REGULAR PAPER

Multimodal motivation modelling and computing towards motivationally intelligent E-learning systems

Ruijie Wang[1] · Liming Chen[2] · Aladdin Ayesh[3]

Abstract

Motivation to engage in learning is essential for learning performance. Learners' motivation is traditionally assessed using self-reported data, which is time-consuming, subjective, and interruptive to their learning process. To address this issue, this paper proposes a novel framework for multimodal assessment of learners' motivation in e-learning environments with the ultimate purpose of supporting intelligent e-learning systems to facilitate dynamic, context-aware, and personalized services or interventions, thus sustaining learners' motivation for learning engagement. We investigated the performance of the machine learning classifier and the most and least accurately predicted motivational factors. We also assessed the contribution of different electroencephalogram (EEG) and eye gaze features to motivation assessment. The applicability of the framework was evaluated in an empirical study in which we combined eye tracking and EEG sensors to produce a multimodal dataset. The dataset was then processed and used to develop a machine learning classifier for motivation assessment by predicting the levels of a range of motivational factors, which represented the multiple dimensions of motivation. We also proposed a novel approach to feature selection combining data-driven and knowledge-driven methods to train the machine learning classifier for motivation assessment, which has been proved effective in our empirical study at selecting predictors from a large number of extracted features from EEG and eye tracking data. Our study has revealed valuable insights for the role played by brain activities and eye movements on predicting the levels of different motivational factors. Initial results using logistic regression classifier have achieved significant predictive power for all the motivational factors studied, with accuracy of between 68.1% and 92.8%. The present work has demonstrated the applicability of the proposed framework for multimodal motivation assessment which will inspire future research towards motivationally intelligent e-learning systems.

Keywords Multimodal motivation modelling · Eye tracking · EEG · E-learning

1 Introduction

E-learning systems have provided an opportunity to track learners' behaviour and states and thus offer opportunities to provide real-time personalized assistance. Users can learn more independently with access to a computer or a mobile device. Compared to traditional classroom learning, users using e-learning systems take more responsibilities and control of their learning progress. Highly motivated users are more likely to interact successfully with e-learning systems and thus learn more effectively. Motivation modelling and assessment are fundamental for applying personalized learning strategies in e-learning environments to address learners' motivational needs and improve learning experience. In the field of motivation assessment, the approaches can be categorised into two kinds of methods, namely subjective methods using learners' self-reported data and objective methods using physiological sensor data. The former suffers from issues such as fatigue and interruption to learning, thus leading to self-reported bias due to interferences with their learning processes. The latter offers a more accurate and efficient solution as data collection from sensors can take place in real-time and without self-reported bias, which is also a crucial step in assisting learners in specific needs

✉ Ruijie Wang
  rwang3@bournemouth.ac.uk

[1] Department of Psychology, Bournemouth University, Poole, UK

[2] School of Computing, Ulster University, Jordanstown, UK

[3] School of Computer Science and Informatics, De Montfort University, Leicester, UK

by providing them with personalized e-learning services and intervention techniques in e-learning environments. Nonetheless, such attempts are still scarce.

Given the scarcity of previous research that attempts to assess learners' motivation based on sensor data in real-time, this paper proposes a novel framework for multimodal assessment of learners' motivation in e-learning environments. The proposed framework will provide significant benefits in terms of the possibility of real-time motivation assessment based on multimodal physiological datasets, and it will facilitate the progress towards motivationally intelligent e-learning systems that entails dynamic, context-aware, and personalized services or interventions. Our generic framework for multimodal motivation assessment includes sensing technologies to capture a multimodal dataset, data processing for feature analyses, and a machine learning classifier based on data collected from multiple modalities in an e-learning environment. We then evaluate the framework in an empirical study using eye tracking and EEG as an instantiation for multimodal motivation assessment.

Motivation in e-learning context is a multi-facet concept involving both affective and cognitive aspects. In this regard, both electroencephalogram (EEG) data and eye tracking data have been widely adopted as indicators of emotional states or cognitive processes related to learners' motivation, e.g., (Derbali et al. 2012; Bixler et al. 2016; Conati 2002; Arroyo et al. 2009; Chanel et al. 2009; Conati et al. 2007). However, few research studies have utilized eye tracking or EEG for learners' motivation assessment, leaving much room for improving the prediction results.

In addition to the scarcity of previous work on motivation assessment based on multimodal data, there seems to be a failure in generalizing the approach into a reusable framework to guide future objective motivation assessment applicable to e-learning environments. Therefore, in this paper, we focus explicitly on two information channels coming from eye movements and the brain's electrical activities. It is expected that combining the sensor data from an eye tracker and an EEG headset will provide great insight in multimodal assessment of learners' motivation.

In this paper, we develop a novel approach to feature analyses, where both statistical analysis and domain knowledge are used to select the most salient features for assessment of different motivational factors. Specifically, the proposed motivation assessment framework consists of the following components: (i) a motivation model containing a range of motivational factors, developed from domain knowledge and empirical data, where the motivational factors represent different dimensions of motivation, among which the level of each factor to be predicted for motivation assessment; (ii) sensing technologies to capture multimodal dataset; (iii) data processing and feature analyses that extract and select features from the raw sensing data for motivation assessment;

and (iv) a machine learning classifier, namely logistic regression, that predicts the level of each motivational factor.

Though a large number of EEG and eye gaze features can be extracted from the raw sensor data, there is a distinct lack of empirical investigation of the prediction power of EEG and eye gaze features in motivation assessment. Therefore, we conduct an experiment for a specific instantiation of the framework. In the experiment, we collect EEG and eye tracking data to predict the level of each motivational factor in the motivation model in Wang et al. (2020). We also describe the process of motivation assessment in the proposed framework, including capturing raw data, analyses for extracting and selecting features, and using the selected features to train a logistic regression classifier. Hence, the main research contributions recorded in this paper are summarized as follows:

- We develop a generic framework that demonstrates how multimodal sensing and machine learning methods can be combined for motivation assessment in e-learning environments. The framework provides enabling technologies for a motivation-aware intelligent e-learning system that can provide personalized services or interventions to address learners' different motivational needs in real-time. This reusable generic framework addresses the scarcity of motivation assessment using sensor data in e-learning environment and will provide great insights to guide future research and practices in this field.
- We introduce a novel approach to feature selection combining data-driven (i.e., statistical tests) and knowledge-driven methods (i.e., knowledge from the motivation model), to train the machine learning classifier for motivation assessment, providing an efficient way of selecting predictors from a large number of extracted features from EEG and eye tracking data.
- Our empirical study provides evidence on the most and least accurately predicted motivational factors, which will inspire the direction of future research and practice in terms of which motivational factors are predictable based on EEG and eye tracking data.
- Our empirical study indicates the role of different EEG and eye tracking features in motivation assessment and identifies those features that are particularly informative, which will inform future research on the rationale behind it regarding how physiological signals are related to learners' motivational states.

The remainder of the paper is organized as follows: Sect. 2 reviews related work, and Sect. 3 describes the motivation model developed using qualitative and quantitative approaches. Section 4 describes the proposed framework for multimodal motivation assessment towards motivationally intelligent e-learning systems. Section 5 presents the

experiment process and results based on EEG and eye tracking data. Finally, discussion and conclusion are provided in Sect. 6.

2 Related work

2.1 Motivation modelling in e-learning context

From psychological perspective, motivation is a multifacet concept, considered by many to have multiple factors. Motivation modelling in the present e-learning context is the process of establishing a conceptual understanding of the factors that determines a user's motivation to engage in an e-learning system. For example, Ryan and Deci (2000) has distinguished extrinsic motivation as "doing something because it leads to a separable outcome" from intrinsic motivation as "doing an activity for the inherent satisfaction of the activity itself". Intrinsic motivation often accompanies increased attention and intrinsic goals. Researchers have also looked into factors contributing to intrinsic motivation, which are perceived challenge, feedback, perceived choice, perceived interest, curiosity and perceived competence consisting of self-efficacy, anxiety or emotion (Shroff et al. 2009; Sun and Sun 2008). Deci and Ryan's self-determination theory supposes that people's motivation is self-determined by the degree to which their innate psychological needs are satisfied, i.e., autonomy, competence and relatedness (Deci et al. 2000).

In e-learning context, learners can learn more independently with access to a computer or a mobile device. Unlike traditional classroom learning, learners in e-learning systems take much more responsibility and control of their own learning progress. Highly motivated users of e-learning systems are more likely to interact effectively with the systems and thus learn more effectively. Motivation in e-learning context is fundamental for effective engagement and learning, and it is a more acute issue for people with learning difficulties such as dyslexia due to reading, writing or other difficulties they usually experience in learning that can potentially lead to frustration and learned helplessness. Therefore, our present study employs people with dyslexia as target audience for constructing the motivation model and evaluating the framework of multimodal motivation assessment.

