves. Though the current set-up seems distant from real-word classrooms, the rationale was to confirm the feasibility of these measures before applying them in a phased manner using more accessible sensors in classrooms. However, the controlled lab-setting of this study eliminated several distractions that one would normally find in a real classroom. Therefore, to run a similar study in real world classrooms would also require measures of students' attention to the task at hand (thus ensuring that the emotions detected are in fact related to the learning) and error correction for external distractions. Additionally, words such as 'annoyed', 'bored' and 'sleepy' accompanied arousal and valence ratings in different quadrants of the Affect Grid. For example, sometimes participants experienced annoyance as a high arousal-negative emotion and sometimes as a low-arousal negative emotion. This is at odds with Russell's (1980) circumplex model that places each emotion in one specific quadrant (in this case, annoyance as high arousal-positive valence emotion). It may be that persistence of annoyance of the former kind leads to other high arousal-negative emotions such as frustration or anger while the second kind leads to other low arousal-negative emotions such as hopelessness or gloominess. However, to provide a clear explanation for this is beyond the scope of this study. Lastly, we acknowledge that generalizability of this study's findings are limited by the demographics included. For example, the sample primarily consisted of undergraduate students from Western Europe, which may not fully represent the broader population or account for physiological variations across different ages or cultural influences on emotional responses.

Mixed results of past research on what are reliable indicators of learner emotions may be attributed to a heterogeneity of (and sometimes a lack of transparency in) methodologies and devices used (Horvers et al., 2021; Lim et al., 2020; Yadegaridehkordi et al., 2019). With this paper, we hope to add to the corpus of clear methods for sensor-based studies in the field of education, thus paving the way for definitive study-design guidelines using such technology.

5 Conclusion

Students experience different emotions when engaging with learning-related tasks and this influences learning outcomes. Sensor technology today allows for (unobtrusive) collection of data that may eventually be used to provide personalized instruction or feedback to improve learning. In this paper, we investigated multimodal sensor data, namely skin conductance, skin temperature and gaze data as indicators of learner emotions operationalized by self-reports. Results indicate that skin conductance response peak count, tonic skin conductance levels, fixation count, duration and dispersion, saccade count, duration and amplitude, and blink count and duration can in fact be indicators of (self-reported) emotional arousal and valence in laboratory conditions. Researchers and designers may use these measures to make digital learning environments emotion-aware. These findings underline the need to move beyond the most extensively used measures – skin conductance and heart rate – and include several relevant factors. Results also reinforce the importance of doing psychophysiological research specific to the context of learning. On the whole, this study is a step toward emotion-aware learning systems.

Data availability statement

The processed data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Behavioural and Management Sciences (BMS) Ethics Committee at the University of Twente which subscribes to the Dutch Code of Ethics for Research in the Social and Behavioural Sciences. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. TE: Supervision, Writing – review & editing. HG: Supervision, Writing – review & editing. BV: Supervision, Writing – review & editing. JS: Formal analysis, Writing – review & editing. WV: Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was funded by the Faculty of Behavioural, Management and Social sciences at the University of Twente.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adão Martins, N. R., Annaheim, S., Spengler, C. M., and Rossi, R. M. (2021). Fatigue monitoring through wearables: a state-of-the-art review. *Front. Physiol.* 12:790292. doi: 10.3389/fphys.2021.790292

Aracena, C., Basterrech, S., Snáel, V., and Velásquez, J. (2015). "Neural networks for emotion recognition based on eye tracking data" in 2015 IEEE international conference on systems, man, and cybernetics (New York: IEEE), 2632–2637.

Avelar, D., Dore, R. A., Schwichtenberg, A. J., Roben, C. K., Hirsh-Pasek, K., and Golinkoff, R. M. (2022). Children and parents' physiological arousal and emotions during shared and independent e-book reading: a preliminary study. *Int. J. Child Comput. Interact.* 33:100507. doi: 10.1016/j.ijcci.2022.100507

Ba, S., and Hu, X. (2023). Measuring emotions in education using wearable devices: a systematic review. *Comput. Educ.* 200:104797. doi: 10.1016/j.compedu.2023.104797

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., and Graesser, A. C. (2010). Better to be frustrated than bored: the incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *Int. J. Human Comput. Stud.* 68, 223–241. doi: 10.1016/j.ijhcs.2009.12.003

Boucsein, W. (2012). Electrodermal activity. New York, NY: Springer.

