

RESEARCH ARTICLE

Open Access



Leveraging computer vision for adaptive learning in STEM education: effect of engagement and self-efficacy

Ting-Ting Wu¹ , Hsin-Yu Lee² , Wei-Sheng Wang² , Chia-Ju Lin² and Yueh-Min Huang^{2*}

*Correspondence:
huang@mail.ncku.edu.tw

¹ Graduate School
of Technological and Vocational
Education, National Yunlin
University of Science
and Technology, Yunlin, Taiwan

² Department of Engineering
Science, National Cheng Kung
University, No. 1, University Road,
Tainan 701, Taiwan, R.O.C.

Abstract

In the field of Science, Technology, Engineering, and Mathematics (STEM) education, which aims to cultivate problem-solving skills, accurately assessing learners' engagement remains a significant challenge. We present a solution to this issue with the Real-time Automated STEM Engagement Detection System (RASEDS). This innovative system capitalizes on the power of artificial intelligence, computer vision, and the Interactive, Constructive, Active, and Passive (ICAP) framework. RASEDS uses You Only Learn One Representation (YOLOR) to detect and map learners' interactions onto the four levels of engagement delineated in the ICAP framework. This process informs the system's recommendation of adaptive learning materials, designed to boost both engagement and self-efficacy in STEM activities. Our study affirms that RASEDS accurately gauges engagement, and that the subsequent use of these adaptive materials significantly enhances both engagement and self-efficacy. Importantly, our research suggests a connection between elevated self-efficacy and increased engagement. As learners become more engaged in their learning process, their confidence is bolstered, thereby augmenting self-efficacy. We underscore the transformative potential of AI in facilitating adaptive learning in STEM education, highlighting the symbiotic relationship between engagement and self-efficacy.

Keywords: Adaptive learning, Engagement assessment, ICAP framework, STEM education, Computer vision, Self-efficacy

Introduction

STEM education aims to integrate knowledge from interdisciplinary fields, fostering learners' problem-solving abilities and playing a vital role in higher education (Reinholz et al., 2021). At the higher education level, learners are faced with increasingly complex problems and challenges. The interdisciplinary nature of STEM education provides the necessary tools to tackle a diverse range of issues (Wong et al., 2022). In essence, STEM education is a learner-centric instructional model, where learners are expected to explore and plan their learning journey, facilitating their knowledge construction (Borda et al., 2020). Unlike traditional education, which tends to focus solely on evaluating learning outcomes, STEM education places significant emphasis on the learning

engagement itself (Lee et al., 2023a, 2023b). Therefore, effectively assessing learners' learning engagement can provide insights into their performance within the STEM education framework, contributing to the development of well-rounded individuals in higher education (Smith Iv et al., 2020).

Learning Engagement, a common indicator of learners' involvement and active participation in the learning process, reflects the time and energy invested by learners (Wang et al., 2016). The relationship between learning engagement and academic performance has been confirmed by numerous studies (Bai et al., 2021; Qureshi et al., 2021). The ICAP is a well-established framework for assessing levels of engagement, particularly cognitive engagement. It categorizes cognitive engagement into four distinct levels: passive, active, constructive, and interactive. This categorization assists educators in devising more effective learning strategies (Chi & Wylie, 2014). ICAP framework has been widely incorporated into STEM activities in various studies to serve as a benchmark for assessing cognitive engagement (Hsiao et al., 2022; Lee et al., 2023a, 2023b; Lee et al., 2023a, 2023b). Moreover, a study conducted by Lee et al., (2023a, 2023b) further substantiated the critical role each indicator within the ICAP framework plays in fostering student engagement during hands-on activities in STEM fields.

According to a systematic review by Gao et al. (2020), self-reporting and observation methods dominate the measurement of learning engagement in STEM education. However, self-reporting, typically conducted through questionnaires or interviews after the course, may lead to inaccurate recall or be influenced by societal expectations (Zimmerman, 2008). Observation methods require researchers to code images or records of activities to assess learners' engagement, a method that can be time-consuming and costly (Harari et al., 2017). To overcome these disadvantages, we combine the ICAP framework with computer vision technology to develop a system, called Real-time Automated STEM Engagement Detection System (RASEDS), effectively and instantaneously measuring the learners' learning engagement. Computer vision technology is used to capture learners' behaviors during group activities, and this information is mapped to the four levels of the ICAP framework, enabling automated measurement of learners' engagement.

On the other hand, self-efficacy and engagement in STEM education share a dynamic and reciprocal relationship (Han et al., 2021). Self-efficacy refers to an individual's confidence and expectations of their abilities, significantly influencing their enthusiasm, effort, and persistence when facing challenges (Chen et al., 2001). Conversely, engagement refers to the level of active interest and involvement in learning tasks (Fredricks et al., 2004). When learners actively engage in STEM learning, they experience success, mastery, and competence, fostering positive experiences that validate their abilities, bolster confidence, and nurture self-efficacy (Bandura, 1977; Kuchynka et al., 2021). Continuous and meaningful engagement not only reinforces self-efficacy but also enhances opportunities for further success and mastery, allowing learners to showcase effective problem-solving skills (Han et al., 2021). Engagement in STEM education encourages learners to assume responsibility for their learning, promoting personal agency and fostering a sense of control and ownership over their educational experiences (Bandura, 1977; Prince, 2004). From this, it is clear the pivotal role both engagement and self-efficacy play in enhancing STEM education.

However, traditional one-size-fits-all teaching methods often fail to cater to the varied knowledge bases, learning styles, interests, and abilities of all learners, leading to feelings of insecurity and helplessness. This not only affects self-efficacy but can also result in lower learning outcomes, reduced engagement, and in extreme cases, dropout or abandonment (Cook et al., 2018; Hu, 2022; Zhu & Wang, 2020). The introduction of adaptive learning can mitigate the shortcomings of traditional teaching methods (El-Sabagh, 2021; Guerrero-Roldán et al., 2021). However, past technological limitations made it difficult to timely and automatically evaluate learning engagement in STEM education and implement adaptive learning strategies. Recently, the advancement in Artificial Intelligence (AI) technology has fostered a growing body of research utilizing AI to support adaptive learning in classrooms. Despite this progress, research focusing on STEM activities remains limited due to their complex, abstract, and multi-dimensional nature compared to traditional subjects (Wang et al., 2022).

Addressing these research gaps, we developed RASEDS, grounded in AI and computer vision technology. This system can integrate ICAP framework to instantly assess learners learning engagement. Ultimately, it recommends tailored learning materials to facilitate adaptive learning, aiming to revolutionize STEM education through a nuanced understanding of engagement dynamics.

Related work

The measurement of engagement in STEM education

STEM education, which integrates knowledge and skills across the domains of Science, Technology, Engineering, and Mathematics, plays a pivotal role in fostering learners' innovative capabilities and problem-solving abilities (Hsiao et al., 2022). It is instrumental in grappling with intricate global issues, propelling economic growth, and enhancing societal well-being (Christensen et al., 2015). Nonetheless, STEM education faces several challenges including a deficiency in learners' interest and confidence, a high rate of talent attrition, and a dearth of diversity (Sithole et al., 2017). Bolstering learners' engagement in STEM education is paramount in ameliorating learning outcomes and piquing learners' interest (Miller et al., 2018). Learners' engagement, defined as the degree of behavioral, emotional, and cognitive investment in the learning process, significantly impacts academic performance, motivation, satisfaction, and graduation rates (Huang & Wang, 2023). The primary methods to measure engagement in STEM education are self-reporting and observation, each with its inherent limitations (Gao et al., 2020). Self-reporting, although allowing learners to appraise their engagement through surveys or interviews, is susceptible to subjectivity (Baumeister et al., 2007; Paulhus & Vazire, 2007). The observational method, on the other hand, objectively assesses learners' behaviors and interactions, but it may introduce bias and prove time-consuming (D'Mello et al., 2017).

To circumvent these constraints, recent research has begun to leverage AI technology in educational settings to glean insights into learners' engagement during the educational process. For instance, Zhang et al. (2019) collected and analyzed data on learners' facial expressions and mouse movements within an online learning environment, thereby enabling an assessment of learners' engagement during online learning. Similarly, Flanagan et al. (2022) identified reading behavior features from digital learning

materials, using these extracted features to predict learners' performance and engagement through Support Vector Machines (SVM). This led to the establishment of an early warning intervention system in higher education learning environments. More recently, Zheng et al. (2023) created a system for self-analysis and feedback on group engagement in Computer-Supported Collaborative Learning (CSCL) environments, grounded in Deep Neural Network models (DNNs). Through a quasi-experimental design, they validated the system's contribution to learning outcomes. Despite these advances, the construction of automated systems for assessing engagement in STEM education remains a significant challenge (Lee et al., 2023a, 2023b; Wang et al., 2022).

In response to this challenge, the present study merges the ICAP (Interactive, Constructive, Active, Passive) framework with computer vision technology to develop a real-time, automated engagement recognition system for STEM education. This system, capable of effectively and instantaneously measuring learners' engagement during the learning process, forms the foundation for recommending adaptive learning materials tailored to individual learners' needs.

ICAP framework

Chi and Wylie (2014) proposed the ICAP framework, a categorization of learning engagement into four unique modes, enabling researchers to associate specific modes with cognitive engagement and ultimately understand the evolution of a learner's knowledge. The ICAP framework classifies in-class cognitive engagement into four modes: passive, active, constructive, and interactive, as detailed below:

- **Passive:** Learners merely assimilate information from teaching resources without active participation in the learning process. For example, they may attentively listen without taking notes or recording information.
- **Active:** Learners demonstrate observable actions or physical interactions. These may include actions such as pausing, fast-forwarding, or rewinding educational videos, annotating and note-taking on learning content, or employing gestures to manipulate learning materials for problem-solving.
- **Constructive:** Learners produce additional externalized outputs or products that extend beyond the given learning materials. This may involve activities such as assembling and operating robots to acquire programming knowledge or creating concept maps to grasp their knowledge construction process.
- **Interactive:** Interaction is defined by two criteria: the statements of both participants must be primarily constructive, and sufficient turn-taking must be observed (Chi & Wylie, 2014). Chi and Wylie (2014) stress that constructive behavior from both parties is essential for productive discussions. Adequate turn-taking during the interaction facilitates the integration of each participant's domain knowledge and the adjustment of cognitive states. Interaction partners can include teachers, parents, or even robots, as long as these conditions are satisfied.