However, most e-learning systems still focus on improving users' knowledge and skills, and the few attempts to provide motivational strategies lack a fundamental basis and guidance from an empirically tested motivation model in e-learning context. From the perspective of technology use, people's motivation changes are reflected in the degree of their acceptance of the technology, where perceived ease of use and perceived usefulness are also important factors.

Thus, users' motivation can be contextualized as continued use intention in the context of interacting with e-learning systems. Technology Acceptance Model (TAM) is the most widely adopted model to explain users' acceptance of technology by two drivers, perceived ease of use and perceived usefulness (Davis and Davis 1989). However, this model has been criticized for its overemphasis on extrinsic motivation, so there have been many attempts to extend the model with intrinsic motivation or other factors, as stated by Chang et al. (2013) who has extended TAM with perceived convenience and playfulness that influence continued use intention for a mobile learning system. Tawafak et al. (2018) have integrated academic performance, student satisfaction, support assessment and effectiveness with TAM to explain the continuance of intention to use the universities' learning management systems. Herrador-Alcaide et al. (2019) have conducted research that targeted at students of financial accounting and revealed that students' perceptions of both the e-learning environment and their own skills have effects on their overall feelings of satisfaction. Hanif et al. (2018) have also extended TAM where subjective norm, perception of external control, system accessibility, enjoyment, and result demonstrability have a positive influence on undergraduate students' use of e-learning systems.

In summary, motivation modelling provides a way to define the users' motivational needs as well as the relevant factors related to system factors perceived by users that may influence their motivation. In e-learning context, only when an e-learning system has information of users' individual motivational needs, the motivational strategies applied can then support different users' needs in a more personalized and thus more effective manner. Motivation has been modelled from both the psychological perspective and technology acceptance perspectives. However, research grounded in motivation theories for learners' continued intention to engage in e-learning environments has been scarce to date. In this paper we investigate the assessment of multiple motivational factors that are grounded in our motivation model developed based on motivational theories and empirical research (Wang et al. 2020) to provide a comprehensive view of the applicability of the proposed method for motivation assessment.

2.2 Multimodal motivation assessment towards motivationally intelligent e-learning systems

E-learning systems can support real-time monitoring of users' behavioural and physiological responses that indicate learning desires, effects and various mental processes or states, thus offering opportunities for enhanced learning through dynamic provision of personalized learning assistance. In mobile or web-based e-learning systems, it

is possible for users' motivational states to be detected in real-time and thus the personalized motivational strategies to be applied during the interaction process between users and systems.

Motivation consists of various factors from intrinsic motivation and extrinsic motivation. These factors represent learners' various motivational needs; thus, each factor should be assessed in order to design and implement personalized feedback to address the corresponding need. Once the motivational needs of a user are detected in an e-learning system, personalized reactions using motivational strategies can be output to the user to address the corresponding motivational needs dynamically to sustain or improve motivation and experience.

Important indicators of motivational factors include time spent on a completing a learning task, quiz scores, and various sensor data. The identification of the physiological or behavioural indicators of motivation is still at its initial stage, though researchers have stated that through motivation-diagnostic input data, appropriate tactical and strategic pedagogic moves are applicable toward motivationally intelligent systems (du Boulay et al. 2010). Increasingly more attention has been paid to the use of physiological sensor data in detection of mental states. Some researchers have used physiological sensors to assess learners' motivation and linked sensor data to some factors relevant to motivation such as attention and confidence (e.g., (Derbali and Frasson 2010; Rebolledo-mendez et al. 2010)), but most of them focused on the assessment of affective states or the specific aspect relevant to motivation such as attention. For instance, Conati (2002) has used biometric sensors (heart rate, skin conductance, electromyogram) and facial expression analysis to develop a probabilistic model of detecting students' affective states towards emotionally intelligent educational games. Arroyo and colleagues (2009) have used four different sensors (camera, mouse, chair, and wrist) in a multimedia adaptive tutoring system to recognize students' affective states and embed emotional support. Among various sensing technologies, EEG and eye tracking have been widely used and shown much potential to inform mental states.

EEG is an electrophysiological technique to record the electrical activity generated by the human brain via electrodes placed on the scalp, and it can reflect a variety of mental processes such as cognitive workload or active concentration. Derbali and Frasson (2012) have found theta wave in the frontal brain region and high-beta wave in the left central region played a significant role in predicting motivation of players using EEG in a serious game. In Jenke et al. (2013), statistical t-test and a univariate feature selection method using Cohen's effect size f^2 from analysis of variance were implemented for electrode and feature selection. Electrodes and features found by these approaches resulted in a small variance in classification accuracies across subjects. Knott

et al. (2001) used three-way multivariate analysis of variance (MANOVA) and t-test for absolute power, relative power and hemispheric asymmetry measures. He found that absolute and relative power in the beta frequency band, but not in the delta, theta or alpha frequency bands, differentiated the depressed group and control/normal group. Chanel et al. (2009) asked participants to recall an episode in their life that corresponded to positive emotions and one that corresponded to negative emotions. A classification accuracy of 63% was reported using the short-time Fourier transform for feature extraction and a linear Support Vector Machine (SVM) for classification.

As a way of collecting data about human eye movements, eye tracking has been widely used for gaze analysis for assessment purpose or as gaze input for interaction purpose. Eye tracking can also provide insights into real-time motivation assessment from prior research about using eye tracking to detect or assess various mental processes or behaviour. For example, Conati and Merten (2007) have used eye tracking data to assess user meta-cognitive behaviour during interaction with an environment for exploration-based learning. Bixler and D' Mello (2016) investigated the use of eye gaze and contextual cues to automatically detect mind wandering during reading with a computer interface. They have applied a correlation-based feature selection process to remove features that convey the same information. Twenty different supervised machine learning techniques including logistic regression and SVM were applied in their study, and Global gaze features (gaze patterns independent of content, such as fixation durations) were found more effective for mind wandering detection than content-specific local gaze features. In addition, eye tracking techniques have been used to detect driver fatigue (Horng et al. 2004). Moreover, cognitive load of a learner could be determined using eye tracking and pupil measuring (Beatty 1982). Moreover, some studies have been conducted on investigating eye responses on emotional stimuli (Partala and Surakka 2003; Wang et al. 2019). In a previous study (Wang et al. 2019), we have attempted to evaluate the motivational strategies and developed logistic regression models based on eye tracking data to assess learners' motivation in an e-learning environment and identified the most significant predictors such as fixation number and pupil diameter.

There are research using EEG or eye tracking data for assessment of mental states such as emotion, cognitive load which are all related to motivation (Crocker et al. 2013). However, to the best of our knowledge, no studies have employed a multimodal dataset combining EEG and eye tracking to infer the levels of motivational factors for users of e-learning systems, and the predictive power of combining EEG and eye tracking data for motivation assessment is yet to be investigated using empirical studies. This study fills this gap in the field of combining eye tacking and EEG

data for multimodal motivation assessment in e-learning environments, which lays the foundation for motivationally intelligent e-learning systems that can provide personalized services or assistance to improve motivation and learning for users in need.

3 Qualitative and quantitative motivation modelling

Using both qualitative and quantitative approaches, we have developed the motivation model for people with dyslexia in our previous study (Wang et al. 2020) which will be briefly described here for clarity. In the qualitative modelling phase, the model was firstly constructed based on domain knowledge including both intrinsic and extrinsic factors of motivation with adjustment to e-learning context. We then conducted an empirical, qualitative study with dyslexic students. The study provided participants with a real-world learning scenario on a mobile learning application followed by individual interviews to elicit the key motivational factors. The qualitative motivation model was developed by combining thematic analysis with previous research in literature.

In the quantitative modelling phase, we constructed a questionnaire to assess self-reported motivation based on the motivation model. It used multi-item 5-point Likert-style questions (where 1 is strongly disagree and 5 is strongly agree) to rate the statements pertaining to the factors in the motivation model. The questionnaire data resulted in quantification of the interrelationships through Structural Equation Modelling (SEM). The motivation model developed using both the qualitative and quantitative approaches is shown in Fig. 1, where the quantified relationships between the factors resulted from covariance-based SEM, showing how different factors function together with direct and/or indirect effects on continued intention to use e-learning systems. The factors in the motivation model include 4 extrinsic motivational factors (i.e., Visual Attractiveness, Perceived Control, Perceived Ease of Use, Perceived Usefulness), 4 intrinsic motivational factors (i.e., Feedback, Self-Efficacy, Perceived Privacy, Attitudes Toward School), 4 motivational factors acting as mediators (i.e., Confirmed Fit, Reading Experience, Utilization, Learning Experience), and the motivational consequence (i.e., Continued Use Intention). They are all referred to as motivational factors in the present paper. Details about development of the motivation model using SEM have been demonstrated in the previously published paper (Wang et al. 2020).