Brouwer, A. M., Beurden, M. V., Nijboer, L., Derikx, L., Binsch, O., Gjaltema, C., et al. (2017). "A comparison of different electrodermal variables in response to an acute social stressor" in International workshop on symbiotic interaction (Cham: Springer), 7–17.

Chanel, G., Rebetez, C., Bétrancourt, M., and Pun, T. (2008). Boredom, engagement, and anxiety as indicators for adaptation to difficulty in games. Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era (MindTrek '08). Association for Computing Machinery. 13–17.

Clements, D. H., Battista, M. T., Sarama, J., and Swaminathan, S. (1997). Development of students' spatial thinking in a unit on geometric motions and area. *Elem. Sch. J.* 98, 171–186. doi: 10.1086/461890

Craig, S., Graesser, A., Sullins, J., and Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *J. Educ. Media* 29, 241–250. doi: 10.1080/1358165042000283101

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York: Harper & Row.

D'Mello, S., and Graesser, A. (2012). Dynamics of affective states during complex learning. *Learn. Instr.* 22, 145–157. doi: 10.1016/j.learninstruc.2011.10.001

Darnell, D. K., and Krieg, P. A. (2019). Student engagement, assessed using heart rate, shows no reset following active learning sessions in lectures. *PLoS One* 14:e0225709. doi: 10.1371/journal.pone.0225709

Faber, M., Bixler, R., and D'Mello, S. K. (2017). An automated behavioral measure of mind wandering during computerized reading. *Behav. Res. Methods* 50, 134–150. doi: 10.3758/s13428-017-0857-y

Foulsham, T., Farley, J., and Kingstone, A. (2013). Mind wandering in sentence reading: decoupling the link between mind and eye. *Canadian J. Exp. Psychol.* 67, 51–59. doi: 10.1037/a0030217

Giannakakis, G., Pediaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P., et al. (2017). Stress and anxiety detection using facial cues from videos. *Biomed. Signal Proces. Control* 31, 89–101. doi: 10.1016/j.bspc.2016.06.020

Gomez-Ibañez, A., Urrestarazu, E., and Viteri, C. (2014). Recognition of facial emotions and identity in patients with mesial temporal lobe and idiopathic generalized epilepsy: an eye-tracking study. *Seizure* 23, 892–898. doi: 10.1016/j.seizure.2014.08.012

Harley, J. M., Pekrun, R., Taxer, J. L., and Gross, J. J. (2019). Emotion regulation in achievement situations: an integrated model. *Educ. Psychol.* 54, 106–126. doi: 10.1080/00461520.2019.1587297

Harmat, L., de Manzano, Ö., Theorell, T., Högman, L., Fischer, H., and Ullén, F. (2015). Physiological correlates of the flow experience during computer game playing. *Int. J. Psychophysiol.* 97, 1–7. doi: 10.1016/j.ijpsycho.2015.05.001

Hernández-Cuevas, B. Y., and Crawford, C. S. (2021). A literature review of physiological-based Mobile educational systems. *IEEE Trans. Learn. Technol.* 14, 272–291. doi: 10.1109/TLT.2021.3098315

Hessels, R. S., Niehorster, D. C., Nyström, M., Andersson, R., and Hooge, I. T. (2018). Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers. *R. Soc. Open Sci.* 5:180502. doi: 10.1098/rsos.180502

Horvers, A., Tombeng, N., Bosse, T., Lazonder, A. W., and Molenaar, I. (2021). Detecting emotions through electrodermal activity in learning contexts: a systematic review. *Sensors* 21:7869. doi: 10.3390/S21237869

iMotions. (2022a). Analysis: GSR peak detection & aggregation. Available online at: https://help.imotions.com/v2/docs/en/analysis-gsr-peak-detection-amp-aggregation-imotions-80-and-previous

iMotions. (2022b). Fixation and saccade classification with I-VT filter. Available online at: https://help.imotions.com/v2/docs/en/fixation-and-saccade-classification-with-the-i-vt-filter