The ICAP framework views learning engagement as a progressive enhancement of learners' cognitive states and engagement with the material (Chi & Wylie, 2014). Researchers often utilize the ICAP framework as an evaluative tool in various studies.

For instance, Raković et al. (2020) applied the model to analyze learners' messages in forums, aiding educators in understanding knowledge construction and predicting learners' exam and debate performance. Liu et al. (2022) developed the BERT-CNN model to automatically detect learners' cognitive and emotional engagement, using the ICAP framework to evaluate the degree of cognitive engagement among learners in MOOC forum discussions. The practical approach of the ICAP framework facilitates a systematic understanding of the cognitive engagement during courses. As a result, we also develop a system based on the ICAP framework to assist researchers and educators in gaining a deeper understanding of learning engagement in STEM education and establishing a foundation for educational assessment in this field.

Adaptive learning in STEM education

In recent years, the advent of digital technology, particularly the application of big data, cloud computing, and artificial intelligence, has increasingly highlighted the importance of adaptive learning. Adaptive learning is an innovative educational approach that uses advanced educational technology, data analytics, and artificial intelligence to personalize learning content, processes, and methods. This allows educators to respond more effectively to the diverse differences among learners, creating a learning experience that is more suitable for individual characteristics (Peng et al., 2019; Wang et al., 2023). Numerous studies have verified the advantages of adaptive learning for learners. For instance, El-Sabagh (2021) developed an adaptive learning platform based on learners' learning styles and confirmed its benefits for learners' engagement through quasi-experimental design. Kabudi et al. (2021) discussed the positive impacts of incorporating AI technology into adaptive learning in terms of enhancing learners' performance and motivation. Wang et al. (2023) examined the influence of China's first adaptive learning system, Squirrel AI learning, on large and small class teaching in a mathematics course, and confirmed the system's effect on learners' mathematical learning.

However, the implementation of adaptive learning faces several challenges in STEM education (Liu, 2022). First, STEM education emphasizes experimental and inquiry-based learning to foster learners' understanding and application of scientific concepts (Wang et al., 2022). This implies that STEM curricula frequently involve lab activities and hands-on experiences, which are challenging to fully realize through computer-assisted learning systems (Chang & Chen, 2022; Lin et al., 2021). Therefore, incorporating adaptive learning into these non-computerized teaching activities requires researchers to explore new technologies and methods to capture learners' statuses and provide appropriately adaptive support (Afini Normadhi et al., 2019). Second, STEM education encourages interaction and collaboration among learners to promote problem-solving and the development of critical and creative thinking (Margot & Kettler, 2019). In this context, learners' learning process are often highly dynamic and complex, involving knowledge and skills from multiple disciplinary areas (Wang et al., 2022). Traditional adaptive learning systems may struggle to accurately identify and evaluate such diversified, interdisciplinary learning processes, and consequently, provide appropriate adaptive support (Mirata et al., 2020). Additionally, STEM curricula typically emphasize assessment of the learning process rather than solely relying on final exam scores (Gao et al., 2020). Thus, implementing adaptive learning in STEM education requires adopting

more diverse, comprehensive assessment indicators to capture learners' learning outcomes at different levels, which poses higher demands for the design and implementation of adaptive learning systems. To address this research gap, we develop a system, called RASEDS based on computer vision technology and the ICAP framework. This system immediately comprehends learners' engagement in STEM activities and provides corresponding adaptive learning materials based on the level of learners' engagement in STEM activities, aiming to achieve the goal of implementing adaptive learning in STEM education.

Research purpose and questions

From the aforementioned discussion, it becomes apparent that there is a current shortage of automated tools to assess learners' engagement in STEM education. This deficit indirectly hampers the implementation of adaptive learning within STEM education, which emphasizes practical applications and collaborative work in small groups. In response to this issue, we leverage computer vision technology and the ICAP (Interactive, Constructive, Active, Passive) framework to develop a system, called Real-time Automated STEM Engagement Detection System (RASEDS). This system identifies the level of learners' engagement during STEM activities, and based on the degree of engagement, it provides corresponding adaptive learning materials. Ultimately, we aim to validate the efficacy of our approach through a quasi-experimental design. This approach tests whether adaptive learning materials, based on engagement levels in STEM education, result in improved engagement and self-efficacy among learners during STEM activities. Consequently, we attempt to answer the following research questions:

- i. Can the Real-time Automated STEM Engagement Detection System (RASEDS) proposed in this study effectively identify learners' engagement in STEM education?
- ii. Can the recommendation of adaptive learning materials based on RASEDS results enhance learners' engagement in STEM activities?
- iii. Does the enhancement of learners' engagement in STEM activities, facilitated by the recommendation of adaptive learning materials based on RASEDS results, subsequently boost their self-efficacy?

Methodology

To address the research question, we conducted multiple "STEM Workshop: Python and Raspberry Pi Practical Activity" prior to the experiment. This setup facilitated the collection of data required for the development and validation of RASEDS, which responds to the first research question. Upon establishing the effectiveness of the RASEDS, the system's output can be utilized to construct STEM performance prediction model. This model, in turn, can recommend adaptive learning materials based on its predictions. A quasi-experimental design was employed in this research to validate whether the recommendation of adaptive learning materials via RASEDS enhance learners' engagement and self-efficacy in STEM activities, ultimately addressing the second and third research questions. The research workflow is illustrated in Fig. 1.

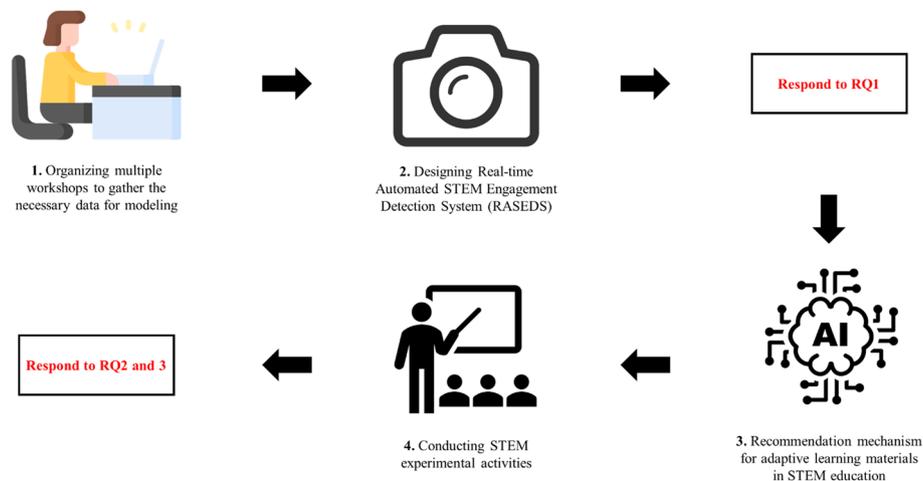


Fig. 1 The research workflow

STEM workshop: Python and Raspberry Pi Practical Activity

To accumulate the requisite data for this modeling project, the study organized three single-day sessions of the “STEM Workshop: Python and Raspberry Pi Practical Activity” ahead of the experimental phase. A total of 86 learners, ranging in age from 16 to 22, engaged in these sessions. These workshops were designed to foster programming skills and to encourage problem-solving capabilities through the infusion of computer science principles, honoring the foundations of STEM education. During the workshops, participants utilized their self-designed codes in a tangible environment through Raspberry Pi, enhancing not only their understanding but also their appreciation for programming concepts, aligning with the tangible learning framework proposed by Marshall (2007). To foster collaboration and idea exchange, learners were encouraged to interact and discuss with one another throughout the activity.

To document each participant’s learning trajectory, cameras were strategically positioned to capture the learners’ interactions with the hands-on materials essential for engagement identification by the RASEDS. Special emphasis was placed on selecting camera angles that would vividly showcase both the participants’ hands and the learning resources involved in the process, as depicted in Fig. 2. Acknowledging the sensitive nature of recording minors, we instituted stringent protocols to obtain informed consent from each participant or their guardians (for participants under 18) before initiating the recording process. This proactive step ensured the ethical handling of visual materials featuring the participants’ faces and adhered to responsible research practices. Following the workshops, we amassed a total of 4515 images, which were bifurcated into training and test datasets at an 80:20 ratio, yielding 3,612 images for the training set and 903 for the test dataset. This data will be employed to train the YOLOR model, advancing the objectives of this research endeavor.

After the conclusion of the workshop, the students presented their STEM projects which were deeply anchored in the field of the Internet of Things (IoT). Leveraging the knowledge they acquired during the workshop regarding sensors and programming concepts, the students were tasked with creating innovative solutions to

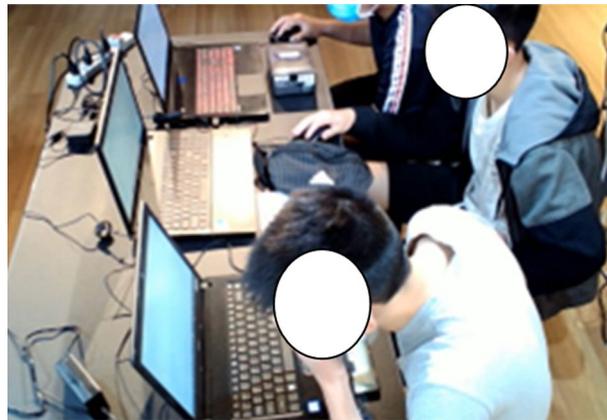


Fig. 2 The setting of camera angle

Table 1 Scorer reliability of the CPAM (Besemer, 1998)

Subscales	indicators	Scorer reliability
1. Novelty	1.1 Original	0.83 ^{***}
	1.2 Surprising	0.80 ^{***}
2. Resolution	2.1 Valuable	0.82 ^{***}
	2.2 Logical	0.79 ^{***}
	2.3 Useful	0.68 ^{***}
	2.4 Understandable	0.71 ^{***}
3. Elaboration & Synthesis	3.1 Organic	0.73 ^{***}
	3.2 Elegant	0.80 ^{***}
	3.3 Well-Crafted	0.84 ^{***}

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

real-world problems using IoT technologies. Their projects encompassed a range of ideas, including smart home solutions, energy-efficient systems, and automation processes that facilitate more sustainable living and working environments.