Although the motivation model results from both domain knowledge and empirical data from qualitative and quantitative studies, it is referred to as knowledge-based motivation model in the present paper in that the factors in the
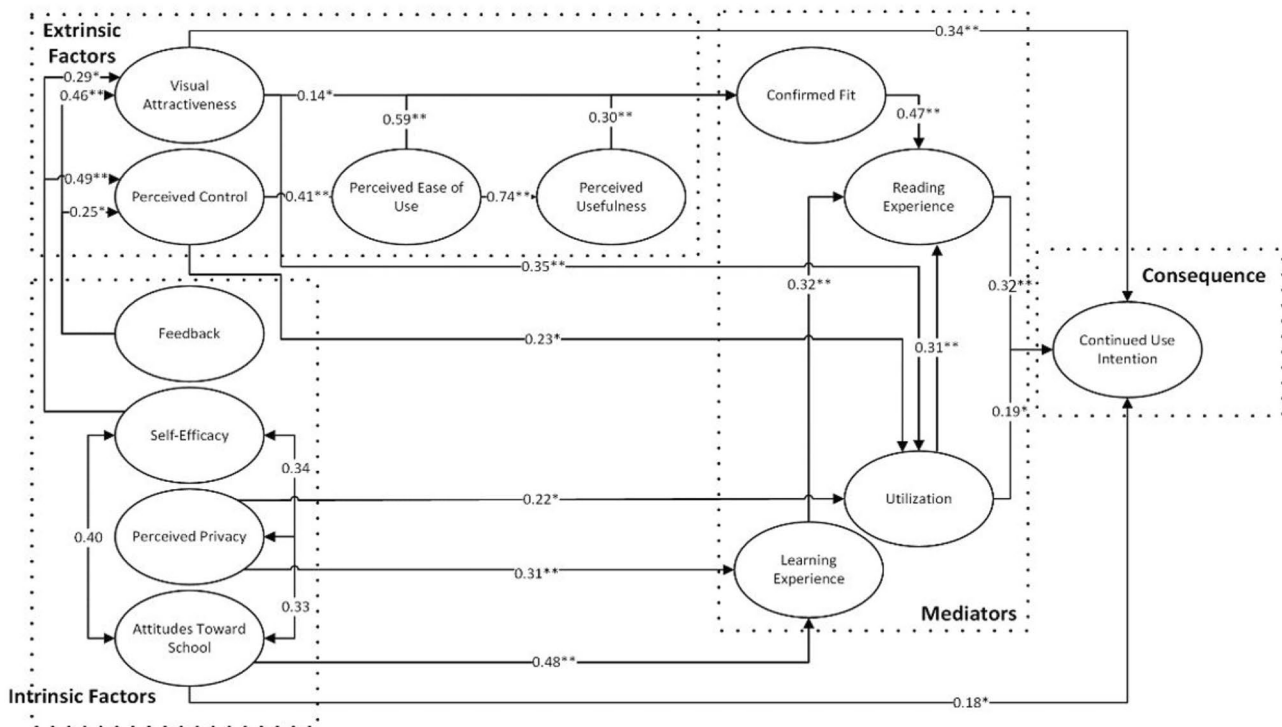

Fig. 1 The knowledge-based motivation model showing standardized regression weights (all paths are significant, *p-value < 0.05; *p-value < 0.01) (Wang et al. 2020)

motivation model purely come from multidisciplinary domain knowledge, compared to features selected from the multimodal sensor data. The knowledge-based motivation model contains factors from intrinsic motivation (Self-efficacy, Attitudes Toward School, Perceived Privacy and Feedback), extrinsic motivation (Visual Attractiveness, Perceived Control, Perceived Ease of Use and Perceived Usefulness) and mediators (Confirmed Fit, Utilization, Learning Experience and Reading Experience).

To address different individuals' motivational needs in e-learning environments, the motivation assessment in our research includes computation of different dimensions of motivation, represented by the factors in the knowledge-based motivation model. The high-level structure of the motivation model that integrates multimodal sensor data is presented in Fig. 2 to provide a holistic picture, which implies that multimodal sensing such as eye tracking and EEG can be employed for automatic assessment of the motivational factors.

4 The framework for multimodal motivation assessment

Figure 3 depicts the proposed framework for motivation assessment based on multimodal sensor data. The main idea is that various physiological signals captured by sensing technologies can be used as indicators of learners' motivational states. The execution process using the framework is elaborated as follows.

The framework consists of four components, namely multimodal sensing, feature extraction and selection, machine learning classifier, and motivational factors. Central to the
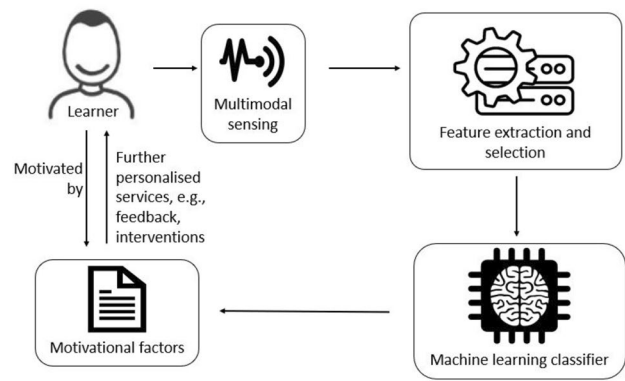

Fig. 3 The motivation assessment framework based on multimodal sensing in e-learning environments

process are feature extraction and selection and machine learning classifier which takes multimodal datasets as inputs and produce predicted results as outputs pertaining to the levels of motivational factors. Multimodal data can be collected from a variety of sensing technologies. In the present paper, we use EEG and eye tracking as a specific instantiation to describe the process in the proposed framework.

4.1 Multimodal sensing, feature extraction and selection

First, raw sensor data are collected from learners using a Tobii eye tracker and an Emotiv 14-channel EEG device while they are learning in an e-learning environment. Second, a series of features are extracted from the raw data for motivation assessment. The extracted EEG features are 5 power bands * (1 mean + 2 extreme value + 4 brain lobe
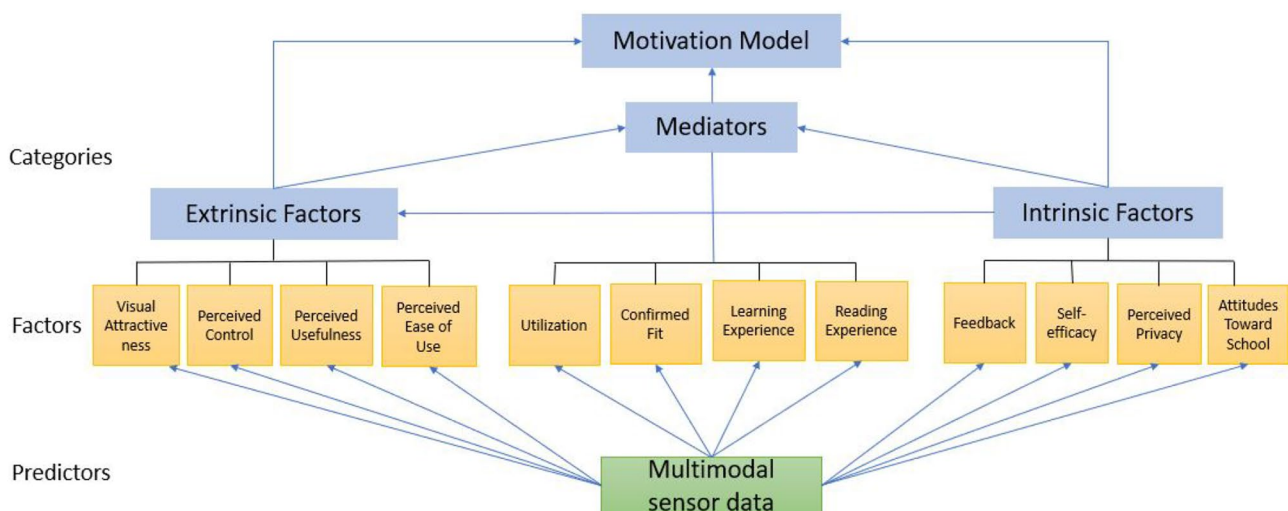

Fig. 2 The high-level structure of the motivation model with incorporation of multimodal sensor data

mean + 2 hemisphere asymmetry). In detail, they are categorized into four:

- The mean power (dB) of theta, alpha, low beta, high beta and gamma bands among all channels.
- The extreme values (both maximum and minimum) of each of the five bands.
- The mean power of each of the five bands for each of the four regions, i.e., occipital, parietal, frontal and temporal lobe.
- The hemisphere asymmetry of each of the five bands, including both the intra-hemispheric power asymmetry and inter-hemispheric power asymmetry. According to the "neurometrics" formulas from John et al. (1988) and Prichep and John (1992), inter-hemispheric power asymmetry for each band is computed with the formula [(R-L)/(R + L)], where R and L refers to the right hemisphere and left one, respectively, and the intra-hemispheric asymmetry is computed with the formula [(A-P)/(A + P)], where A and P refers to the anterior (i.e., frontal) region and posterior (i.e., back) one, respectively.

Amongst the eye gaze features extracted, 9 of them from fixation domain, 3 of them from saccade domain and 5 others are all specified below. Some gaze features are extracted from data collected for specific areas of screens called Areas of Interest (AOIs). That allows the screen areas related to learning contents to be separated from the blank areas. Z score standardization is performed for pupil diameter, and the unit of all the time measures is unified as seconds, and all length measures are computed as pixels. Specifically, they are categorized into three:

- 9 features from fixation domain: fixation number in AOI, fixation duration in AOI, fixation number in all screen areas (during a lesson overall and last 10 s of the lesson), fixation connection length, fixation spatial density, path velocity, regressions (i.e., regressive eye movements, during a lesson overall and last 10 s of the lesson).
- 3 features from saccade domain: average saccade velocity, saccade duration, average saccade length.
- 5 others: fixation saccade ratio, pupil diameter (mean and maximum), samples out of monitor, data loss (due to blinks and out of monitor).

Following the feature extraction, feature selection is performed based on: (i) statistical analysis on sensor data and the motivational factors; and (ii) the relationships between factors in the knowledge-based motivation model. First, statistical analysis is conducted to generate salient EEG features in a data-driven manner from Spearman correlation test and ANOVA test (alternatively Kruskal Wallis Test for variables with a non-normal data distribution).

Significant EEG features and eye gaze features which can differentiate the high level from low level of the motivational factors are computed from all the extracted features with Spearman correlation and ANOVA tests (or Kruskal Wallis Test for variables with non-normal distribution). Second, a knowledge-driven approach to feature selection is employed according to the motivation model described in Sect. 3 to minimize the possibility of missing the features that may improve the classification accuracies. Specifically, the relations between the motivational factors in the motivation model are also considered, and the features selected for the direct independent factors of a motivational factor are also be adopted for the assessment of the levels of the motivational factor. For example, as shown in Fig. 1, Perceived Ease of Use (Factor A) is the direct independent factor of Perceived Usefulness (Factor B), so the features selected for Factor A will also be selected as inputs for assessing the level of Factor B. Particularly, the significance (i.e., p-value) of the effect of Factor A on Factor B is required to be less than 0.001; if the p-value is between 0.001 and 0.05, only the features for assessment of Factor A with the significance (i.e., of the features' correlation with Factor A or difference between the two levels of Factor A) of less than 0.001 are chosen for assessment of Factor B.

4.2 Machine learning classifier and prediction for motivational factors

Based on a series of data analyses described in Sect. 4.1, an integrated set of features are selected to develop a machine learning classifier for classification of the level of each motivational factor. We use logistic regression in the present study as an exemplar to describe the process of generating a classification model for inferring the high/low level of each motivational factor based on the multimodal dataset, as it has been proved effective at assessing learners' motivation based on eye tracking data in our previous study (Wang et al. 2019).

Logistic regression performs the classification by computing a probability of a motivational factor M_1 being at high level $P \in [0, 1]$, using:

$$P(M_1) = \frac{1}{1 + e^{-z}}, where Z = \sum_{i=0}^{n} \beta_i X_i and X_0 = 1. \qquad (1)$$

Coefficients β_i measures the effect of a predictor X_i being significant on the probability of high level of the motivational factor. Thus, for positive β_i the greater the value of predictor X_i, the greater the increase in the probability of motivational factor being high level and vice versa. β_0 is the constant which is the log of the odds when

all X_i equal 0. Then π is deduced from P via a threshold $\gamma \in [0, 1]$ for the assumed uncertainty of the solution:

$$\pi(M_1) = \begin{cases} 1, P(M_1) > \gamma \\ 0, otherwise \end{cases} \quad (2)$$

where π refers to the level of a motivational factor, and $\pi(M_1)$ being 1 or 0 represents that M_1 is classified into the high or low level, respectively.

Finally, the predictive results for motivation assessment can be used further for providing personalized services in the e-learning environment to support learners' motivational needs.

5 Experiment and evaluation

In the present study, we conducted an experiment to illustrate and evaluate the proposed framework for multimodal motivation assessment. The experiment aims at investigating possibility of using the features extracted from EEG and eye tracking data to infer the levels of the motivational factors in an e-learning environment. The participants and procedure of the experiment are firstly described in this section, and then the proposed framework for motivation assessment is evaluated based on the multimodal sensor data collected from the eye tracker and EEG device. The experiment was approved by the Faculty Research Ethics Filter Committee at De Montfort University.

5.1 Experiment method

5.1.1 Participants and learning materials

Twenty-five participants (16 females and 9 males) were recruited for the experiment. All of them came from Leicestershire, most of which were university students with one from a middle school; the mean age of them was 25.5 (SD = 8.4). Thirteen of them have been diagnosed as dyslexic, and the others have self-reported learning difficulties without formal diagnosis. We recruited participants with learning difficulties as our knowledge-based motivation model was developed for people with dyslexia, who usually suffer from lack of motivation in learning, and using this target group in experiments to evaluate our approach to motivation assessment can also help avoid ceiling effect.

The learning materials consist of three lessons with each taking about 5–10 min to complete. Each lesson contains both text and picture as well as quizzes at the end. Lessons were designed to teach transferable skills about learning and reading such as reading strategies (Lesson 1 and Lesson 2) and time management skills to avoid procrastination (Lesson 3). As our participants were people with learning difficulties including dyslexia, meaning most of which having reading difficulties, we supposed that Lesson 1 and Lesson 2 were more related to their specific learning needs and thus more likely to attract their attention at the very beginning. Our participants might have a relatively low level of motivation due to suffering from learning difficulties compared to those without, so we presented Lesson 1 and Lesson 2 prior to Lesson 3 which is about general learning skills about time management and organization, in order to avoid floor effect. The sequence of Lesson 1 and Lesson 2 allows smooth transition between basic reading strategies to facilitate understanding and effective reading (Lesson 1) and quick reading tips to facilitate efficient reading (Lesson 2). All participants did not learn about the same knowledge before the experiment. Teaching knowledge about transferable skills was to minimize the effect of difficulty levels of the learning materials compared to the like of scientific lessons.

5.1.2 Simplified motivation questionnaire

We employed a simplified motivation questionnaire to collect self-reported data on the motivational factors, which provided labels to develop the logistic regression classifier to predict the levels of motivational factors based on the sensor data from eye tracking and EEG. Each participant was asked to fill in a simplified multi-item motivation questionnaire after each lesson (see Fig. 4 for a screenshot). The questionnaire was constructed based on the conceptual motivation model built from a qualitative approach in the initial stage (Wang et al. 2017). The original questionnaire (Wang et al. 2020) consists of 61 statements in total with about 3 to 5 statements for each motivational factor, while the simplified version has totally 33 statements. The questionnaire

Fig. 4 The multi-item motivation questionnaire in Google Form

was simplified in order to reduce the time learners spent on the questionnaire between the lessons and thus the effect of interruption during their learning process.

To examine whether the simplified motivation questionnaire is reliable for measuring each motivational factor or not, a Cronbach's Alpha was employed on the questionnaire data. The Cronbach's Alpha for Continued Use Intention, Perceived Usefulness, Perceived Ease of Use, Confirmed Fit, Feedback, Visual Attractiveness, Learning Experience, and Reading Experience was 0.91, 0.78, 0.81, 0.91, 0.70, 0.70, 0.76, and 0.77, respectively. However, Perceived Control, Utilization and Perceived Privacy were removed, as the reliability of the corresponding statements did not pass the threshold. The short questionnaire used to measure intrinsic motivation at the beginning of the experiment was also examined with Cronbach's Alpha, and the reliability of Self-efficacy and Attitudes Toward School was 0.74 and 0.60, respectively. The results showed that the simplified motivation questionnaire used in the present study was overall reliable, with the removal of three factors.

5.1.3 Experiment procedure

The experiment was conducted individually for each participant in the same laboratory environment and setting (see Sect. 5.2.1 for setting). Before the experiment, the participants were provided with an information sheet and a consent form to be signed, which clearly explained the study objectives, data collection process and privacy protection, the rights of participants, etc. After the experiment, a voucher worth 10 British pounds from Amazon/Tesco/John Lewis was given to each participant as compensation.

The experiment procedure is summarized in a flow diagram in Fig. 5. At the beginning of the experiment, participants were asked to complete a short questionnaire pertaining to intrinsic motivation including attitudes toward school and self-efficacy. The two intrinsic motivational factors were measured in the pretest instead of during the learning process, because they were formed by learners' long-term learning and life experience, not likely to change due to different circumstances in a short time period. After that, the eye tracker and EEG headset were calibrated, and then each participant was asked to complete three learning tasks (i.e., three lessons) with a quiz and the motivation questionnaire after each lesson.