Jang, E., Byun, S., Park, M., and Sohn, J. (2019). Reliability of physiological responses induced by basic emotions: a pilot study. *J. Physiol. Anthropol.* 38:15. doi: 10.1186/s40101-019-0209-y

Jang, E. H., Park, B. J., Park, M. S., Kim, S. H., and Sohn, J. H. (2015). Analysis of physiological signals for recognition of boredom, pain, and surprise emotions. *J. Physiol. Anthropol.* 34, 25–12. doi: 10.1186/s40101-015-0063-5

Jennings, J. R., Kamarck, T., Stewart, C., Eddy, M., and Johnson, P. (1992). Alternate cardiovascular baseline assessment techniques: Vanilla or resting baseline. *Psychophysiology* 29, 742–750. doi: 10.1111/j.1469-8986.1992.tb02052.x

Jindrová, M., Kocourek, M., and Telenský, P. (2020). Skin conductance rise time and amplitude discern between different degrees of emotional arousal induced by affective pictures presented on a computer screen. *bio Rxiv*. doi: 10.1101/2020.05.12.090829

Killgore, W. D. S. (1998). The affect grid: a moderately valid, nonspecific measure of pleasure and arousal. *Psychol. Rep.* 83, 639–642. doi: 10.2466/pr0.1998.83.2.639

Korda, A. I., Giannakakis, G., Ventouras, E., Asvestas, P. A., Smyrnis, N., Marias, K., et al. (2021). Recognition of blinks activity patterns during stress conditions using CNN and Markovian analysis. *Signals* 2, 55–71. doi: 10.3390/signals2010006

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biolo. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010

Lal, S., Eysink, T. H., Gijlers, H. A., Verwey, W. B., and Veldkamp, B. P. (2021). Detecting emotions in a learning environment: a multimodal exploration. In 16th European conference on technology enhanced learning, EC-TEL 2021 (pp. 71–82).

Larmuseau, C., Cornelis, J., Lancieri, L., Desmet, P., and Depaepe, F. (2020). Multimodal learning analytics to investigate cognitive load during online problem solving. *Br. J. Educ. Technol.* 51, 1548–1562. doi: 10.1111/bjet.12958

Lee, H., Chen, Z., Yeh, S., Hsiao, J. H., and Wu, A. (2021). When eyes wander around: mind-wandering as revealed by eye movement analysis with hidden Markov models. *Sensors* 21:7569. doi: 10.3390/s21227569

Lee, V. R., Fischback, L., and Cain, R. (2019). A wearables-based approach to detect and identify momentary engagement in afterschool makerspace programs. *Contemp. Educ. Psychol.* 59:101789. doi: 10.1016/j.cedpsych.2019.101789

Liang, N., Yang, J., Yu, D., Prakah-Asante, K. O., Curry, R., Blommer, M., et al. (2021). Using eye-tracking to investigate the effects of pre-takeover visual engagement on situation awareness during automated driving. *Accid. Anal. Prev.* 157:106143. doi: 10.1016/j.aap.2021.106143

Lim, J. Z., Mountstephens, J., and Teo, J. (2020). Emotion recognition using eye-tracking: taxonomy, review and current challenges. *Sensors* 20:2384. doi: 10.3390/S20082384

Lindenmuth, T. L. (n.d.). Shape matching: Scored. Available online at: https://www.geogebra.org/m/rvz58cma#material/eMCXcErd

Liu, Z., Ren, Y., Kong, X., and Liu, S. (2022). Learning analytics based on wearable devices: a systematic literature review from 2011 to 2021. *J. Educ. Comput. Res.* 60, 1514–1557. doi: 10.1177/07356331211064780

Maffei, A., and Angrilli, A. (2019). Spontaneous blink rate as an index of attention and emotion during film clips viewing. *Physiol. Behav.* 204, 256–263. doi: 10.1016/j.physbeh.2019.02.037

Maier, M., Elsner, D., Marouane, C., Zehnle, M., and Fuchs, C. (2019). DeepFlow: detecting optimal user experience from physiological data using deep neural networks. *IJCAI Int. Joint Conf. Artif. Intel.*, 1415–1421. doi: 10.24963/IJCAI.2019/196

Mayer, R. E. (2014). Computer games for learning: an evidence-based approach. Cambridge, MA: MIT Press.