Each project was required to integrate sensor technologies to collect data and to use programming concepts to create a functional IoT system. The students utilized various sensors to gather data, and applied programming concepts to analyze the data and automate responses in the systems they developed. These projects encouraged students to think critically and creatively, pushing them to devise solutions that were both innovative and technically sound.

After the students presented their projects, they were evaluated by two experts in the STEM field. The evaluation process adhered to the criteria outlined in the Creative Product Analysis Matrix (CPAM) model, which encompassed three dimensions and nine scoring indicators, as detailed in Table 1. Drawing from the expertise of Besemer (1998), who verified the effectiveness of the CPAM model in a separate study, this evaluation method ensured a comprehensive and meticulous assessment of the students' project.

The scoring was conducted using a five-point Likert scale, a tool that facilitated a nuanced understanding of the strengths and weaknesses of each project. The

inter-rater reliability, gauging the level of agreement between the two experts, was substantiated by a correlation coefficient ranging between 0.68 and 0.84, pointing to a significant level of consensus in the scoring process. This metric not only attested to the coherence in the evaluation but also affirmed the credibility of the scores allocated, establishing a reliable foundation for the validation data needed for the STEM learning performance prediction model.

Real-time Automated STEM Engagement Detection System (RASEDS)

To address the highly dynamic and complex nature of learning in STEM activities, we developed a system, called Real-time Automated STEM Engagement Detection System (RASEDS) to evaluate the engagement level of learners automatically and instantaneously. The architectural diagram of the system is shown in Fig. 3. Initially, RASEDS employs object detection technology, specifically YOLOR, to identify both the hands of learners and all learning materials used in the activities. The interactions between learners’ hands and the learning materials are recognized and serve as a measure of the learners’ immediate behaviors. Ultimately, these behaviors are mapped to the ICAP framework to assess the engagement level of learners during STEM activities.

YOLOR

In this study, we adopted the YOLOR object detection model proposed by Wang et al. (2021) to identify learners’ hands and learning materials. YOLOR is one of the most robust models for object detection tasks currently available. It leverages the integration of tacit and explicit knowledge to learn universal representations, thereby significantly enhancing model performance with one ten-thousandth of the parameters and computational capacity, and executing inference for multiple computer vision tasks (Wang et al., 2021). However, since YOLOR’s pre-trained weights do not include commonly used learning materials in STEM activities, we employ Transfer Learning to retrain the YOLOR model specifically for identifying learning materials used in STEM activities. The training parameters used for this purpose are detailed in Table 2. As a result of this retraining process, the YOLOR model becomes capable of recognizing six objects

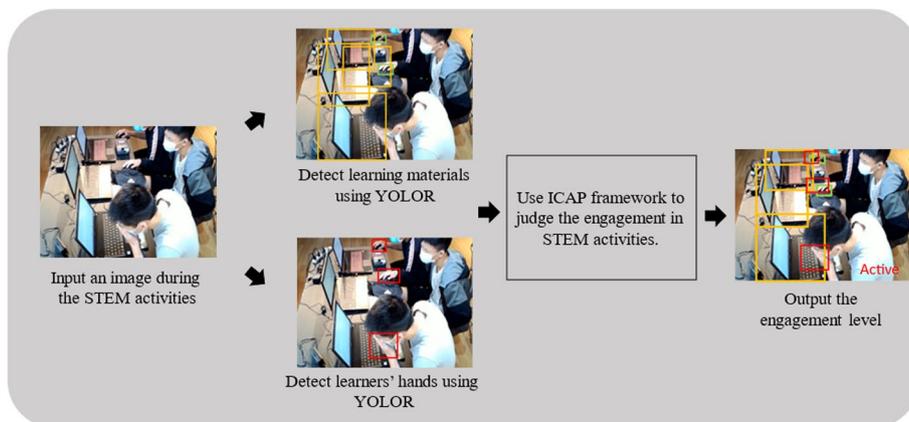


Fig. 3 The architecture of RASEDS

Table 2 The training parameters in YOLOR

Parameter	Batch_size	Epochs	Height	Width	Class
Value	16	400	640	640	6

Table 3 Relationship between ICAP framework and learning behaviors in STEM education

Indicator	Definition	Material being manipulated
Passive (P)	<ul style="list-style-type: none"> Learners' hands do not touch objects related to STEM activities 	
Active (A)	<ul style="list-style-type: none"> Learners actively operates tablet containing learning materials to solve their questions instead of passively receiving knowledge from instructors 	<ul style="list-style-type: none"> Tablet
Constructive (C)	<ul style="list-style-type: none"> Learners use laptop and mouse to complete project Learners assemble Raspberry Pi components to solve project 	<ul style="list-style-type: none"> Laptop Mouse Raspberry Pi
Interactive (I)	<ul style="list-style-type: none"> Learners assist their peers in solving problems and completing tasks and projects 	<ul style="list-style-type: none"> Hands
Other (O)	<ul style="list-style-type: none"> RASEDS fails to recognize learners' hands Learners use cellphone to do something unrelated to STEM workshop 	<ul style="list-style-type: none"> Cellphone

commonly encountered in STEM activities: hand, tablet, laptop, mouse, Raspberry Pi, and cellphone.

It should be noted that we seek to understand learners' current behaviors by identifying the interaction between the learner's hand and learning materials. Hence, the Intersection over Union (IoU) is used as a metric to judge the level of interaction between the learner's hand and the learning material (as Formula 1). IoU is a method for calculating overlap regions and is frequently employed to determine the degree of overlap between two bounding boxes in object detection tasks. In this study, the IoU threshold was set at 0.7. If the IoU exceeds 0.7, it indicates an interaction between the learner's hand and the given learning material. Ultimately, the learner's behaviors can be defined based on this interaction with the learning material. For instance, if there is an interaction between the learner's hand and the laptop, it can be inferred that the learner is currently using the computer.

$$IOU = \frac{\text{Learners' hand} \cap \text{learning materials}}{\text{Learners' hand} \cup \text{learning materials}} \tag{1}$$

ICAP framework

The RASEDS defines the learner's current behavior by identifying the interaction between the learner's hands and learning materials. As such, different learning behaviors are discerned as the learner's hands interact with various learning materials. By aligning the identified learning behaviors with the indicators of the ICAP framework, it is possible to clearly understand the learner's learning engagement and changes in their engagement levels. Table 3 provides a comparative chart between the ICAP framework and learning behaviors exhibited during STEM activities.

The setup of RASEDS

To establish RASEDS, we first need to clone and build the YOLOR project as described by Wong (2022). We utilized Python 3.9 and PyTorch 1.8.0 with operation system Ubuntu 20.04 for development. Additionally, we provide pseudocode to illustrate the functionality of RASEDS, which is depicted in Table 4. RASEDS employs YOLOR every 5 s to extract the learner's hand and all learning materials, thereby determining the learner's behavior during STEM activities. Whether the learner's hand interacts with the learning materials depends on whether the IoU value exceeds 0.7. The ICAP framework is adopted to correlate learning behaviors with engagement. RASEDS summarizes the engagement indicators within one minute (i.e., 12 entries), and outputs the most frequently occurring engagement indicator within the minute as the learner's engagement level for the past minute. In the end, each learner's engagement level in the STEM activities is recorded at a frequency of one minute and formatted as a.csv file. These records serve as the basis for subsequent adaptive learning material recommendations.

Table 4 The pseudocode of RASEDS

Algorithm: RASEDS

Input: *Video V*

Output: *MostLearnerEngagement MLE*, where $MLE = [I, C, A, P, O]$

1. Create an array LE
2. **repeat**
3. Capture an Image *I* from *V* every 5 seconds
4. Detect learning materials *M* by YOLOR
5. Detect learners' hands H_i by YOLOR, where *i* is the number of learners
6. **if** $IOU(H_i, M) \geq 0.7$ **then**
7. **if** $M == \text{tablet}$ **then**
8. $LE.append("A")$
9. **else if** $M == \text{laptop, mouse, raspberry pi}$ **then**
10. $LE.append("C")$
11. **else**
12. $LE.append("O")$
13. **endif**
14. **else**
15. **if** $IOU(H_i, H_j) \geq 0.7$, where $i \neq j$ **then**
16. $LE.append("I")$
17. **else**
18. $LE.append("P")$
19. **endif**
20. **endif**
21. **if** the length of LE = 12 **then**
22. Output the *MostLearnerEngagement MLE* in LE
23. Save *MLE* to a CSV file
24. Clean LE
25. **until** the end of the *V*

Recommendation mechanism for adaptive learning materials in STEM education

In order to provide adaptive learning materials based on learners’ engagement in STEM activities, we first established a STEM performance prediction model. This model uses the engagement levels identified by RASEDS to predict final learning outcomes. The STEM performance prediction model was developed based on data collected from 86 participants in STEM Workshop: Python and Raspberry Pi Practical Activity. RASEDS was used to analyze the participants’ engagement during the workshop, producing percentages for each engagement indicators for every learners. These engagement indicators percentages served as independent variables, with the project scores of each participant serving as dependent variables, in a multiple linear regression analysis. The initial regression formula is shown as Formula 2, where I, C, A, P, and O represent the percentages of Interaction, Construction, Active, Passive, and Other indicators, respectively. Here, we need to find a1, a2, a3, a4, a5, and b based on the data from the 86 workshop participants.

$$achievement = a_1I + a_2C + a_3A + a_4P + a_5O + b \tag{2}$$

During data collection, the five independent variables of I, C, A, P, and O were converted into percentages to represent the proportion of each engagement type, hence these variables have a linear relationship (i.e., I + C + A + P + O = 100). This necessitates the consideration of multicollinearity among the independent variables, requiring the use of specialized regression methods. Common regression models for this situation include Ridge Regression, Least Absolute Shrinkage and Selection Operator Regression (LASSO), and Elastic Net Regression. Ridge Regression mitigates the effects of multicollinearity by introducing a penalty term (L2 regularization) in the objective function, improving model stability and generalizability. LASSO Regression uses an L1 regularization term to conduct feature selection and reduce the impact of multicollinearity. Elastic Net Regression combines the features of Ridge Regression (L2 regularization) and LASSO Regression (L1 regularization), overcoming some of the limitations of LASSO Regression when dealing with highly correlated features, while retaining feature selection capabilities. Metrics such as Mean Squared Error (MSE), R-Squared, and Adjusted R-Squared were used to evaluate model performance. The results, shown in Table 5, suggest that Elastic Net Regression has the lowest MSE and the highest R-Squared, indicating the smallest prediction error and the highest model explanatory power. Thus, Elastic Net Regression was chosen as the STEM performance prediction model.