5.2 Data collection and analysis

5.2.1 Setting for multimodal sensing

The Open Gaze And Mouse Analyzer (OGAMA) 5.0, an open-source software, was used for eye tracking data recording and analysis (see Fig. 6a for an example screenshot of
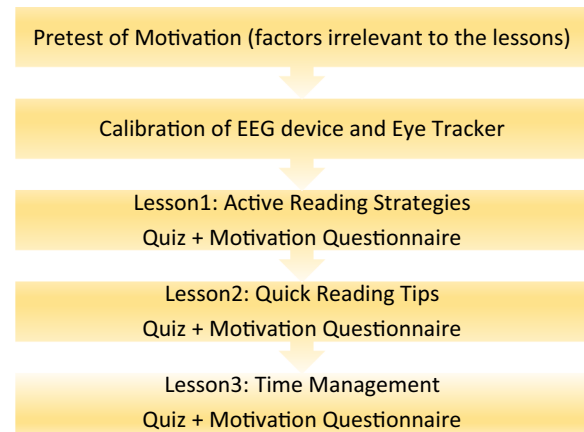

Fig. 5 The flow diagram of the experiment procedure

the OGAMA learning environment). The three learning lessons were adapted to the OGAMA environment. Tobii X120 tracker with a sampling rate of 60 Hz was employed to collect eye movements. The eye tracker is a standalone device that did not restrain participants from head movements. A wearable EEG headset, provided by Emotiv Inc. called Emotiv EPOC + with the bandwidth of 0.16-43 Hz, was employed to collect brainwave data. The EEG device has 14 electrodes with metal contacts and felt sensors which need saline solution for adequate contact quality. The electrodes are in line with the international 10–20 system, with placements of the electrodes (i.e., AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4) shown in Fig. 6b, where the corresponding brain regions are also annotated (Hou et al. 2015). It is wireless, mobile and able to transmit data via Bluetooth, thus with little discomfort compared to traditional EEG with wires and gel solution.

Meanwhile, EmotivPro 1.8 was used in conjunction with Emotiv EPOC + to observe and record EEG data. The headset was configured at 128 Hz as the EEG sample rate. The software also enables observation of a real-time Fast Fourier Transform (FFT) plot of raw data and recording of the power in each of the five frequency bands, i.e., theta (θ: 4–8 Hz), alpha (α: 8–12 Hz), low beta (low β: 12–16 Hz), high beta (high β: 16–25 Hz) and gamma (Υ: 25–45 Hz) (see Fig. 7 for an example screenshot of a raw EEG interface). The experiment setup for multimodal sensing from eye tracking and EEG devices is displayed in Fig. 8.

5.2.2 Feature extraction and selection

The data was collected from the participants including a total of seventy-five trials (i.e., lessons) from the experiment. We split the data in a pseudo-randomized manner by
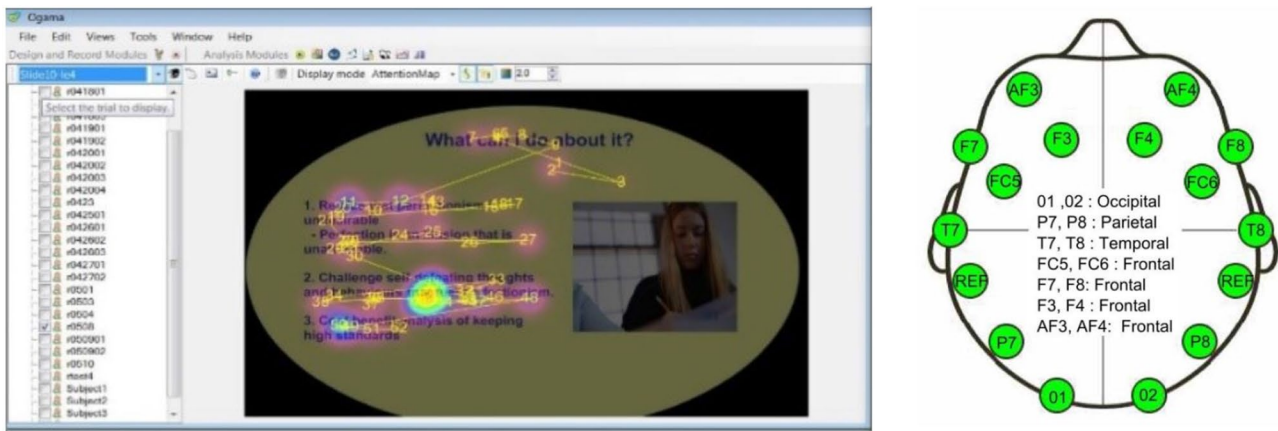
Fig. 6 On the left **a**: the screenshot of an e-learning interface with the attention map in OGAMA; on the right **b** the international 10–20 system with electrode positions corresponding to brain lobes (Hou et al. 2015)

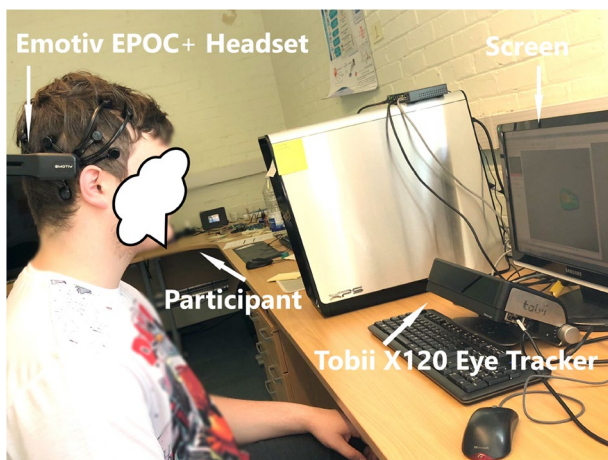**Fig. 7** The raw EEG interface in EmotivPro

Fig. 8 System setup for multimodal sensing

using the dataset from the first thirty-nine trials (i.e., the first thirteen participants in the study) to train the model and

testing it on the entire dataset. After removing the outliers of EEG data and eye tracking data according to the descriptive statistics, we extracted the features from the multimodal sensor data, following the process described in Sect. 4.1. In addition, time duration and quiz score of a lesson for each participant were recorded as two features.

Afterwards the features were selected generating classification models, following the process described in Sect. 4.1. Specifically, the extracted features from EEG and eye tracking data were analyzed using statistical tests for their relations with the motivational factors. The results from Spearman correlation and ANOVA tests (or Kruskal Wallis Test for variables with non-normal distribution), are reported in the Appendices. The significant features involve all categories of data sources where the features were extracted: EEG, Eye (i.e., eye tracking), and OtherBehaviour (i.e., lesson duration and quiz performance), indicating that the features extracted from the employed sensor data and other behaviour data during learners' interaction process with the e-learning environment are promising at facilitating the assessment of

the motivational factors. These significant features to be used as inputs of a classification algorithm were all selected in a data-driven manner from the statistical tests.

We then applied the knowledge-drive process of feature selection described in Sect. 4.1, and the results of the selected features are reported in Table 1. Using the three steps of statistical analyses and knowledge-driven method, we selected 4–26 features for each motivational factor; in other words, we removed 38–60 features from the total 64 features for each motivational factor, which is much more efficient comparing to the baseline method of removing these features one by one.

5.2.3 Prediction results for motivational factors

From the abovementioned correlation analyses from both data-driven and knowledge-driven approaches, we got a number of features selected as the input data. Afterwards, we performed the motivation assessment, i.e., inferring the high/low level classification for each of the motivational factors using logistic regression as the machine learning classifier, following the process described in Sect. 4.2. The 0.5 cut-off point was set as default decision threshold to infer the level (i.e., π in (2)) of each motivational factor. To find an optimal cut-off point for each motivational factor, a ROC curve was drawn, using "Sensitivity" as Y-axis and "1-Specificity" as X-axis derived from all possible cut-off values. Sensitivity is the rate of true positive predictions and specificity is the rate of true negativity predictions, so we aimed to find the point which was the closest to the point (0, 1) to maximize both sensitivity and specificity. This process is demonstrated below for the motivational factor "Visual Attractiveness" as an example.

Combined with the chart showing the results in Table 2, the cut-off point, 0.348, with the point (0.89, 0.17) in ROC curve in Fig. 9 should be adopted as the best one for Self-Efficacy classification. In the same way, the cut-off points for all the motivational factors were identified. Our model for inferring the level of each motivational factor used Backward method, to remain only the subset of variables (features of EEG or eye gaze) from the selected features that were more related to the response variable (the motivational factor) to make the model least prone to error according to the statistic of likelihood ratio (LR). The model achieved statistically significant prediction power, reported in Table 3, indicating that adding the features from sensor data including EEG and eye tracking have significantly improved the prediction
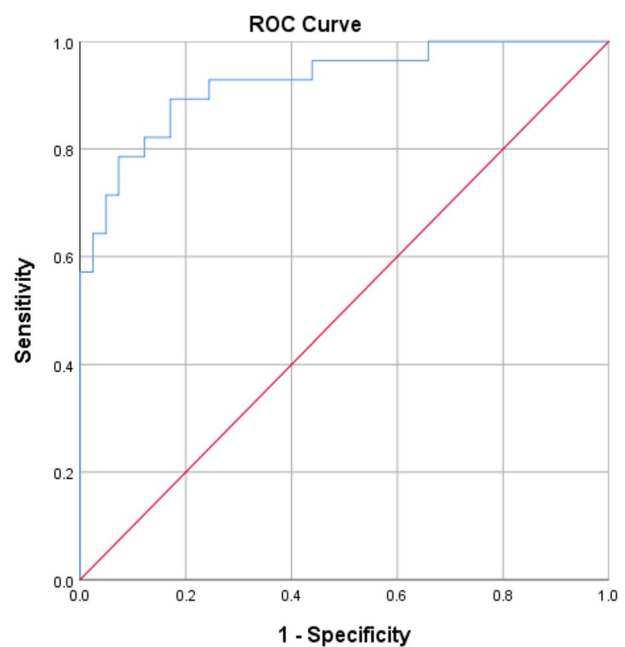

Fig. 9 ROC curve for visual attractiveness

Table 1 The features from direct independent factors

Dependent factors	Independent factors with direct effects
Perceived usefulness	Perceived ease of use
Confirmed fit	Visual attractiveness*, perceived usefulness, perceived ease of use
Learning experience	Attitudes toward school
Reading experience	Confirmed fit, learning experience
Visual attractiveness	Feedback, self-efficacy*
Perceived ease of use	Self-efficacy
Continued use intention	Attitudes toward school*, reading experience, visual attractiveness

*For these independent factors, only features with p < 0.001 selected from the statistical tests are adopted for their dependent factors, because the p-values for their causal relations are between 0.001 and 0.05; for all the other factors in the right column have direct causal effects on the corresponding factors in the left with p-value of less than 0.001

Table 2 Part of coordinates of the ROC curve for visual attractiveness

Positive if greater than or equal to*	Sensitivity	1 - Specificity
0.2879672	0.929	0.244
0.3160239	0.893	0.244
0.3340052	0.893	0.220
0.3408276	0.893	0.195
0.3475657	0.893	0.171
0.3573349	0.857	0.171
0.3669414	0.821	0.171
0.4225623	0.821	0.146
0.4800537	0.821	0.122
0.4915688	0.786	0.122

Table 3 Omnibus tests of model coefficients using logistic regression

Motivational factors	Chi-square	df	Sig.	Nagel-kerke's R Square
Perceived Usefulness	18.072	4	0.001	0.314
Confirmed Fit	21.521	4	0.000	0.359
Feedback	35.448	6	0.000	0.550
Attitudes Toward School	27.170	4	0.000	0.441
Learning Experience	34.087	4	0.000	0.520
Reading Experience	45.880	9	0.000	0.648
Continued Use Intention	28.242	5	0.000	0.452
Visual Attractiveness	45.386	9	0.000	0.651
Perceived Ease of Use	17.872	3	0.000	0.305
Self-efficacy	58.529	5	0.000	0.775

ability to distinguish between high level and low level for all the studied motivational factors. Nagelkerke's R square ranged from 30.5 to 77.5%, indicating a moderately high relation between the predictors and the motivational factors.

The classification accuracies obtained with both the default and optimal cut-off value are shown in Table 4 for all the motivational factors studied, the features having significant contributions with significance level of 0.05 to the prediction models are also shown in the table. The accuracy of prediction based on EEG and eye tracking data has achieved

68.1-92.8% for the motivational factors with the optimal cut-off values. The explanation ability of predictors as well as the prediction accuracy is the weakest for Perceived Ease of Use, indicating that it is worth introducing more predictors other than the EEG and eye gaze features used in the present study for predicting the factor.

6 Discussion and conclusion

Our present study has provided promising results for combining EEG and eye gaze features for motivation assessment, showing advantages compared to previous research involving solely EEG or gaze features. Compared with our previous study that only adopted eye gaze features for the classification task (Wang et al. 2019), introducing EEG features and using our proposed method step by step has been proved effective in the present study at developing regression models with better quality in terms of the significance of the prediction power and the accuracy of classification. Particularly for the factor Self-efficacy and Continued Use Intention, using only eye gaze features resulted in a model with insignificant and near-to-the-threshold significant prediction ability, respectively. In contrast, the present method has significantly improved the prediction ability for both factors. Furthermore, using a similar sample size (thirty-three participants), a previous study (Derbali et al. 2012) explored

Table 4 Classification accuracies for the motivational factors with cut-off points and significant features

Motivational factors	Cut-off points	Classification Accuracy (%)	Significant Features (with significance level of 0.05)
Perceived Usefulness	0.5	72.5	QuizPerformance; Gamma-occipital; HighBeta-temporal; HighBeta-IntraAsymmetry
	0.600	76.8	
Confirmed Fit	0.5	68.1	FixationNumber-overall; Gamma-mean; Gamma-occipital
	0.640	73.9	
Feedback	0.5	72.5	Gamma-max; Gamma-temporal
	0.371	79.7	
Attitudes Toward School	0.5	76.8	FixationNumber-overall; Alpha-InterAsymmetry
	0.496	76.8	
Learning Experience	0.5	81.2	PupilDiameter-average; FixationSpatialDensity; HighBeta-frontal; Gamma-occipital
	0.530	81.2	
Reading Experience	0.5	88.4	LessonDuration; QuizPerformance; PupilDiameter-max; SaccadeLength-average; HighBeta-max; HighBeta-min; HighBeta-occipital
	0.536	88.4	
Continued Use Intention	0.5	75.4	FixationNumber-10s; Theta-mean; Alpha-min; Alpha-max; HighBeta-frontal
	0.438	82.6	
Visual Attractiveness	0.5	84.1	FixationSaccadeRatio; FixationNumber-overall; Theta-max; HighBeta-frontal; HighBeta-temporal; Gamma-temporal; LowBeta-occipital
	0.348	85.5	
Perceived Ease of Use	0.5	62.3	FixationSpatialDensity; Gamma-temporal; Alpha-IntraAsymmetry
	0.416	68.1	
Self-efficacy	0.5	92.8	Gamma-frontal; Gamma-occipital; Alpha-frontal; HighBeta-frontal
	0.481	92.8	

using physiological data including EEG, heart rate, and skin conductance to develop logistic regression models for predicting learners' motivation, which achieved prediction success between 65.5% and 79.3% and found significant feature only from EEG data among the three types of physiological data. In contrast, in the present study, we have achieved prediction success between 68.1% and 92.8% with a large number of significant features from both eye tracking and EEG shown in Table 4, indicating the promising advantage of combing eye tracking and EEG for motivation assessment.

The prediction accuracy is highest (over 85%) for Self-Efficacy, Reading Experience and Visual Attractiveness, and it can be seen that generally more significant predictors lead to bigger prediction success. Interestingly, the models for predicting Self-Efficacy and Feedback have a relatively small number of significant predictors but a relatively high prediction accuracy, and also the significant predictors resulting from the present method are all EEG features, indicating the key role that EEG data plays in assessing the two motivational factors.

Specifically, the differences from Gamma waves and that from the temporal lobe are significant between the two levels of Feedback. As feedback appears in the quiz stage with relevant knowledge information and positive words, this can be explained by the association of Gamma waves with learning, memory and information processing and the fact that the temporal lobe is involved in sensory processing for visual memories, language comprehension and emotion association. As for Self-Efficacy, brainwaves in frontal lobes play an important role in the classification task, and this may be explained by the rich dopamine-sensitive neurons in the frontal brain region related to reward, planning, motivation and short-term memory (Health and Risk 2001). For all the dimensions of motivation studied, it is worth noting that no EEG features from parietal lobe have significant prediction ability, probably because the learning materials used in our study are only involved with the visual information, while the parietal brain lobe is mainly responsible of integration of sensory information from different modalities, including spatial sense and navigation (Goldberg and Goldberg 2001).

Furthermore, it can be found from the significant predictors that Gamma and HighBeta waves are involved in the assessment of most motivational factors studied, and eye gaze features including fixation and saccade domain are useful predictors as well, amongst which fixation number and pupil diameter are the most salient factors at motivation assessment in respect of eye gaze features. It is consistent with prior research about pupil diameter and fixation number, where it has been found that pupil diameter is useful to indicate emotional arousal (Alghowinem et al. 2014)as well as mental effort of viewers when they are doing tasks requiring cognitive effort (Goldberg et al. 1999; Poole et al. 2005), and fixation number is useful indicators of mental

process such as task efficiency (e.g., (Walber et al. 2014)) and interest (e.g., (Kodappully et al. 2016)). However, it is also worth noting that amongst all the motivational factors studied, the features having significant power at motivation assessment contain either both eye gaze and EEG features or solely EEG features for two of them as mentioned (Self-Efficacy and Feedback), corroborating the importance and necessity of introducing EEG data in the task of motivation assessment instead of using sole gaze features as inputs.

In the present paper, we developed a framework that employs multimodal sensing technologies for motivation assessment to enable further personalizing services or interventions towards motivationally intelligent e-learning systems. We designed an experiment to evaluate the assessment process under the framework, where features from EEG and eye tracking and other behaviour data from learner's interaction process with an e-learning system were extracted and selected to infer the level of each motivational factor. We developed a novel method of feature analyses to select the salient features to be input for the classification algorithm by investigating the correlation between each motivational factor and the features including both eye gaze and EEG features, and by identifying the features with significant differences between the two levels of each motivational factor, as well as by referring to our knowledge-based motivation model. Classification task was performed using logistic regression classifier that identified the significance of the model and that of each variable at predicting the response variable. Backward method was applied iteratively by removing the variables with less prediction abilities among the non-significant ones resulting in the most statistically significant prediction models. Finally, the cut-off threshold was decided via ROC curve for each motivational factor to achieve the best prediction success, i.e., classification accuracy.

The experiment results have shown that the combining EEG features relevant to band power, brain lobes and hemisphere asymmetry and eye gaze features relevant to fixation, saccade and pupil is effective for assessing learners' motivation including different motivational factors. More importantly, our empirical data with logistic regression classifier has resulted in significant prediction models for all the motivational factors studied, achieving the assessment accuracy of between 68.1% and 92.8%. The experiment has illustrated and evaluated the proposed framework as well as revealed the features that are most relevant to assessment for each motivational factor. All the EEG and eye gaze feature types studied have played a salient role in motivation assessment, indicating the effectiveness and applicability of combining EEG and eye tracking as a multimodal dataset for motivation assessment. The rationale of the process in the proposed framework can also be applied to other e-learning contexts for a wider group of users and based on a more

generic learner model, which paves the way for motivationally intelligent e-learning systems that can personalize services or intentions to address different users' motivational needs in real time. The present study has its limitations in using a relatively small sample size and thus an overlap of training and testing data to compensate for the small sample size, which, however, increased the risk of overfitting. Future work includes employing a greater sample size to evaluate the framework, integrating new learning algorithms and fuzzy inference techniques, as well as exploring multimodal information fusion other than EEG and eye tracking, which may improve the assessment system from classification techniques or enriched predictors.

Appendix 1

The Significant EEG and Eye Tracking Features (P < 0.05) According to Spearman Correlation.

Motivational Factors	Significant Features	Category
Perceived Usefulness	Gamma-max; Theta-IntraAsymmetry; HighBeta-IntraAsymmetry	EEG
Perceived Ease of Use	LessonDuration	Other- Behaviour
	Gamma-mean; HighBeta-occipital; Gamma-occipital; HighBeta-temporal; Gamma-temporal; HighBeta-frontal	EEG
	FixationNumber-overall; FixationSpatialDensity-overall; Regressions-overall	Eye
Visual Attractiveness	PupilDiameter-average; FixationSpatialDensity-overall	Eye
Feedback	Gamma-max; Theta-InterAsymmetry	EEG
Attitudes Toward School	LowBeta-max; HighBeta-temporal; Gamma-temporal; HighBeta-frontal; Theta-InterAsymmetry	EEG
	FixationNumber-AOI; FixationSpatialDensity-overall	Eye

Motivational Factors	Significant Features	Category
Self-Efficacy	LowBeta-frontal; HighBeta-frontal; Gamma-frontal; Theta- IntraAsymmetry; LowBeta-IntraAsymmetry; HighBeta-IntraAsymmetry	EEG
	FixationNumber-AOI; FixationSaccadeRatio; SamplesOutOfMonitor; AverageSaccadeVelocity; PathVelocity	Eye
Reading Experience	LessonDuration; QuizPerformance	Other-Behaviour
	Theta-max	EEG
	FixationConnectionLength-overall; FixationSpatialDensity-overall; FixationNumber-overall; Regressions-overall	Eye
Learning Experience	Theta-occipital; LowBeta-occipital	EEG
	PupilDiameter-average; PupilDiameter-max; FixationSpatialDensity-overall	Eye
Confirmed Fit	Alpha-max; HighBeta-max; Gamma-max	EEG
	SaccadeLength-overall; SamplesOutOfMonitor; PupilDiameter-average; PupilDiameter-max; FixationConnectionLength-overall	Eye
Continued Use Intention	Theta-min; Alpha-max; Alpha-min; LowBeta-min; HighBeta-min; Theta-mean; Theta-IntraAsymmetry; Gamma-IntraAsymmetry	EEG
	FixationDuration-AOI; FixationNumber-10s; SamplesOutOfMonitor; PathVelocity	Eye

Appendix 2

The Significant EEG and Eye Tracking Features (P < 0.05) According to One-way ANOVA/Kruskal Wallis Test.

Motivational Factors	Test	Significant Features	Category	Motivational Factors	Test	Significant Features	Category
Perceived Usefulness	Krulskal-Wallis Test	QuizPerformance	Other-Behaviour	Attitudes Toward School	One-way ANOVA	Theta-min*	EEG
		HighBeta-IntraAsymmetry	EEG			FixationNumber-AOI*, FixationNumber-overall*, Fixation-ConnectionLength-overall*	Eye
Perceived Ease of Use	One-way ANOVA	FixationNumber-overall*	Eye				
	Krulskal-Wallis Test	LessonDuration	Other- Behaviour		Krulskal-Wallis Test	Alpha-InterAsymmetry	EEG
		Gamma-mean; HighBeta-mean, Gamma-occipital, HighBeta-occipital, Gamma-temporal, HighBeta-temporal, Gamma-pariental, HighBeta-frontal,Alpha-IntraAsymmetry	EEG			FixationSpatialDensity-overall, FixationSaccadeRatio, PupilDiameter-average	Eye
				Self-Efficacy	Krulskal-Wallis Test	Gamma-mean, Gamma-occipital, HighBeta-parietal, Gamma-parietal, LowBeta-temporal, Gamma-temporal, Alpha-frontal, LowBeta-frontal, HighBeta-frontal, Gamma-frontal, Theta-IntraAsymmetry	EEG
		FixationSpatialDensity-overall	Eye				
Visual Attractiveness	Krulskal-Wallis Test	Theta-max, Alpha-max, Gamma-max, Theta-parietal, Theta-InterAsymmetry	EEG				
		FixationSpatialDensity-overall	Eye			FixationSaccadeRatio, SamplesOutOfMonitor	Eye
Feedback	One-way ANOVA	FixationNumber-overall*	Eye	Reading Experience	Krulskal-Wallis Test	LessonDuration, QuizPerformance	Other- Behaviour
	Krulskal-Wallis Test	SamplesOutofMonitor	Eye			Theta-max, Alpha-max, HighBeta-min	EEG
		Theta-max; Gamma-max; Theta-occipital; LowBeta-occipital; HighBeta-occipital; Gamma-occipital; HighBeta-temporal; Gamma-temporal; HighBeta-frontal	EEG			AverageSaccadeLength-10s, FixationSpatialDensity-overall	Eye

Motivational Factors	Test	Significant Features	Category
Learning Experience	Krulskal-Wallis Test	Theta-max, Theta-occipital, LowBeta-occipital, HighBeta-occipital, Gamma-occipital	EEG
		PupilDiameter-average, PupilDiameter-max, FixationSpatialDensity-overall	Eye
Confirmed Fit	One-way ANOVA	Theta-min*	EEG
		FixationConnection-Length-overall*	Eye
	Krulskal-Wallis Test	Alpha-max, HighBeta-max, Gamma-max	EEG
		SamplesOutOfMonitor	Eye
Continued Use Intention	One-way ANOVA	Theta-min*, Alpha-min*, LowBeta-min*	EEG
	Krulskal-Wallis Test	Alpha-max, HighBeta-min, Theta-mean, HighBeta-temporal	EEG
		FixationNumber-10s	Eye

*Normality of distribution and homogeneity of variances are met, tested by Shapiro-Wilk Test (Sig. > 0.05) and Levene's Test (Sig. > 0.05)

Declarations

Conflict of interest The present work is part of the first author's PhD project conducted under the supervision of the second author and the third author. The authors have no competing interests to declare that are relevant to the content of this article.

References

Alghowinem, S., Alshehri, M., Goecke, R., Wagner, M.: Exploring Eye Activity as an Indication of Emotional States Using an Eye-tracking Sensor. In: Chen, L., Kapoor, S., Bhatia, R. (eds.) Studies in Computational Intelligence, pp. 261–276. Springer, Cham (2014)

Arroyo, I., Cooper, D.G., Burleson, W., Woolf, B.P., Muldner, K., Christopherson, R.: Emotion sensors go to school. In: Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling. pp. 17–24: (2009)

Beatty, J.: Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychol. Bull. **91**, 276–292 (1982)

Bixler, R., D'Mello, S.: Automatic gaze-based user-independent detection of mind wandering during computerized reading. User Model. User-Adapt. Interact. **26**, 33–68 (2016)

du Boulay, B., Avramides, K., Luckin, R., Martínez-Mirón, E., Méndez, G.R., Carr, A.: Towards systems that care: a conceptual framework based on motivation, metacognition and affect. Int. J. Artif. Intell. Educ. **20**, 197–229 (2010)

Bradley, M., Miccoli, L., Escrig, M., Lang, P.: The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology. **45**, 602–607 (2008)

Chanel, G., Kierkels, J.J.M., Soleymani, M., Pun, T.: Short-term emotion assessmentin a recall paradigm. Int. J. Hum. Comput. Stud. **67**, 607–627 (2009). https://doi.org/10.1016/j.ijhcs.2009.03.005

Chang, C.-C., Liang, C., Yan, C.-F., Tseng, J.-S.: The impact of college students ' intrinsic and extrinsic motivation on continuance intention to use english mobile learning systems. Asia Pac. Educ. Res. **22**, 181–192 (2013). https://doi.org/10.1007/s40299-012-0011-7

Conati, C.: Probabilistic assessment of user's emotions in educational games. Appl. Artif. Intell. **16**, 555–576 (2002)

Conati, C., Merten, C.: Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation. Knowl. Based Syst. **20**, 557–574 (2007). https://doi.org/10.1016/j.knosys.2007.04.010

Crocker, L.D., Heller, W., Warren, S.L., O'Hare, A.J., Infantolino, Z.P., Miller, G.A.: Relationships among cognition, emotion, and motivation: implications for intervention and neuroplasticity in psychopathology. Front. Hum. Neurosci. **7**, 2 (2013). https://doi.org/10.3389/fnhum.2013.00261

Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. **13**, 319–340 (1989)

Deci, E.L., Ryan, R.M.: The " What " and " Why " of goal pursuits: Human needs and the self-determination of behavior. Psychol. Inq. **11**, 227–268 (2000). https://doi.org/10.1207/S15327965PLI1104_01

Derbali, L., Frasson, C.: Prediction of Players' Motivational States using Electrophysiological Measures during Serious Game Play. In: 2010 10th IEEE International Conference on Advanced

Learning Technologies. pp. 498–502. IEEE: (2010). https://doi.org/10.1109/ICALT.2010.143

Derbali, L., Frasson, C.: Assessment of learners motivation during interactions with serious games: a study of some motivational strategies in food-force. Adv. Hum. Comput. Interact. (2012). https://doi.org/10.1155/2012/624538

Goldberg, M.E.: Parietal Lobe. In: International Encyclopedia of the Social & Behavioral Sciences. pp. 11051–11054: (2001). https://doi.org/10.1016/B0-08-043076-7/03471-9

Goldberg, H.J., Kotval, X.P.: Computer interface evaluation using eye movements: Methods and constructs. Int. J. Ind. Ergon. 24, 631–645 (1999)

Hanif, A., Jamal, F.Q., Imran, M.: Extending the technology acceptance model for use of e-learning systems by digital learners. IEEE Access. 6, 73395–73404 (2018). https://doi.org/10.1109/ACCESS.2018.2881384

Herrador-alcaide, T.C., Hernández-Solís, M., Galván, R.S.: Feelings of satisfaction in mature students of financial accounting in a virtual learning environment: an experience of measurement in higher education. Int. J. Educ. Technol. High. Educ. 16, 2 (2019). https://doi.org/10.1186/s41239-019-0148-z

Horng, W.-B., Chen, C.-Y., Chang, Y., Fan, C.-H.: Driver fatigue detection based on eye tracking and dynamk, template matching. In: IEEE International Conference on Networking, Sensing and Control. pp. 7–12., Taipei: (2004)

Hou, X., Liu, Y., Sourina, O., Mueller-wittig, W.: CogniMeter: EEG-based Emotion, Mental Workload and Stress Visual Monitoring. In: 2015 International Conference on Cyberworlds (CW). pp. 153–160: (2015). https://doi.org/10.1109/CW.2015.58

Jenke, R., Peer, A., Buss, M.: Effect-size-based electrode and feature selection foremotion recognition from EEG. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. pp. 1217–1221., Vancouver: (2013). https://doi.org/10.1109/ICASSP.2013.6637844

John, E., Prichep, L., Easton, P.: Neurometrics: computer assisted differential diagnosis of brain dysfunctions. Science (80-) 239, 162–169 (1988)

Knott, V., Mahoney, C., Kennedy, S., Evans, K.: EEG power, frequency, asymmetry and coherence in male depression. Psychiatry Res. Neuroimaging. 106, 123–140 (2001). https://doi.org/10.1016/S0925-4927(00)00080-9

Kodappully, M., Srinivasan, B., Srinivasan, R.: Towards predicting human error: eye gaze analysis for identification of cognitive steps performed by control room operators. J. Loss Prev. Process. Ind. 42, 35–46 (2016). https://doi.org/10.1016/j.tins.2015.02.002

National Institute of Mental: Health, J.A.: Gene Slows Frontal Lobes, Boosts Schizophrenia Risk, https://web.archive.org/web/20150404205032/http://www.nih.gov/news/pr/may2001/nimh-29.htm, (2001)

Partala, T., Surakka, V.: Pupil size variation as an indication of affective processing. Int. J. Hum. Comput. Stud. 59, 185–198 (2003)

Poole, A., Ball, L., Phillips, P.: In search of salience: a response-time and eye-movement analysis of bookmark recognition. In: Fincher, S., Markopoulos, P., Moore, D., and Ruddle, R. (eds.) People and Computers XVIII — Design for Life SE – 23. pp. 363–378. Springer London: (2005). https://doi.org/10.1007/1-84628-062-1_23

Prichep, L., John, E.: QEEG profiles of psychiatric disorders. Brain Topogr. 4, 249–257 (1992)

Rebolledo-mendez, G., De Freitas, S., Rafael Rojano-caceres, J., Garcia-gaona, R., A.: An Empirical Examination of the Relation Between Attention and Motivation in Computer-Based Education: A Modeling Approach. In: Proceedings of the Twenty-Third International Florida Artificial Intelligence Research Society Conference. pp. 74–79: (2010)

Shroff, R.H., Vogel, D.R.: Assessing the factors deemed to support individual student intrinsic motivation in technology supported online and face-to-face discussions. J. Inf. Technol. Educ. 8, 59–85 (2009)

Sun, S.: An examination of disposition, motivation, and involvement in the new technology context computers in human behavior. Comput. Hum. Behav. 24, 2723–2740 (2008). https://doi.org/10.1016/j.chb.2008.03.016

Tawafak, R.M., Romli, A.B.T., Bin, R., Arshah, A.: Continued intention to use UCOM: Four factors for integrating with a technology acceptance model to moderate the satisfaction of learning. IEEE Access. 6, 66481–66498 (2018). https://doi.org/10.1109/ACCESS.2018.2877760

Walber, T., Scherp, A., Staab, S.: Benefiting from users ' gaze: selection of image regions from eye tracking information for provided tags. Multimed Tools Appl. 71, 363–390 (2014). https://doi.org/10.1007/s11042-013-1390-3

Wang, R., Chen, L., Solheim, I., Schulz, T., Ayesh, A.: Conceptual Motivation Modeling for Students with Dyslexia for Enhanced Assistive Learning. In: SmartLearn '17 Proceedings of the 2017 ACM Workshop on Intelligent Interfaces for Ubiquitous and Smart Learning. pp. 11–18. ACM New York, NY, USA ©2017: (2017). https://doi.org/10.1145/3038535.3038542

Wang, R., Chen, L., Ayesh, A., Shell, J., Solheim, I.: Gaze-based Assessment of Dyslexic Students' Motivation within an E-learning Environment. In: 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation. pp. 610–617. IEEE, Leicester, UK (2019). https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00142

Wang, R., Chen, L., Solheim, I.: Modeling dyslexic students' motivation for enhanced learning in e-learning systems. ACM Trans. Interact. Intell. Syst. 10, 21 (2020)

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

Ruijie Wang is a Lecturer in Psychology at Bournemouth University in the UK. Her research interests lie in cyberpsychology and human computer interaction, particularly digital wellbeing, students' motivation in e-learning context, transparency in online gambling, and the role of human factors in the design and development of intelligent interactive systems. She has multidisciplinary research experience including qualitative and quantitative analyses, statistical modelling, applying machine learning algorithms to automatic assessment of human mind and behaviour based on physiological information such as eye tracking and EEG. She also has experience in user-centred design, usability evaluation, and software prototyping.

Liming Chen is a Professor of Data Analytics and Research Director for the School of Computing at Ulster University in the UK. Prior to joining UU, he was Professor of Computer Science and the Head of the Context, Intelligence and Interaction Research Group (CIIRG) in the School of Computer Science and Informatics, De Montfort University in the UK. He has extensive research expertise and a wide range of research interests with special emphasis being placed on problem solving for real-world applications. His research spans from curiosity-driven theoretical exploration to practical problem-motivated technology and system development covering pervasive computing, data analytics, artificial intelligence, user-centred intelligent systems, smart environment, and digital health and assisted Living.

Aladdin Ayesh is a Professor of Artificial Intelligence in the School of Computer Science and Informatics at De Montfort University in the UK. His research area of interest within Artificial Intelligence is Knowledge Representation and Reasoning in Cognitive Systems. This falls within the crossing boundaries of Knowledge Engineering and Intelligent Agents. His current research focuses on Emotions Modelling, Social Swarms, and Text Mining with their applications in knowledge management, ad hoc networks, e-learning, and games.