McFarland, R. A. (1985). Relationship of skin temperature changes to the emotions accompanying music. *Biofeedback Self Regul.* 10, 255–267. doi: 10.1007/bf00999346

Meer, Y., Breznitz, Z., and Katzir, T. (2016). Calibration of self-reports of anxiety and physiological measures of anxiety while reading in adults with and without reading disability. *Dyslexia* 22, 267–284. doi: 10.1002/DYS.1532

Noroozi, O., Pijeira-Díaz, H. J., Sobocinski, M., Dindar, M., Järvelä, S., and Kirschner, P. A. (2020). Multimodal data indicators for capturing cognitive, motivational, and emotional learning processes: a systematic literature review. *Educ. Inf. Technol.* 25, 5499–5547. doi: 10.1007/s10639-020-10229-w

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341. doi: 10.1007/s10648-006-9029-9

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., and Goetz, T. (2017). Achievement emotions and academic performance: longitudinal models of reciprocal effects. *Child Dev.* 88, 1653–1670. doi: 10.1111/cdev.12704

Piferi, R. L., Kline, K. A., Younger, J., and Lawler, K. A. (2000). An alternative approach for achieving cardiovascular baseline: viewing an aquatic video. *Int. J. Psychophysiol.* 37, 207–217. doi: 10.1016/s0167-8760(00)00102-1

Pijeira-Díaz, H. J., Drachsler, H., Kirschner, P. A., and Järvelä, S. (2018). Profiling sympathetic arousal in a physics course: how active are students? *J. Comput. Assist. Learn.* 34, 397–408. doi: 10.1111/jcal.12271

Randall, J. G., Beier, M. E., and Villado, A. J. (2019). Multiple routes to mind wandering: predicting mind wandering with resource theories. *Conscious. Cogn.* 67, 26–43. doi: 10.1016/j.concog.2018.11.006

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372

Rimm-Kaufman, S. E., and Kagan, J. (1996). The psychological significance of changes in skin temperature. *Motiv. Emot.* 20, 63–78. doi: 10.1007/bf02251007

Ritz, T., Thöns, M., Fahrenkrug, S., and Dahme, B. (2005). Airways, respiration, and respiratory sinus arrhythmia during picture viewing. *Psychophysiology* 42, 568–578. doi: 10.1111/J.1469-8986.2005.00312.X

Roos, A. L., Goetz, T., Voracek, M., Krannich, M., Bieg, M., Jarrell, A., et al. (2021). Test anxiety and physiological arousal: a systematic review and meta-analysis. *Educ. Psychol. Rev.* 33, 579–618. doi: 10.1007/S10648-020-09543-Z

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/H0077714

Russell, J. A., Weiss, A., and Mendelsohn, G. A. (1989). Affect grid: a single-item scale of pleasure and arousal. *J. Pers. Soc. Psychol.* 57, 493–502. doi: 10.1037/0022-3514.57.3.493

Schleicher, R., Galley, N., Briest, S., and Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics* 51, 982–1010. doi: 10.1080/00140130701817062

Shaby, N., and Bokhove, C. (2023). Using EDA measures to detect emotional expressions during family science activities, a methodological perspective. *Int. J. Res. Method Educ.* 47, 378–393. doi: 10.1080/1743727X.2023.2289539

Sharma, K., Papamitsiou, Z., Olsen, J. K., and Giannakos, M. (2020). Predicting learners' effortful behaviour in adaptive assessment using multimodal data. In Proceedings of the tenth international conference on learning analytics & knowledge (pp. 480–489).

Sims, V. K., and Mayer, R. E. (2001). Domain specificity of spatial expertise: the case of video game players. *Appl. Cogn. Psychol.* 16, 97–115. doi: 10.1002/acp.759

Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N., et al. (2023). Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Rev. Biomed. Eng.* 16, 260–277. doi: 10.1109/rbme.2021.3066072

Steindorf, L., and Rummel, J. (2019). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behav. Res. Methods* 52, 162–176. doi: 10.3758/s13428-019-01214-4

Stern, J. A., Schroeder, D. J., Touchstone, R. M., Stoliarov, N., and Boyer, D. J. (1996). Blinks, saccades, and fixation pauses during vigilance task performance: II. Gender and time of day: PsycEXTRA Dataset. Washington, DC: American Psychological Association.