By substituting the parameters of the Elastic Net Regression model into Formula 2, we obtain Formula 3. Following this, we only need to input the percentage of each engagement indicators from any time period in the classroom into Formula 3 to predict the

Table 5 The performance of different STEM Performance Prediction Models

	MSE	R-Squared	Adjusted R-Squared
Ridge Regression	12.05	0.83	0.76
LASSO Regression	12.06	0.83	0.76
Elastic Net Regression	11.67	0.84	0.77

learner's performance during that period. As the CPAM is scored using a nine-item five-point Likert scale, we have chosen a score of 27 (all nine items scored as 3) as the boundary between high and low achievement.

$$\text{achievement} = 1.956 * I + 3.1188 * C - 0.354 * A - 3.500 * P - 1.091 * O + 29.794 \quad (3)$$

According to the self-efficacy theory proposed by Bandura et al. (1999), learners feel more motivated, interested, satisfied, and accomplished when they believe they can complete challenging tasks. Conversely, if tasks are too easy or too difficult, learners may feel bored, frustrated, or give up. Thus, based on the theory of self-efficacy, we provide more challenging adaptive learning materials for high-achieving learners and, conversely, simpler materials with more annotations for low-achieving learners to facilitate easier completion.

Experimental design

Participants

In this study, we enlisted 87 learners from the Department of Engineering Science at a university in southern Taiwan, all of whom were taking part in the Networks Embedded System and Application course spanning two semesters. Before initiating the study, we ensured to obtain informed consent from every participant to record and use videos that included their faces for the sole purpose of this research. This step was undertaken to adhere to ethical guidelines pertaining to privacy and consent. As described in "[STEM Workshop: Python and Raspberry Pi Practical Activity](#)" section, none of the learners had previously attended the "STEM Workshop: Python and Raspberry Pi Practical Activity." The participants were split into two groups: the Control Group (CG) comprising 41 learners from the first semester, and the Experimental Group (EG) with 46 learners from the second semester. The division was designed such that none of the participants were aware of the distinct treatments they were set to receive during the study. In the CG, all participants utilized a uniform set of learning materials, whereas the EG benefited from adaptive learning material recommendations, which were tailored based on individual engagement levels to aid in course completion. To ensure a fair experimental setup, a single instructor taught both groups and maintained a consistent classroom setting. The RASEDS system was employed in both settings to monitor fluctuations in learner engagement levels. The pivotal difference between the two groups was in the application of the data derived from the RASEDS system; while the CG's data was collected and archived, the EG's data actively informed adaptive learning material recommendations designed to enhance engagement levels.

Procedure

A quasi-experimental design was employed to examine whether adaptive learning materials recommendation via RASEDS in STEM education helps improve learners' engagement and self-efficacy. The experimental activities were carried out within the 'Networks Embedded System and Application' course over two semesters. The course was organized individually, but peer interactions and discussions were permitted during project creation. The course, which focuses on IoT and AI, encourages learners to apply their

software and hardware knowledge to address real-life problems, thus aligning with the core concepts of STEM education (as illustrated in Fig. 4). Each session of the course, lasting three hours per week, began with two hours of theoretical instruction and fundamental programming principles. These lessons formed the basis for the week’s project and ensured a prerequisite comprehension. The remaining hour emphasized practical project work, designed to apply and reinforce the principles of STEM.

However, the key distinguishing factor between the EG and CG was the methodology adopted for the delivery of learning materials in the final hour of the sessions. In the Experimental Group (EG), learning resources were not standardized; instead, they were tailored to individual achievement levels determined during the first two hours of each session. The principle behind the distribution of adaptive learning materials was as follows: after learners completed two hours of coursework, the RASEDS system would calculate individual engagement metric percentages for each learner, inputting them into Formula 3 to project anticipated achievement levels. Based on these projections, teachers would assign more challenging materials to those with scores above 27, and vice versa. This approach aims to enhance learning capacity and retain student interest by providing materials suited to each learner’s achievement level.

Conversely, the Control Group (CG) adhered to a more traditional approach, where all participants received standardized learning materials in the last hour, regardless of their individual achievement levels discerned in the initial two-hour period. This approach, while simpler, did not allow for the adaptive personalization facilitated in the EG, remaining static and neglecting the diverse achievement levels of the participants.

For instance, in week 3 of the course, the topic was "AI Application in IoT". For the first two hours, both groups received the same instructional material, covering the theoretical concepts of AI and its intersection with IoT, and basic programming principles involved in creating AI-based IoT applications. In the final hour, the EG received adaptive learning resources. Suppose a participant showed high achievement during the first two hours, asking questions about advanced AI algorithms for IoT. His adaptive learning material for the final hour might include a challenging coding exercise on implementing a neural network for an IoT device, accompanied by

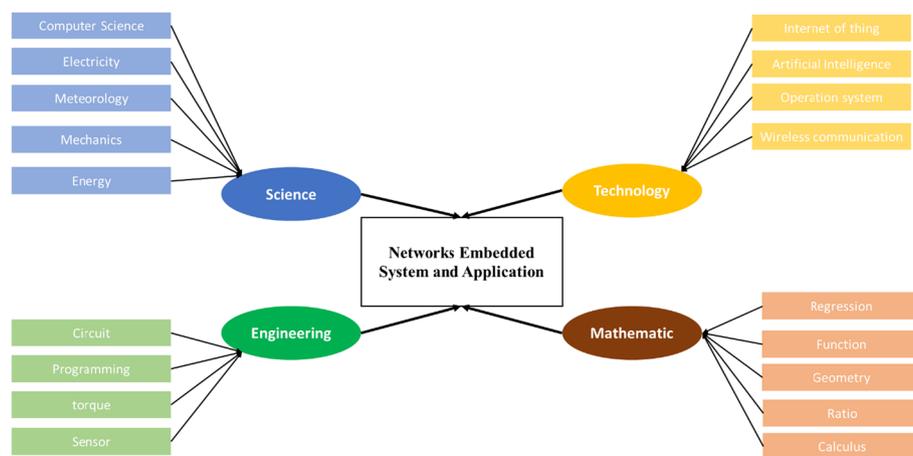


Fig. 4 STEM concepts in Networks Embedded System and Application

resources on best practices and advanced techniques. Meanwhile, another participant who struggled with the basic programming principles, would receive material focusing on reinforcing these fundamentals. His material might include a simpler coding exercise, along with additional explanations and examples to help solidify her understanding of the topic. On the other hand, for the CG, regardless of their individual achievement levels or difficulties during the initial two hours, all participants were provided with the same material in the final hour. This material was a standard one, providing a medium-difficulty coding exercise on implementing a basic AI algorithm for IoT, along with some generic resources. It was not tailored to the specific interests or struggles of any participant, unlike the adaptive approach employed with the EG. These different strategies embody the key divergence between the EG and CG – the former group experienced an adaptive, personalized learning approach based on their measured achievement levels, while the latter group did not.

The experiment ran for a total of five weeks, and both groups underwent pre- and post-tests. These tests measured two critical parameters: engagement and self-efficacy. The tests were conducted at the start of the activity (in the first week) and at the end (in the fifth week), as demonstrated in Fig. 5.

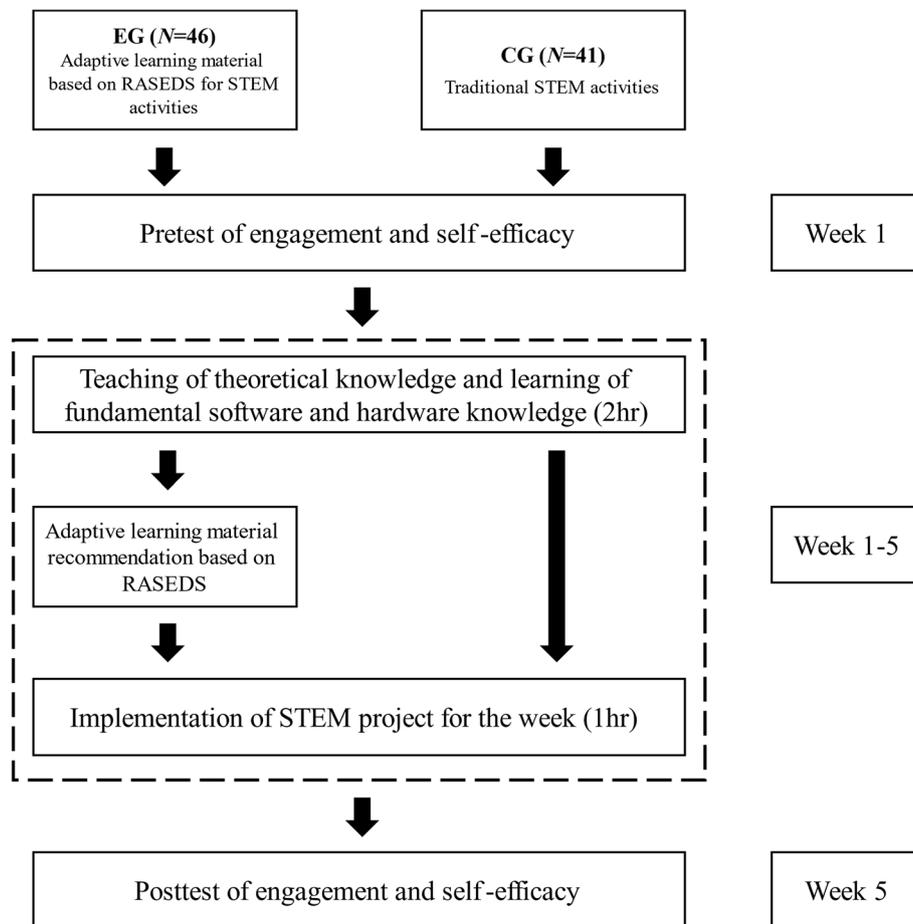


Fig. 5 Experimental procedure

Research tools

The learning engagement questionnaire used in this study was adapted from the Math and Science Engagement Scales proposed by Wang et al. (2016). The questionnaire divides engagement into four dimensions: cognitive, behavioral, emotional, and social. Cognitive engagement refers to self-regulated learning and the use of necessary cognitive strategies to understand complex ideas; behavioral engagement involves engagement in academic and classroom activities, the presence of positive behavior, and the absence of disruptive behavior. Emotional engagement is defined by the presence of positive emotional responses towards teachers, peers, and classroom activities, as well as interest in and value placed on the learning content. Social engagement denotes the quality of social interaction with peers and the willingness to establish and maintain relationships during the learning process (Wang et al., 2016). The questionnaire was designed as a five-point Likert scale and has been proven to have high reliability and validity in Wang et al. (2016). To suit this research, we translated the questionnaire into Chinese and conducted another reliability analysis. Table 6 presents the original and revised reliability of the questionnaire. The results show that the revised reliability all exceed 0.7, indicating sufficiently high reliability (Nunnally, 1978).

The New General Self-Efficacy Scale proposed by Chen et al. (2001) was used in this study. This scale, revised from the General Self-Efficacy Scale by Schwarzer and Jerusalem (1995), addresses concerns of low content validity and multidimensionality. Self-efficacy, as defined by the scale, is the belief in one's capacity to mobilize motivation, cognitive resources, and actions to meet specific situational demands. Essentially, self-efficacy is akin to confidence, characterized by a learner's belief in their ability to perform effectively within an academic setting. The reliability and validity of this scale were previously affirmed by Chen et al. (2001). Based on a five-point Likert scale, this questionnaire has demonstrated high reliability and validity in prior research. For the purposes of the current study, the scale was translated, followed by an additional reliability analysis. The resulting reliability coefficient was 0.88, indicating a high degree of reliability, consistent with the standards outlined by Nunnally (1978).

Results

The performance of Real-time Automated STEM Engagement Detection System (RASEDS)

To understand the performance of RASEDS in identifying the engagement of STEM learners, we first employed the confusion matrix, a table used to describe the performance of a classification model (or "classifier") on a set of data for which the true values are known. It presents the true positives, true negatives, false positives, and false negatives, allowing for a more detailed analysis of the system's performance in recognizing various engagement indicators (i.e., I, C, A, P, O). Following this, we calculated

Table 6 Reliability analysis of engagement scale

	Cognitive engagement	Behavioral engagement	Emotional engagement	Social engagement
Original reliability	0.75	0.82	0.89	0.74
Revised reliability	0.74	0.84	0.88	0.75

RASEDS's precision, recall, and F1 score, metrics derived from the confusion matrix, to assess RASEDS's performance. Here, the precision is the number of true positives divided by the number of true positives and false positives, indicating the proportion of correctly identified positive observations. The recall, also known as sensitivity or true positive rate, is the number of true positives divided by the number of true positives and the number of false negatives, showing the ability of the system to find all the positive samples. The F1 score is the harmonic mean of the precision and recall, where an F1 score reaches its best value at 1 (perfect precision and recall) and worst at 0, giving a well-rounded view of the model's accuracy. Due to the current lack of a comparative benchmark, we compared the outputs of RASEDS with the encodings of two experts for the same data and calculated the Cohen's kappa value. If the output of RASEDS has a high enough consistency with the encodings of the two experts, it means that RASEDS has high enough accuracy to replace expert encoding. Moreover, as RASEDS is an automatic and real-time encoding solution, it greatly reduces the time and manpower costs of expert encoding.

The results of RASEDS's confusion matrix are shown in Fig. 6. For model validation, we randomly selected 474 one-minute video clips from the learning videos of the control group, resulting in a total of 474 data points for confusion matrix analysis. The reason for using the control group's learning videos for validation rather than the workshop learning videos described in "[STEM Workshop: Python and Raspberry Pi Practical Activity](#)" section is that RASEDS was trained through the workshop data. Therefore, to enhance the credibility of model validation, it is necessary to avoid using the same or similar data for validation. As can be seen from Fig. 6, the model did not exhibit any

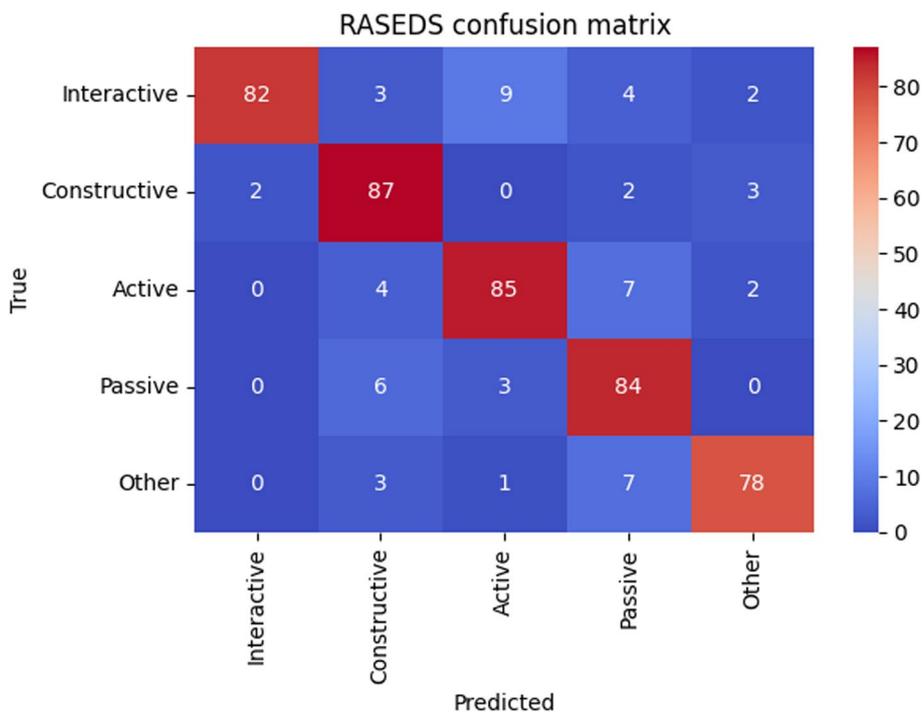


Fig. 6 The confusion matrix of RASEDS

noticeable misjudgments in its recognitions, with only a few possible misjudgments caused by angle or obstruction issues affecting RASEDS.

Based on the results of the confusion matrix in Fig. 6, we can calculate the precision, recall, and F1 score of RASEDS for each engagement level. The results are shown in Table 7. The average precision, average recall, and average F1 score of RASEDS are 0.883, 0.879, and 0.878, respectively.

Given the current lack of a benchmark for comparing the performance of RASEDS, we calculated Cohen’s kappa to compare the output of RASEDS and expert encoding. Two experts and RASEDS encoded the same 10 ten-minute learning videos. The encoding frequency of the experts was the same as that of RASEDS, with an encoding required every minute, resulting in 100 encoding opportunities. The resulting Cohen’s kappa between Expert A and Expert B was 0.82, between Expert A and RASEDS was 0.85, and between Expert B and RASEDS was 0.81, all exceeding 0.70 (Landis & Koch, 1977). These results indicate that the reliability between raters was sufficiently high; that is, there was no difference between expert encoding and RASEDS encoding. Therefore, to answer research question 1, RASEDS can effectively and accurately measure learners’ engagement, and it can automate encoding while achieving the same accuracy as expert encoding, significantly improving the time and manpower costs required by traditional observation methods.

The impact of adaptive learning material recommendation via RASEDS on engagement in STEM education

To investigate the effect of adaptive learning materials recommendations via RASEDS on learners’ engagement in STEM activities, we used ANCOVA analysis. The pre-test scores of engagement were used as covariates, and the post-test scores of engagement as dependent variables. The homogeneity of variances was assessed using Levene’s test. The results indicated homogeneity for cognitive engagement ($F = 0.03, p = 0.862$), behavioral engagement ($F = 2.44, p = 0.122$), emotional engagement ($F = 0.575, p = 0.450$), and social engagement ($F = 1.17, p = 0.282$). Therefore, the results confirmed the robustness of variance equality, and ANCOVA was deemed suitable for use.

Descriptive statistics of engagement and ANCOVA analysis results are shown in Tables 8 and 9, respectively. As seen in Table 9, there were significant differences in cognitive ($F = 21.86, P < 0.001$), behavioral ($F = 43.5, P < 0.001$), and emotional ($F = 26.81, P < 0.001$) engagement between the experimental group and the control group. Furthermore, according to Table 8, the post-test scores of cognitive, behavioral, and emotional engagement were significantly higher in the experimental group than in the control group. Therefore, the results suggest that introducing adaptive learning materials

Table 7 The performance of RASEDS

	Interactive	Constructive	Active	Passive	Other	Average
Precision	0.867	0.845	0.976	0.918	0.808	0.883
Recall	0.867	0.926	0.82	0.876	0.903	0.879
F1 score	0.867	0.883	0.891	0.897	0.853	0.878

Table 8 Descriptive results for engagement

	EG (N = 46)				CG (N = 41)			
	Pretest		Posttest		Pretest		Posttest	
	M	SD	M	SD	M	SD	M	SD
Cognitive engagement	15.8	2.58	19.8	2.85	15.8	2.68	17.0	2.92
Behavioral engagement	17.8	2.45	18.7	2.63	17.2	1.91	15.5	1.78
Emotional engagement	16.0	2.32	18.8	2.89	15.6	2.62	15.8	3.03
Social engagement	16.1	2.28	17.7	2.24	15.6	2.57	17.2	2.05

Table 9 ANCOVA results for engagement

Variable	SS	df	Mean Square	F	p	Partial η ²
Cognitive engagement	171.3	1	171.3	21.86	< 0.001 ***	0.206
Behavioral engagement	201.4	1	201.4	43.5	< 0.001 ***	0.341
Emotional engagement	214.9	1	214.94	26.81	< 0.001 ***	0.242
Social engagement	7.78	1	7.78	1.79	0.185	0.021

Bold values represent significant difference

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10 Descriptive results for self-efficacy

	EG (N = 46)				CG (N = 41)			
	Pretest		Posttest		Pretest		Posttest	
	M	SD	M	SD	M	SD	M	SD
Self-efficacy	26.6	3.57	30.0	4.57	26.1	5.84	25.3	4.85

recommendation via RASEDS in STEM activities can significantly improve learners’ engagement, particularly in cognitive, behavioral, and emotional aspects.

The impact of adaptive learning material recommendation via RASEDS on self-efficacy in STEM education

In order to understand the effect of adaptive learning materials recommendations via RASEDS on learners’ self-efficacy in STEM activities, we also used ANCOVA analysis. The pre-test scores of self-efficacy were used as covariates, and the post-test scores of self-efficacy as dependent variables. The homogeneity of variances was assessed using Levene’s test. The results indicated homogeneity for self-efficacy ($F = 0.05$, $p = 0.831$). Therefore, the results confirmed the robustness of variance equality, and ANCOVA was again deemed suitable for use.

The descriptive statistics and ANCOVA analysis results of self-efficacy are shown in Tables 10 and 11 respectively. As seen in Table 11, there was a significant difference in self-efficacy ($F = 26.5$, $P < 0.001$) between the experimental group and the control group. Furthermore, according to Table 10, the post-test scores of self-efficacy were significantly higher in the experimental group than in the control group. Therefore, the results suggest that introducing adaptive learning material recommendation via RASEDS in STEM activities can significantly improve learners’ self-efficacy.

Table 11 ANCOVA results for self-efficacy

Variable	SS	df	Mean Square	F	p	Partial η^2
Self-efficacy	448	1	448	26.5	<0.001***	0.240

Bold values represent significant difference

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Discussion

Adaptive learning material recommendation via RASEDS in STEM activities

In this study, we propose a system, called Real-time Automated STEM Engagement Detection System (RASEDS) for automatically and objectively understanding learners' engagement in STEM activities. Compared with traditional methods of assessing learners' engagement in STEM education, such as self-reporting and observation method, RASEDS leverages AI and computer vision to measure learners' engagement in STEM education in a standardized and fair manner, potentially overcoming the limitations of self-reporting. Furthermore, as an automated system, RASEDS reduces the time and labor required for expert coding in observation method. Our findings suggest that RASEDS performs excellently in detecting STEM engagement and can achieve identification results similar to those of expert coding. RASEDS effectively addresses the lack of automated assessment tools for STEM education mentioned in systematic review of Gao et al. (2020), providing insights for assessing engagement in STEM education.

On the other hand, most of the current research on adaptive learning is focused on e-learning environments because researchers can easily obtain learners' learning trajectories and develop corresponding adaptive learning mechanisms (El-Sabagh, 2021; Premlatha & Geetha, 2015). However, STEM education often involves laboratory activities and hands-on experiences, making it challenging to develop adaptive learning mechanisms in STEM education (Chang & Chen, 2022; Lin et al., 2021). Therefore, based on the engagement results identified by RASEDS, we predict learners' performance in STEM activities and uses this as a basis for recommending adaptive learning materials. Ultimately, it provides insights for adaptive learning in STEM education.

The impact of adaptive learning material recommendation via RASEDS on engagement in STEM education

In the evolving landscape of STEM education, empowering learners to steer their educational journey is becoming increasingly pivotal. The shift from a teacher-centered approach to a learner-centric paradigm necessitates tools that can facilitate effective learner engagement (Fang et al., 2022; Li et al., 2020). This study sought to address this gap through the development of RASEDS, a real-time student engagement monitoring system equipped with artificial intelligence and data analysis capabilities to recommend adaptive learning materials during STEM activities.

Our findings delineated in Tables 8 and 9 validate the efficacy of RASEDS in enhancing cognitive, behavioral, and emotional engagement, converging with a burgeoning body of literature emphasizing the potential of adaptive learning in fostering knowledge construction (Mou et al., 2022; Xie et al., 2019). RASEDS affords a nuanced understanding of learners' engagement levels, thereby guiding them to materials congruent with their

learning phase, a strategy echoed in other studies (Sein, 2022). This harmonized learning pathway augments cognitive engagement, corroborating earlier research underscoring the significance of aligned learning materials in bolstering cognitive strategies (Wu et al., 2023).

Moreover, the mitigation of learning interruptions encountered during challenging phases stands as a testament to RASEDS' potential in fostering behavioral engagement. This resonates with prior works highlighting the role of adaptive learning in sustaining students' zest for learning, thereby preventing early disengagement due to perceived difficulties (El-Sabagh, 2021; Ross et al., 2018).

Furthermore, the favorable shift in learners' attitude and emotional response towards STEM activities underscore the emotional benefits reaped through adaptive learning systems like RASEDS. This is buttressed by prior research spotlighting the positive repercussions of adaptive learning on learners' emotions and attitudes (Fatahi, 2019; Martin et al., 2020; Megahed & Mohammed, 2020). Encouragingly, the nurtured proactive learning stance fostered by RASEDS finds echoes in studies that advocate for a tailored learning approach in reducing frustration and cultivating a positive learning ambiance (Amin et al., 2023; Standen et al., 2020).

In conclusion, this study furthers the discourse on the instrumental role of adaptive learning systems in advancing STEM education. Through the lens of RASEDS, it becomes manifest that real-time engagement monitoring paired with adaptive learning material recommendations can be a linchpin in facilitating a holistic learning environment, nurturing cognitive, behavioral, and emotional engagement. Future studies may delve deeper, exploring the multifaceted dimensions of learner engagement to pave the way for a richer, more interactive, and learner-centric STEM education landscape.

The impact of adaptive learning material recommendation via RASEDS on self-efficacy in STEM education

Self-efficacy, a term coined by Bandura (1977) and later elaborated on by Bandura and Watts (1996), refers to a learner's belief in their ability to achieve their objectives. This concept, which centers around individuals' confidence and expectations regarding their capacities, is particularly pivotal in STEM education where a student-centered approach is predominant (Kuchynka et al., 2021). The self-efficacy demonstrated by students in STEM education has a direct bearing on their learning outcomes and their sustained interest in participating (Luo et al., 2021). Yet, the intricate nature of STEM subjects can sometimes be a double-edged sword, potentially dampening self-efficacy when students encounter hurdles, thereby affecting their academic performance (Luo et al., 2021).

To counter this, we propose the utilization of RASEDS, a system designed for the real-time monitoring of student engagement, thereby facilitating the recommendation of adaptive learning materials tailored to individual needs. As reflected in the data presented in Tables 10 and 11, leveraging RASEDS significantly amplifies self-efficacy during STEM activities. It is vital to emphasize the symbiotic relationship between self-efficacy and engagement — a surge in one invariably promotes growth in the other. Higher engagement translates to active participation and a deeper comprehension of the subject matter, subsequently fostering a stronger sense of self-assuredness.

This augmentation in self-efficacy corresponds directly to the enhanced engagement observed amongst students immersed in STEM tasks (Han et al., 2021; Kuchynka et al., 2021).

By offering learning materials fine-tuned to suit learners' aptitudes, RASEDS alleviates the challenges posed by potentially overwhelming obstacles, nurturing not only a deeper engagement with STEM topics but also fortifying students' confidence in handling STEM tasks, thereby reinforcing self-efficacy. This echoes previous studies that advocate for adaptive learning as a means to synergistically bolster confidence and self-efficacy through heightened engagement (Graham, 2022; Seon Ahn & Bong, 2019).

In conclusion, although the inherent difficulties of STEM education can pose a threat to students' self-efficacy (Luo et al., 2021), our study illuminates the rehabilitative power of adaptive learning interventions. RASEDS emerges as a formidable asset in this endeavor, fostering a conducive learning ecosystem that encourages confidence and fosters self-efficacy. While the results are promising, it remains essential to substantiate these initial findings through ongoing research, aiming to deepen our understanding and to carve pathways for more nuanced, learner-focused strategies in STEM education.

Conclusion

We aim to develop a system, called Real-time Automated STEM Engagement Detection System (RASEDS), based on computer vision and the ICAP framework, and to examine the impact of adaptive learning material recommendation via RASEDS on students' engagement and self-efficacy in STEM activities. The main findings and contributions of this research are as follows:

- RASEDS effectively identifies students' engagement in STEM activities by recognizing the interaction between their hand and learning materials (using YOLOR), and mapping these to the four modes of the ICAP framework.
- By recommending adaptive learning materials via RASEDS, it can enhance learners' engagement and self-efficacy in STEM activities by providing adaptive support and learning materials according to their learning needs and preferences.

We demonstrate the potential of integrating AI technologies and educational theories to support adaptive learning in STEM education. It also provides a novel and practical approach for measuring and enhancing the learning process and outcomes in STEM activities. However, this study has some limitations. Firstly, the small sample size of participants in the experiment ($N = 87$) may affect the validity of the statistical analysis. Since RASEDS is based on computer vision technology, the system is limited by the camera angle and the problem of occlusion, which leads to misrecognition.

Despite these limitations, we contribute to the literature on adaptive learning in STEM education by developing and evaluating a system, called RASEDS based on computer vision and the ICAP framework. Future research can verify the causal effect of learners' engagement on self-efficacy through more rigorous experimental designs and apply this system to different types of STEM activities and environments to examine its robustness and scalability. Furthermore, future work can expand the findings of this study by applying RASEDS to different STEM fields and contexts, exploring other factors that affect

learners' engagement and self-efficacy, and evaluating the long-term impact of RASEDS on students' STEM literacy and career aspirations.

Acknowledgements

We are extremely grateful to the research assistants and students who participated in this study.

Author contributions

T-TW is the leader of this research, he is in charge of the research design, conducting teaching and learning experiment, data analysis. H-YL is responsible for assisting in the conduct of experiments and surveying related literature, writing the manuscript, and proofreading the manuscript. W-SW is responsible for assisting in the conduct of experiments and surveying related literature. C-JL is responsible for assisting in the conduct of experiments. Y-MH is responsible for designing research experiments, providing fundamental education theories and comments to this research, and he is also responsible for revising the manuscript. All authors spent more than 2 months to discuss and analyze the data. The author(s) read and approved the final manuscript.

Funding

This project was funded by the National Science and Technology Council (NSTC), of the Republic of China under Contract numbers NSTC 109-2511-H-006-011-MY3, NSTC 110-2511-H-006012-MY3, NSTC 110-2511-H-006-008-MY3.

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 14 June 2023 Accepted: 15 September 2023

Published online: 29 September 2023

References

- AfiniNormadhi, N. B., Shuib, L., Md Nasir, H. N., Bimba, A., Idris, N., & Balakrishnan, V. (2019). Identification of personal traits in adaptive learning environment: Systematic literature review. *Computers & Education*, *130*, 168–190. <https://doi.org/10.1016/j.compedu.2018.11.005>
- Amin, S., Uddin, M. I., Alarood, A. A., Mashwani, W. K., Alzahrani, A., & Alzahrani, A. O. (2023). Smart E-learning framework for personalized adaptive learning and sequential path recommendations using reinforcement learning. *IEEE Access*, *11*, 89769–89790. <https://doi.org/10.1109/ACCESS.2023.3305584>
- Bai, S., Hew, K. F., Sailer, M., & Jia, C. (2021). From top to bottom: How positions on different types of leaderboard may affect fully online student learning performance, intrinsic motivation, and course engagement. *Computers & Education*, *173*, 104297. <https://doi.org/10.1016/j.compedu.2021.104297>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A., Freeman, W. H., & Lightsey, R. (1999). Self-efficacy: The exercise of control. *The Journal of Cognitive Psychotherapy*. <https://doi.org/10.1891/0889-8391.13.2.158>
- Bandura, A., & Watts, R. E. (1996). Self-efficacy in changing societies. *The Journal of Cognitive Psychotherapy*. <https://doi.org/10.1891/0889-8391.10.4.313>
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: What ever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. <https://doi.org/10.1111/j.1745-6916.2007.00051>
- Besemer, S. P. (1998). Creative product analysis matrix: Testing the model structure and a comparison among products-three novel chairs. *Creativity Research Journal*, *11*(4), 333–346. https://doi.org/10.1207/s15326934crj1104_7
- Borda, E., Schumacher, E., Hanley, D., Geary, E., Warren, S., Ipsen, C., & Stredicke, L. (2020). Initial implementation of active learning strategies in large, lecture STEM courses: Lessons learned from a multi-institutional, interdisciplinary STEM faculty development program. *International Journal of STEM Education*, *7*(1), 4. <https://doi.org/10.1186/s40594-020-0203-2>
- Chang, C.-C., & Chen, Y. (2022). Using mastery learning theory to develop task-centered hands-on STEM learning of Arduino-based educational robotics: Psychomotor performance and perception by a convergent parallel mixed method. *Interactive Learning Environments*, *30*(9), 1677–1692. <https://doi.org/10.1080/10494820.2020.1741400>
- Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, *4*(1), 62–83. <https://doi.org/10.1177/109442810141004>
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Christensen, R., Knezek, G., & Tyler-Wood, T. (2015). Alignment of Hands-on STEM Engagement Activities with Positive STEM Dispositions in Secondary School Students. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-015-9572-6>
- Cook, C. R., Kilgus, S. P., & Burns, M. K. (2018). Advancing the science and practice of precision education to enhance student outcomes. *Journal of School Psychology*, *66*, 4–10. <https://doi.org/10.1016/j.jsp.2017.11.004>

- D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist*, *52*(2), 104–123.
- El-Sabagh, H. A. (2021). Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *International Journal of Educational Technology in Higher Education*, *18*(1), 53. <https://doi.org/10.1186/s41239-021-00289-4>
- Fang, J.-W., He, L.-Y., Hwang, G.-J., Zhu, X.-W., Bian, C.-N., & Fu, Q.-K. (2022). A concept mapping-based self-regulated learning approach to promoting students' learning achievement and self-regulation in STEM activities. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2022.2061013>
- Fatahi, S. (2019). An experimental study on an adaptive e-learning environment based on learner's personality and emotion. *Education and Information Technologies*, *24*(4), 2225–2241. <https://doi.org/10.1007/s10639-019-09868-5>
- Flanagan, B., Majumdar, R., & Ogata, H. (2022). Early-warning prediction of student performance and engagement in open book assessment by reading behavior analysis. *International Journal of Educational Technology in Higher Education*, *19*(1), 41. <https://doi.org/10.1186/s41239-022-00348-4>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. <https://doi.org/10.3102/00346543074001059>
- Gao, X., Li, P., Shen, J., & Sun, H. (2020). Reviewing assessment of student learning in interdisciplinary STEM education. *International Journal of STEM Education*, *7*(1), 24. <https://doi.org/10.1186/s40594-020-00225-4>
- Graham, S. (2022). Self-efficacy and language learning—what it is and what it isn't. *The Language Learning Journal*, *50*(2), 186–207. <https://doi.org/10.1080/09571736.2022.2045679>
- Guerrero-Roldán, A.-E., Rodríguez-González, M. E., Bañeres, D., Elasmri-Ejjaber, A., & Cortada, P. (2021). Experiences in the use of an adaptive intelligent system to enhance online learners' performance: A case study in Economics and Business courses. *International Journal of Educational Technology in Higher Education*, *18*(1), 36. <https://doi.org/10.1186/s41239-021-00271-0>
- Han, J., Kelley, T., & Knowles, J. G. (2021). Factors influencing student STEM learning: Self-efficacy and outcome expectancy, 21st century skills, and career awareness. *Journal for STEM Education Research*, *4*(2), 117–137. <https://doi.org/10.1007/s41979-021-00053-3>
- Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences*, *18*, 83–90.
- Hsiao, J.-C., Chen, S.-K., Chen, W., & Lin, S. S. J. (2022). Developing a plugged-in class observation protocol in high-school blended STEM classes: Student engagement, teacher behaviors and student-teacher interaction patterns. *Computers & Education*, *178*, 104403. <https://doi.org/10.1016/j.compedu.2021.104403>
- Hu, Y.-H. (2022). Effects and acceptance of precision education in an AI-supported smart learning environment. *Education and Information Technologies*, *27*(2), 2013–2037. <https://doi.org/10.1007/s10639-021-10664-3>
- Huang, Y., & Wang, S. (2023). How to motivate student engagement in emergency online learning? Evidence from the COVID-19 situation. *Higher Education*, *85*(5), 1101–1123. <https://doi.org/10.1007/s10734-022-00880-2>
- Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, *2*, 100017. <https://doi.org/10.1016/j.caeai.2021.100017>
- Kuchynka, S., Reifsteck, T. V., Gates, A. E., & Rivera, L. M. (2021). Developing self-efficacy and behavioral intentions among underrepresented students in STEM: The role of active learning [Original Research]. *Frontiers in Education*. <https://doi.org/10.3389/educ.2021.668239>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174.
- Lee, H.-Y., Cheng, Y.-P., Wang, W.-S., Lin, C.-J., & Huang, Y.-M. (2023a). Exploring the learning process and effectiveness of STEM education via learning behavior analysis and the interactive-constructive-active-passive framework. *Journal of Educational Computing Research*, *61*(5), 951–976. <https://doi.org/10.1177/07356331221136888>
- Lee, H.-Y., Lin, C.-J., Wang, W.-S., Chang, W.-C., & Huang, Y.-M. (2023b). Precision education via timely intervention in K-12 computer programming course to enhance programming skill and affective-domain learning objectives. *International Journal of STEM Education*, *10*(1), 52. <https://doi.org/10.1186/s40594-023-00444-5>
- Li, S., Du, H., Xing, W., Zheng, J., Chen, G., & Xie, C. (2020). Examining temporal dynamics of self-regulated learning behaviors in STEM learning: A network approach. *Computers & Education*, *158*, 103987. <https://doi.org/10.1016/j.compedu.2020.103987>
- Lin, K.-Y., Lu, S.-C., Hsiao, H.-H., Kao, C.-P., & Williams, P. J. (2021). Developing student imagination and career interest through a STEM project using 3D printing with repetitive modeling. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2021.1913607>
- Liu, S., Liu, S., Liu, Z., Peng, X., & Yang, Z. (2022). Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement. *Computers & Education*, *181*, 104461. <https://doi.org/10.1016/j.compedu.2022.104461>
- Liu, T. C. (2022). A case study of the adaptive learning platform in a Taiwanese elementary school: Precision education from teachers' perspectives. *Education and Information Technologies*, *27*(5), 6295–6316. <https://doi.org/10.1007/s10639-021-10851-2>
- Luo, T., So, W. W. M., Wan, Z. H., & Li, W. C. (2021). STEM stereotypes predict students' STEM career interest via self-efficacy and outcome expectations. *International Journal of STEM Education*, *8*(1), 36. <https://doi.org/10.1186/s40594-021-00295-y>
- Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education*, *6*(1), 2. <https://doi.org/10.1186/s40594-018-0151-2>
- Marshall, P. (2007). *Do tangible interfaces enhance learning?* Proceedings of the 1st international conference on Tangible and embedded interaction, Baton Rouge, Louisiana. <https://doi.org/10.1145/1226969.1227004>
- Martin, F., Chen, Y., Moore, R. L., & Westine, C. D. (2020). Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. *Educational Technology Research and Development*, *68*(4), 1903–1929. <https://doi.org/10.1007/s11423-020-09793-2>
- Megahed, M., & Mohammed, A. (2020). Modeling adaptive E-Learning environment using facial expressions and fuzzy logic. *Expert Systems with Applications*, *157*, 113460. <https://doi.org/10.1016/j.eswa.2020.113460>

- Miller, K., Sonnert, G., & Sadler, P. (2018). The influence of students' participation in STEM competitions on their interest in STEM careers. *International Journal of Science Education, Part B*, 8(2), 95–114. <https://doi.org/10.1080/21548455.2017.1397298>
- Mirata, V., Hirt, F., Bergamin, P., & van der Westhuizen, C. (2020). Challenges and contexts in establishing adaptive learning in higher education: Findings from a Delphi study. *International Journal of Educational Technology in Higher Education*, 17(1), 32. <https://doi.org/10.1186/s41239-020-00209-y>
- Mou, C., Tian, Y., Zhang, F., & Zhu, C. (2022). Current situation and strategy formulation of college sports psychology teaching following adaptive learning and deep learning under information education [Original Research]. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2021.766621>
- Nunnally, J. C. (1978). An overview of psychological measurement. In B. B. Wolman (Ed.), *Clinical diagnosis of mental disorders: A handbook* (pp. 97–146). Springer US. https://doi.org/10.1007/978-1-4684-2490-4_4
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of Research Methods in Personality Psychology*, 1(2007), 224–239.
- Peng, H., Ma, S., & Spector, J. M. (2019). Personalized adaptive learning: An emerging pedagogical approach enabled by a smart learning environment. *Smart Learning Environments*, 6(1), 9. <https://doi.org/10.1186/s40561-019-0089-y>
- Premlatha, K. R., & Geetha, T. V. (2015). Learning content design and learner adaptation for adaptive e-learning environment: A survey. *Artificial Intelligence Review*, 44(4), 443–465. <https://doi.org/10.1007/s10462-015-9432-z>
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231. <https://doi.org/10.1002/j.2168-9830.2004.tb00809.x>
- Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, S. A., & Yousufi, S. Q. (2021). Factors affecting students' learning performance through collaborative learning and engagement. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2021.1884886>
- Raković, M., Marzouk, Z., Liaqat, A., Winne, P. H., & Nesbit, J. C. (2020). Fine grained analysis of students' online discussion posts. *Computers & Education*, 157, 103982. <https://doi.org/10.1016/j.compedu.2020.103982>
- Reinholz, D. L., White, I., & Andrews, T. (2021). Change theory in STEM higher education: A systematic review. *International Journal of STEM Education*, 8(1), 37. <https://doi.org/10.1186/s40594-021-00291-2>
- Ross, B., Chase, A.-M., Robbie, D., Oates, G., & Absalom, Y. (2018). Adaptive quizzes to increase motivation, engagement and learning outcomes in a first year accounting unit. *International Journal of Educational Technology in Higher Education*, 15(1), 30. <https://doi.org/10.1186/s41239-018-0113-2>
- Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. *J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs*, 35, 37.
- Sein, M. (2022). AI-assisted knowledge assessment techniques for adaptive learning environments. *Computers and Education: Artificial Intelligence*, 3, 100050. <https://doi.org/10.1016/j.caeai.2022.100050>
- Seon Ahn, H., & Bong, M. (2019). Self-Efficacy in Learning: Past, Present, and Future. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge Handbook of Motivation and Learning* (pp. 63–86). Cambridge University Press. <https://doi.org/10.1017/9781316823279.005>
- Sithole, A., Chiyaka, E. T., McCarthy, P., Mupinga, D. M., Bucklein, B. K., & Kibirige, J. (2017). Student attraction, persistence and retention in STEM programs: Successes and continuing challenges. *Higher Education Studies*, 7(1), 46–59. <https://doi.org/10.5539/hes.v7n1p46>
- Smith Iv, D. H., Hao, Q., Dennen, V., Tsikerdekis, M., Barnes, B., Martin, L., & Tresham, N. (2020). Towards Understanding Online Question & Answer Interactions and their effects on student performance in large-scale STEM classes. *International Journal of Educational Technology in Higher Education*, 17(1), 20. <https://doi.org/10.1186/s41239-020-00200-7>
- Standen, P. J., Brown, D. J., Taheri, M., Galvez Trigo, M. J., Boulton, H., Burton, A., Hallewell, M. J., Lathe, J. G., Shopland, N., Blanco Gonzalez, M. A., Kwiatkowska, G. M., Milli, E., Cobello, S., Mazzucato, A., Traversi, M., & Hortal, E. (2020). An evaluation of an adaptive learning system based on multimodal affect recognition for learners with intellectual disabilities. *British Journal of Educational Technology*, 51(5), 1748–1765. <https://doi.org/10.1111/bjet.13010>
- Wang, C.-Y., Yeh, I.-H., & Liao, H.-Y. M. (2021). You only learn one representation: Unified network for multiple tasks. *arXiv preprint arXiv:2105.04206*. <https://doi.org/10.48550/arXiv.2105.04206>
- Wang, L.-H., Chen, B., Hwang, G.-J., Guan, J.-Q., & Wang, Y.-Q. (2022). Effects of digital game-based STEM education on students' learning achievement: A meta-analysis. *International Journal of STEM Education*, 9(1), 26. <https://doi.org/10.1186/s40594-022-00344-0>
- Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The Math and Science Engagement Scales: Scale development, validation, and psychometric properties. *Learning and Instruction*, 43, 16–26. <https://doi.org/10.1016/j.learninstruc.2016.01.008>
- Wang, S., Christensen, C., Cui, W., Tong, R., Yarnall, L., Shear, L., & Feng, M. (2023). When adaptive learning is effective learning: Comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*, 31(2), 793–803. <https://doi.org/10.1080/10494820.2020.1808794>
- Wong, B., Chiu, Y.-L.T., Murray, O. M., & Horsburgh, J. (2022). End of the road? The career intentions of under-represented STEM students in higher education. *International Journal of STEM Education*, 9(1), 51. <https://doi.org/10.1186/s40594-022-00366-8>
- Wong, K. (2022, May 5). *YOLOR*. Retrieved December 26, from <https://github.com/WongKinYiu/yolor>
- Wu, T.-T., Lee, H.-Y., Li, P.-H., Huang, C.-N., & Huang, Y.-M. (2023). Promoting self-regulation progress and knowledge construction in blended learning via ChatGPT-based learning aid. *Journal of Educational Computing Research*. <https://doi.org/10.1177/07356331231191125>
- Xie, H., Chu, H.-C., Hwang, G.-J., & Wang, C.-C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, 140, 103599. <https://doi.org/10.1016/j.compedu.2019.103599>
- Zhang, Z., Li, Z., Liu, H., Cao, T., & Liu, S. (2019). Data-driven Online Learning Engagement Detection via Facial Expression and Mouse Behavior Recognition Technology. *Journal of Educational Computing Research*, 58(1), 63–86. <https://doi.org/10.1177/0735633119825575>

- Zheng, L., Long, M., Niu, J., & Zhong, L. (2023). An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL. *International Journal of Computer-Supported Collaborative Learning*, 18(1), 101–133. <https://doi.org/10.1007/s11412-023-09386-0>
- Zhu, Q., & Wang, M. (2020). Team-based mobile learning supported by an intelligent system: Case study of STEM students. *Interactive Learning Environments*, 28(5), 543–559. <https://doi.org/10.1080/10494820.2019.1696838>
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. <https://doi.org/10.3102/0002831207312909>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ting-Ting Wu is a distinguished professor and director of the Graduate School of Technological and Vocational Education, National Yunlin University of Science and Technology, Taiwan. Her research covers such areas as technology-enhanced language learning and applications of advanced technology to support nursing education.

Hsin-Yu Lee received his B.S. degree in the Department of Technology Application And Human Resource Development, National Taiwan Normal University, Taiwan, R.O.C., in 2019, and received the M.S. degree in the Department of Engineering Science, National Cheng-Kung University, Taiwan, R.O.C., in 2021. He is currently pursuing his Ph.D. degree in engineering science at National Cheng Kung University, Taiwan, R.O.C. His research interests include educational technology, learning analytics, computer vision and artificial intelligence.

Wei-Sheng Wang received her B.S. and M.S. degrees from the Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taiwan, R.O.C., in 2018 and 2020, respectively. He is currently pursuing his Ph.D. degree in engineering science at National Cheng Kung University, Taiwan.

Chia-Ju Lin received her B.S. and M.S. degrees from the Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taiwan, R.O.C., in 2019 and 2021, respectively. She is currently pursuing her Ph.D. degree in engineering science at National Cheng Kung University, Taiwan. Her research interests include educational technology, learning analytics, computer networks and artificial intelligence.

Yueh-Min Huang received the M.S. and Ph.D. degrees in Electrical Engineering from the University of Arizona in 1988 and 1991 respectively. He is a Chair Professor in Department of Engineering Science, National Cheng-Kung University, Taiwan. He has co-authored 3 books and has published more than 250 refereed journal research papers. His research interests include e-Learning, multimedia communications, wireless networks, and artificial intelligence. He is in the editorial board of several international journals in the area of educational technology, computer communications, and web intelligence, including 3 SSCI-indexed e-Learning ones. In e-Learning area, except of publishing more than 100 SSCI-indexed journal papers and editing 3 special issues in SSCI-indexed journals, he was also serving as the directors of Disciplines of Applied Science Education and Innovative Engineering Education in Taiwan's Ministry of Science and Technology. Dr. Huang is a senior member of the IEEE and became Fellow of British Computer Society in 2011. Dr. Huang has received many research awards, such as Taiwan's National Outstanding Research Award in 2011 and 2014, given to Taiwan's top 100 scholars. According to a paper published in BJET, he is ranked no. 3 in the world on terms of the number of educational technology papers published in the period 2012 to 2017