Tetris. (1985). Tetris ® & © 1985~2024 Tetris Holding, used with permission.

Thayer, R. E. (1967). Measurement of activation through self-report. *Psychol. Rep.* 20, 663–678. doi: 10.2466/pr0.1967.20.2.663

Thayer, R. E. (1978). Toward a psychological theory of multidimensional activation (arousal). *Motiv. Emot.* 2, 1–34. doi: 10.1007/bf00992729

Tsang, V. (2018). Eye-tracking study on facial emotion recognition tasks in individuals with high-functioning autism spectrum disorders. *Autism* 22, 161–170. doi: 10.1177/1362361316667830

Tsuruda, K. M. (2010). A closer look at Tetris: analysis of a variant game. *Atlantic Electron. J. Math.* 4, 23–34.

Uzzaman, S., and Joordens, S. (2011). The eyes know what you are thinking: eye movements as an objective measure of mind wandering. *Conscious. Cogn.* 20, 1882–1886. doi: 10.1016/j.concog.2011.09.010

Van Bruinessen, I. R., van den Ende, I. T. A., Visser, L. N. C., and Van Dulmen, S. (2016). The impact of watching educational video clips on analogue patients' physiological arousal and information recall. *Patient Educ. Couns.* 99, 243–249. doi: 10.1016/j.pec.2015.08.022

Williams, S., and Bright, G. W. (1991). "Spatial learning in one video game" in Proceedings of the annual meeting of the north American chapter of the International Group for the Psychology of mathematics education (13th, Blacksburg, Virginia, October 16–19, 1991) (Vols. 1–2). ed. R. G. Underhill (Karlsruhe: International Group for the Psychology of Mathematics Education), 139–145.

Wu, C. H., Huang, Y. M., and Hwang, J. P. (2016). Review of affective computing in education/learning: trends and challenges. *Br. J. Educ. Technol.* 47, 1304–1323. doi: 10.1111/bjet.12324

Yadegaridehkordi, E., Noor, N. F. B. M., Ayub, M. N. B., Affal, H. B., and Hussin, N. B. (2019). Affective computing in education: a systematic review and future research. *Comput. Educ.* 142:103649. doi: 10.1016/j.compedu.2019.103649

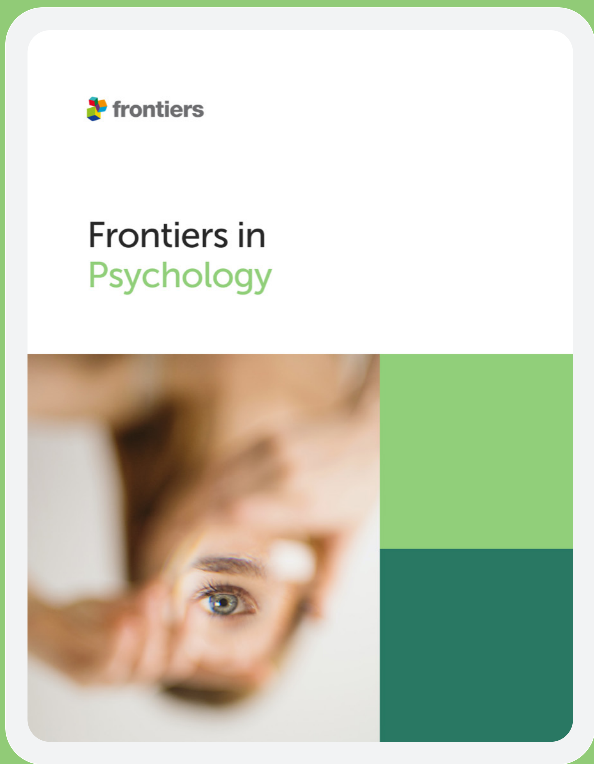
Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact