Towards Detection of Interest Using Physiological Sensors

Areej Babiker 1, Yahia Baashar 2,3, Ammar Ahmed Alkahtani 3,* , Ibrahima Faye 4 and Gamal Alkawsi 3

Abstract: The positive effects of interest on different aspects, e.g., learning and education, economy, psychological well-being, and social relations, have been widely addressed by many psychological and physiological studies in the last two decades. While the psychological work has investigated this impact of interest theoretically, the physiological studies have focused more on the modulatory effects. However, some studies have addressed both sides of the effects. In this work, we conduct a comprehensive review of physiological studies on interest detection, from different perspectives carried out between 2003 and 2019. A lack of connection between the psychological and physiological studies was identified. Therefore, this paper aims to integrate the unique psychological and physiological aspects and characteristics of interest to form a base for future research by considering the pros and cons of the included studies. For example, considering the two types of interest (situational and individual) the detected interest in learning, gaming, and advertisement’s physiological experiments could be referring specifically to situational interest. Hence, bridging the gap between both physiological and psychological studies is essential for improving the research on interest. Furthermore, we propose several suggestions for future work direction.

Keywords: electroencephalography; classification algorithms; wearable sensors; interest definition; interest levels

1. Introduction

Many theories, definitions, hypotheses, and experiments regarding interest were proposed in the late 19th century. Each had its own perspective, addressed one or more aspects of interest, and was based partially on previous research and findings and/or new experimental work. This work led to the proposal of a variety of definitions, and thus concepts, regarding interest. It is therefore difficult to combine the conclusions of this research in one concept because of the uncontrolled settings of the experimental studies, or often, missing connections between the different fields of study. What could be so interesting about interest that grasped researchers’ attention and led to efforts to study and explore this area of research? According to a large number of researchers, e.g., Krapp [1], Rotgans and Schmidt [2], and Harackiewicz and Hulleman [3], interest is highly related and linked to academic achievement. Many researchers have proposed the idea of interest as a motivational variable and an important and basic element of productive learning. These researchers include Silvia [4], U. Schiefele [5], and Hidi [6]. It has been suggested that interesting information is significantly better remembered and more easily understood than non-interesting information. Renninger and Wozniak [7] have studied the effects of interest on recognition and recall in young children. Interest is also an emotion that is often studied in the context of television commercials and consumers’ responses. It is advised to utilize technology to trigger consumer interest toward specific products in order to increase...
sales and enhance brand image [8,9]. There are other implicit rewards for “interest” that benefit individuals personally by raising their self-efficacy. In addition, interest not only has cognitive effects but also physiological effects.

This paper reviews researches of different objectives that employed different methods in order to provide interested researchers with direction for future studies. The motivation for writing this review was to help researchers recognize the astonishing benefits of triggering, detecting, and maintaining interest in individuals in different aspects of life, including well-being, education, society, and economy. This review also discusses the strengths and weaknesses of some current research on interest from an objective point of view. The paper is structured as follows. Section 2 describes our review method and research questions, along with the nature and level of interests. Section 3 highlights the physiological patterns and features of interest. These patterns include EEG, ECG, Eye tracking and other patterns. The correlation between self-report and physiological sensors is also discussed. Section 4 illustrates the effect of experimental design on the output of the experiment. This includes task, stimulation, number of subjects, and their age. The paper concludes with some recommendations for future studies on the detection of interest using physiological methods.

2. Methodological Choices and Research Questions

Here, we review all the related peer-reviewed articles that address interest or its levels using physiological sensors. Specifically, we reviewed recent work written in English from the past last 13 years, comprising articles written from 2003 until 2019. We intend to include the earliest studies with high impact on the current research and to precisely discuss and analyze recently published innovative findings and insights.

A quick review of the literature indicated that many studies address interest subjectively, but very few studies with objective support. Many psychological, educational, and social studies lack support from neuroscientific research that examine how interest modulates physiological parameters. Encouraging researchers to carry out physiological studies would lead to the publication of research that would not only support, but also explain and clarify ambiguities regarding interest, such as interest as a cognitive construct or emotional state, and the development of interest over time. These studies would also lead to a better understanding of the two common types of interest (personal and situational) and would assist future researchers in discovering solutions to reinforce the presence of either or both of these two types of interest. This would, in turn, improve student learning, well-being, and goal accomplishment.

Hence, this review focuses on presenting research work that examines the physiological effects of interest using the following physiological parameters: electroencephalogram (EEG), electrocardiogram (ECG), heart rate (HR), galvanic skin response (GSR), eye movements, and pressure seat. We aim to answer the following research questions in this review:

RQ1: What basic definitions of interest do the authors refer to?
RQ2: What is interesting about the different levels and types of “interest”?
RQ3: What are the physiological sensors that have been used so far to assess interest and what are their associated patents?
RQ4: What are the possible effects of differences in experiment design, including differences in stimulation, number of subjects, age of subjects, and length of experiment, on the experiment output?
RQ5: How can we improve research related to the detection of interest using physiological sensors?

The terms used in our search included “situational interest detection”, “detection of interest using EEG”, “individual interest detection”, “detection of interest using eye tracker”, “detection of interest using ECG and HR”, “interest in academic content”, “customer interest detection using physiological sensors”, “emotion detection using physiological sensors”, and “recognition of interest. The main purpose of this paper is to review objec-
tive studies on interest. Hence, subjective studies are only discussed to form a basis for the understanding of the physiological (objective) research. Physiological studies were included or excluded based on the following criteria: (i) written in English, (ii) published between 2003 and 2019, (iii) the study addressed the detection of interest or emotions or cognition leading to the detection of interest, (iv) used physiological sensors, (v) had a valid stimulation paradigm to trigger interest, (vi) resulted in the detection of interest or a new output. Studies that did not fulfill these criteria were excluded. The success of each of the studies discussed here was measured by assessing the experiment’s output. These outputs included the participant’s grades [3], if applicable, the participant’s self-report, and the classification of interest as suggested by the physiological parameter of interest. We included nineteen (19) studies in our review. The methodological choices and research questions are summarized in Figure 1.

![Figure 1. Methodological choices and research questions.](image)

2.1. The Nature of Interest

Interest has been extensively studied in the last decades from different perspectives and based on expanded empirical and theoretical approaches. There is probably no uniform definition for interest. However, it is accepted that interest is “a construct that characterizes a person’s special relationship with an object (contents, topic, special subject, object domain, etc.)” in the psychological, sociological, and educational domains [10]. According to many researchers, namely, Renninger, Hidi, and Krapp, focused attention and/or engagement with a specific content reflect interest in the characteristics of that particular content or object [11–13]. Harackiewicz et al. [14] considered interest to be “both a psychological state of attention and affect toward a particular object or topic, and an enduring predisposition to reengage overtime”. Silvia [4] described interest as a facet of human motivation and emotion and suggested using interest to solve practical problems in learning, education, and motivation. In fact, many studies have shown that interest enhances learning (for a summary, see [5,15,16]). Dewey [17] described interest in an object or a topic as “being engaged, engrossed, or entirely taken up”. Furthermore, interest has been discussed by many researchers in connection with individual states, e.g., emotions by Izard [18], attention by J. A. Deutsch and D. Deutsch [19], curiosity by Kashdan and
Silvia [20], flow by M. Csikszentmihalyi [21], and “a thinking state, rather than an emotion” by Ekman [22]. Although there is a valid justification for each consideration, studying interest as related to only one of these phenomena narrows its meaning. Hockenbury and Hockenbury [23] have stated that emotions consist of three different components: physiological response, behavioral expressive response, and subjective experience. Interest appears to have these components: it encompasses a subjective experience as described by Izard [24], physiological and behavioral expressive responses associated with increased cognition, activation, and concentration, and approach-oriented action [4, 18, 25]. These arguments support the idea that interest has both affective and cognitive components. Hidi [6] has suggested that interest involves affect when it is initially triggered, which makes it appropriate to consider interest as an emotion at this stage. Interest is then integrated into one’s cognition as it develops [13, 26].

Krapp [27] studied interest as an independent variable such as certain aspects of learning outcomes, dependent variables such as academic achievement, or hypothetical mediators such as attention and emotional experience. He also discussed the relationships between situational and individual interest, which are widely accepted as different types of interest (for a review, see [3]), and learning outcomes. This will be discussed thoroughly in the next section.

This brief review highlights the variety of publications establishing or studying the numerous meanings of interest. Nevertheless, the studies outlined here generally agree that “interest” is a multidimensional construct, i.e., it has both emotional and cognitive components [28]. Establishing that interest has both emotional and cognitive components is necessary to assess the feasibility of different methods and sensors used to measure interest in the physiological studies that will be discussed in the following sections.

Here we will briefly review studies discussing interest as a physiological construct to explore differences in addressing interest in different domains. Belle et al. [29] used interesting and non-interesting video clips to assess the levels of attention and engagement in study subjects. They used interesting clips in order to keep viewers attentive and engaged with the content presented. Non-interesting clips were intended to induce boredom in the subjects and to reduce attentiveness. The researchers then studied differences in physiological measures following the viewing of the two categories of video clips in the participants. Here, the terms “interest,” “attention,” “engagement,” and “boredom” were used to describe the stimuli and/or the participants’ reactions to the stimuli. Shen et al. [30] described interest as one of eight basic learning emotions in their affective learning model. Interest was also regarded as an academic emotion, along with confidence, excitement, and frustration by Azcarraga in [31]. Other studies have investigated students’ interest in learning in terms of positive emotions displayed in the classroom. For example, Nor et al. [32] studied student interest in learning science by observing four emotions, namely happiness, fear, sadness, and calm. In a study by Shigemitsu and Nittono [33], the authors assessed the level of interest while the subjects watched movies. The authors explained interest as a representation of how strongly an audio-visual experience attracts a viewer’s interest. The rest of the reviewed papers, did not specifically define interest [34]. Rather, interest in these papers was described as an affective state [35, 36] or described in detail in terms of its different levels [37] and/or associated physiological patterns [38]. More details are given in the appropriate sections of this paper.

Obviously, many of the objective studies presented here are not associated with or refer to an appropriate psychological or theoretical model. Many theories and models regarding interest development have been proposed. These models include, but are not limited to the four-phase model for interest development by Hidi and Renninger [13], the integrated model of interest development by Harackiewicz and Hulleman described in [3], the psychology of constructive capriciousness by Silvia [39], the model of domain learning by Alexander [40], and the person-object theory of interest [27, 41]. The lack of connection between physiological and psychological models, and the results obtained based on these models, may cause confusion with similar cognitive and emotional concepts.
For example, the study by Belle et al. [29] used interesting versus non-interesting video clips to assess participant’s attention and engagement. Therefore, it is unclear if the observed physiological effects in this study were related to interest, engagement, attention, or all three parameters at the same time. In fact, as mentioned earlier, interest has been described as cognitive or emotional states, including engagement or pleasant tension. Furthermore, Krapp stated that interest is a construct with cognitive and emotional intrinsic qualities and values [42]. One possible way to distinguish interest from other states, e.g., engagement or attention, is to monitor its development rather than studying it as a single emotional or cognitive state. Therefore, it is advisable to suggest or provide evidence for links between physiological and psychological studies in order to establish an informative, well-established, and well-developed concept. However, as argued by Krapp and Krapp, Renninger, and Hoffmann [27,43], one should not define or limit the meaning of interest to only one specific aspect in order to establish a common definition. Instead, one should develop a theoretical framework referring to different interest phenomena. Therefore, we provide the reader with an overview of the different perspectives regarding interest in this section. Although each model or theory is distinct from the others by different reasoning and arguments, the models still overlap in some areas, e.g., the impact of interest on individuals.

2.2. The Levels of Interest

The vast majority of research on interest based on the psychological point of view suggests that there are two types of interest: individual and situational interest [11,27,44,45]. Individual interest is reflective of the preferences of an individual regarding a particular subject or activity and their contributions to cognitive performance. In contrast, situational interest comprises interest caused by environmental factors in most individuals, regardless of their individual differences [14]. Nevertheless, it is still early to come to a conclusion regarding the types or levels of interest from a physiological point of view. This may be due to the smaller number of the physiological studies on “interest” compared to psychological studies and the lack of adequate measures due to high cost, problems with temporal resolution, long computation time, etc. Here we highlight the numerous methods used to quantify interest. Yeasin et al. [37] proposed using facial expressions to define levels of interest based on the intensity and weight of emotion. The intensity of interest was computed by considering the number of relative facial expression images for an emotion. The weight of interest was computed using a three-dimensional effect space (arousal, stance, and valence) discussed in [46]. Positive values of arousal, valence, and stance were associated with positive expressions, while negative values of arousal, valence, and stance were associated with negative expressions. This suggests that positive and negative values for interest are reflective of high and low interest, respectively. However, it is not clear whether interest can be initiated in the presence of negative affect. Interestingly, Hidi [6] has stated that interest can also operate in affectively negative situations, although it is suggested that affective states associated with interest tend to be positive. Nevertheless, negative affective states have been reported to be associated with interest on rare occasions [47]. Using the interactive system (iTourist) created by Qvarfordt and Zahi [38], the authors studied eye gaze patterns to assess the subject’s interest. They measured the likelihood that a user was interested in an object based on the object’s arousal level and eye-gaze intensity. Eye-gaze intensity is modified by other eye-gaze factors, such as frequency of gaze and relative size to baseline, which either inhibit or increase the object’s excitability. Mota and Picard [35] studied postures to assess the learner’s interest. They classified interest as high, medium, or low. However, the performance of the classifier in the above study was poor in identifying medium interest. In addition, the classifier often had problems differentiating between high and low interest. As a result, interest levels were only classified as “low interest”, “high interest”, or “taking a break” in the final report. Studies of brain activity that use EEG utilized different approaches to assess interest. The study by Babiker et al. [48] used 8 EEG channels and classified interest into high situational interest and low situational interest.
This study classified situational interest (SI) into high SI and low SI with high classification accuracy. Nomura and Mitsukura [49] used a one-channel EEG device. In order to analyze the data, the authors used a commercial analyzer that assigned percentage values to five emotional states, including “interest” (0%: low degree, to 100%: high degree). Interest has also been assessed by combining data of multiple modalities from three different channels: face, posture, and game information [36]. The authors classified interest level as high interest, low interest, or refreshing (they defined refreshing as a state associated with a short break in the task to restore concentration). The above study extended Mota’s work [35], which studied postures to determine the level of interest. This explains the similarity between the reported interest levels in the two studies. The rest of the studies addressed general interest without specifying the level of interest. Though, there is an important physiological parameter that has been overlooked in physiological research despite its significance for the accurate and proper understanding of the phenomenon of interest, namely the effects of external stimuli versus the subject’s internal state while viewing the stimuli. This parameter does not only affect the length and intensity of the subject’s engagement [14] but may also affect the subject’s state of interest in general. This observation is supported by the psychological studies that suggest the existence of two types of interest: individual interest and situational interest [1,2,50]. Therefore, further research should be carried out to clarify the initialization process, the development stage, and the maintenance of these two types of interest. Hidi [6] suggested that the strengths of the two components of “interest” (emotional and cognitive) are associated with the level of interest development. In other words, when the psychological state of interest is initiated, the affect component is stronger than the cognitive component. However, as interest develops and is maintained, both the cognitive and affect components contribute to the experience of interest.

3. Physiological Patterns and Features of Interest

This section of the paper is crucial for understanding interest from physiological and neuroscientific points of view and to support or interpret the psychological studies in this field. The physiological parameters discussed in this review that have been used to assess interest include EEG, ECG, eye tracker, HR, GSR, seat pressure, mouse pressure, and facial expressions. These methods are shown in Figure 2.

![Physiological sensors used in interest research.](image)

The main advantage of using these methods is that they depend on the subject’s physiological parameters, which cannot be manipulated voluntarily. The disadvantage of
these methods is the set-up time that has been significantly reduced with the advance in technology in recent years. Furthermore, some of these sensors limit the structure of the experimental stimulation, e.g., the mouse, which requires a computer application that uses a mouse. The eye-tracker requires certain positions to ensure detection of the pupil of the eye, which is not applicable in real settings.

Similarly, the seat pressure and the facial gestures require a standstill position that is difficult in certain scenarios such as the classroom. In order to overcome the shortcomings of these methods, they are combined in a way that sensor A detects the activities that sensor B cannot detect. These sensors are also used separately or by combining two or more of them depending on the application. The careful selection of a suitable physiological sensor helps to achieve the best possible result.

To facilitate the study of sensors used to measure these parameters, this section was divided into four subsections. Three subsections focused on groups of related sensors, while the last section discussed the correlation between these physiological sensors and self-report. A brief introduction is provided for each sensor in the appropriate subsection.

### 3.1. EEG and ECG Patterns

As shown in Table 1 below, the EEG channels and devices used vary from one study to another. This diversity in equipment specifications could be related to the differences in the purpose of the study, brain areas of interest, availability of devices, cost of devices, and required features. We are interested in patterns of brain activity leading to high detection and/or prediction rates for the state of interest. EEG may be the widely most studied tool in computational studies of emotional systems [51]. This is because new EEG products are more flexible and portable, relatively low cost, and produce high-quality data using fewer channels. In addition, some EEG devices are using dry sensors. EEG devices record brain activity in the form of electrical signals measured in microvolts and/or millivolts. These signals are produced by the brain and are recorded by placing specific electrodes on the scalp. Compared to functional magnetic resonance imaging, EEG produces data with a better temporal resolution, as it can acquire signals in milliseconds. Raw EEG data must undergo pre-processing, which includes, but is not limited to cleaning, that includes noise removal and filtering techniques, in order to be suitable for processing and analysis. Table 1 also lists other sensors that are normally used along with EEG, including ECG, GSR sensors, HR sensors, and electrooculography (EOG), which are known to produce data indicative of the emotional state of the subject. ECG devices record changes in electrical activity across the heart. ECG recordings are performed with each heartbeat over time in a process similar to HR sensor. Increases in HR may reflect cognitive processing, while decreased HR is indicative of attention to environmental stimuli [52]. GSR sensors are attached to the fingers of the subject’s non-dominant hand. Correlations between HR and GSR parameters and valence and arousal are highlighted in several studies, e.g., [53].

EOG is sensitive and is used for correcting and cleaning eye blinking artifacts. All the EEG studies described below utilized questionnaires to confirm, classify, or support the findings based on data from the physiological sensors. We will not discuss this aspect of the studies, as this paper aims to examine physiological parameters. For further review of the subjective measures of interest, the reader is advised to see [54,55]. We, therefore, only mention the use of questionnaires when necessary. Another topic of interest is the baseline session or the resting-state period. These sessions are used to obtain data from the subject in neutral or stimulus-free conditions. The duration of the baseline recording varies from one experiment to another and often ranges between 1 and 5 min in EEG studies. Because the physiological state of an individual is subjective and varies regularly, recording the baseline state is essential when comparing the subject’s state in the presence or absence of a stimulation. Differences between these two conditions (baseline and stimulation) are later used for further analyses and calculations.
Table 1. Detecting interest using EEG and/or ECG.

| Study | Purpose | Participants | Emotion | Technique | EEG Channels | Stimuli | Measurement | Feature Extraction Technique | Classifier | Remarks |
|-------|---------|--------------|---------|-----------|--------------|---------|-------------|-----------------------------|------------|---------|
| Suzuki, J.; et al. 2005 [34] | Examining whether level of interest in visual materials can be assessed using ERP | 8 men, 4 women. Age range: 20 to 25 years | Interest level | EEG-EDG | FP1, FP2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2, M1, and M2 | -180 stimuli (70 ms each) in 4 conditions: pre-control, interesting video, non-interesting video, and post control. Length: 5 min. for each condition. | -Self-report-Behavioral | Time-domain features | -Amplitude of spectra 0.05-60 Hz | N/A | Both target and deviant tones had large P3 amplitudes that were smaller when participants viewed interesting video clips when compared to neutral video clips or still images. |
| Shigemitsu and Nittono 2008 [33] | Assessment of interest level during the watching of movies | 5 men, 13 women. Age range: 20 to 26 years | Interest level | EEG-EDG | 33 scalp sites | -45 stimuli (200 ms each) in 3 conditions: interesting video, boring video, and still picture. Length: 6.5 min. for each condition. | -Self-report | -Amplitude of spectra 0.05-100 Hz. | N/A | -Amplitude of N140 was significantly smaller when viewing interesting clips when compared to boring clips or still images. -Alpha-band power was attenuated in interesting condition when compared to the boring condition. |
| Belle, et al., 2012 [29] | Assessment of engagement and attention systems using ECG and EEG | ECG: 21 EEG:12 Age range: NA | Cognitive attention during task execution | EEG-EDG | FP1 and FP2 | Interesting and tedious series of pre-selected video clips Length: 20 min. | Depended on the preselection of videos | -Amplitude of spectra 0.2-55 Hz | C4.5, classification via regression and random forest | Accuracy using random forest: EEG = 86%, ECG = 77% Accuracy using C4.5: EEG = 81%, ECG = 74.22% Accuracy using classification via regression: EEG = 83%, ECG = 72% |
| Azcarraga and Suarez 2013 [31] | Prediction of academic emotions based on brainwaves, mouse behavior, and personality profile | 14 male subjects, 11 female subjects. Age range: 17 to 20 years | Interest, excitement, frustration, and confidence | EEG-mouse clicks and movements | AF3, F7, F3, FC5, T7, T3, O1, O2, P8, T8, FC6, F4, F8, and AF4 | -Apliusx math learning software -Solving 4 algebra equations | Computation of signal value at each EEG channel | -Amplitude of N140 was attenuated in interesting clips or still images. | MLP, SVM | -Accuracy: 43% for SVM, 65% for MLP -Accuracy when including mouse movements: 51% for SVM, 73% for MLP -Average accuracy when using outliers and MLP: 92% |
| Fukai et al. 2013 [56] | Design of a preference acquisition detection system using EEG | 28 men, 26 women. Age range: 20 to 50 years | Degrees of interest | EEG | FP1 | 10 types of television commercials. Length: 15 s each | 5 questions after watching all 10 television commercials | Amplitude spectra of 4-22 Hz | Euclidean distance was a factor for classification | Multiple regression analysis was used to construct the estimation model to estimate the results of the questionnaires using EEG features. | -For spontaneous recall, women scored higher on recall and appreciation than men in the perfume category. -Physiological tools, especially ECG, are effective in determining customer preference and memory, and can result in cost reduction. |
| Vecchiato et al. 2014 [8] | Investigation of consumer gender differences during observation of television commercials | 16 men, 12 women. Age range: 20 to 24 years | Emotions (positive vs. negative), Interest, and memory | EEG-EDG-CSR | 21-min. documentary interrupted 2 times. Each interruption contained 6 commercial video clips. Length: 30 s for each clip | -Interview-Self-report | -ICA -Referenced using CAR-IAF was calculated. -Amplitude of spectra 2-30 Hz | Z-score analysis | - |
| Study | Purpose | Participants | Emotion | Technique | EEG Channels | Stimuli | Measurement | Feature Extraction Technique | Classifier | Remarks |
|-------|---------|---------------|---------|-----------|--------------|---------|-------------|-------------------------------|------------|---------|
| Nor et al., 2015 [32] | Detection of student interest in learning science using four emotions | 7 boys, 8 girls. Age range: 10 to 11 years | Fear, sadness, calmness, and happiness | EEG | 16 channels | -Emotional video clips -One set of mathematics and science tests | 2-D (ASM) | MFCC | Neural network (MLP) | Authors claimed that the presence of negative emotions (sadness and fear) prevented interest in answering the test questions. |
| Nomura and Mistukura 2015 [49] | Detection of latent emotions induced by television commercials and examination of their effect on consumer memory using EEG | 15 men, 11 women. Age range: 20 to 50 years | Liking, interest, concentration, drowsiness, and stress. | EEG FP1 | 100 television Commercials divided equally into two categories: admired and ordinary television commercials. Length: 15 s and 30 s | Two spontaneous questions after every 10 television commercials. One question one month later | FFT and the amplitude spectra of 1–30 Hz | Instant analysis using KANSEI analyzer | Both categories had the same memory retention in the long run. EEG data indicated that different emotions were correlated based on characteristics of the television commercials. |
| Morillo et al., 2016 [9] | Proposition of a discrete classification technique to recognize interest in television advertisements | 5 men, 5 women. Age range: 26 to 63 years | Like and dislike | EEG | AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 | 14 short television advertisements. Length: 20 s to 60 s long | -Self-report | | |
| Mateusz, P., 2018 [57] | Analyzing the level of interest in social issue advertising | 16 men, 14 women. Age range: 18 to 50 years | Levels of interest | EEG | Fp1, Fp2, F7, F3, Fz, F4, and F2 | A film with 2 advertising breaks. Each break consisted of 6 advertisements. | Analyzing video scene of the advertisement with respect to index of interest. | Using the average poser of EEG to obtain Index fo interest. | N/A | The study suggested the use of index of interest to identify the level of interest during advertisement. |
| Babiker, A., et al. 2019 [48] | Measuring Situational interest of students in classroom | 43 Age range: 20 to 25 years | Situational interest | EEG | Fp1, Fp2, F3, F4, (O1, O2)/(Pz, P4) | Classroom lecture given by university lecturer. | -Self-Report -video camera -informal interview | -EMD to extract the features -ROC to select the best ones | SVM -kNN | -In dataset 1: SVM 93.3 %, kNN 87.5%. -In dataset 2: SVM 87.5%, kNN 86.7%. -Gamma and delta could be efficient in detecting situational interest. |
| Shestyuk, A. Y., et al. 2019 [58] | Using EEG measure of attention, memory, and motivation to predict the population level TV viewership and Twitter engagement | 331 Age range: 21 to 54 years | Interest in program content and Engagement with social media | EEG | Fp1, AF3, F7, F3, FC5, FC3, Fpz, Fz, FCz, Fp2, AF4, F4, F8, FC4, FC6 | -TV episode -Twitter | -spectral decomposition of EEG data -power normalization by z-score | N/A | The study highlighted the viability of using EEG measure to predict the success of a TV program and determine the cognitive process related to engagement with such content. |

Abbreviations—N/A: not applicable, EEG: electroencephalogram, ECG: electrocardiogram, GSR: galvanic skin response, SVM: support vector machine, KNN: K-nearest neighbor, MLP: multi-layer perceptron, ERP: event-related potential, FFT: fast Fourier transform, MFCC: mel-frequency cepstral coefficient, ASM: affective space model, DWT: discrete wavelet transform, ICA: independent component analysis, CAR: common average reference, IAF: individual alpha frequency, EMD: empirical mode decomposition, ROC: receiver operator characteristics.
Vecchiato et al. [8] have referred to interest as an approach-withdrawal (AW) index. They computed this index by assessing the difference between the average power of the alpha band recorded from left and right EEG frontal lobes channels. Specifically, increased AW indicates an increase in the subject’s interest and vice versa. We can interpret the relationship between this measure and interest by the two fundamental motivational dimensions: approach and withdrawal (also called avoidance). The approach system may be related to appetitive behavior, interest, movement toward a goal, or the “I want” response. In contrast, the avoidance system or withdrawal is associated with defensive withdrawal, pain, failure, or the “I do not want” response [6]. This is discussed by Larsen and Augustine [59], who stated that positive effect promotes exploratory behavior and that if an individual is experiencing positive affect, he or she would seek to broaden and continue this experience. The authors also stated that “the experience of interest encourages one to approach novel situations, ideas, and individuals that are related to the object of interest”.

Peterson et al. [60] found a relationship between greater left frontal activation and approach motivation, while greater right frontal activity is associated with withdrawal motivation [61]. This is true, regardless of the valence of emotion [62]. Many studies have reported anger, which is an approach-oriented negative emotion, to be associated with relatively greater left frontal brain activation than right frontal activation. Thus, interest may be described as effect reflective of high levels of approach, as stated by Harmon-Jones [62]. Davidson et al. [63] have proposed that greater left-sided dorsolateral activity might be related to approach-related, goal-directed action planning. Discussion of emotional processing in the prefrontal and frontal cortex and the approach-withdrawal model by Davidson [64] explains the use of the AW index by Vecchiato et al. [8] and the utilization of only the frontal channels in EEG. It is worth mentioning that greater left frontal activation than right brain activation was not observed in all studies. As proposed by Harmon-Jones et al. [65], this may be because the stimulation failed to involve the subjects personally and/or failed to provide the expectation of an action, which is essential for approach motivation.

The alpha band can be used to assess interest, as it can be used to determine the AW index. Recently, it was reported [66] that alpha asymmetry measures yielded significant correlation patterns for frontal and parietal asymmetry. The findings of the above study indicate that greater behavioral activation system activity is related to greater left-sided activation. However, Vecchiato et al. did not report whether data from bands other than the alpha band can be used to obtain similar AW index values to quantify interest. It likely takes less time to use data from a single band, although using more than one band may enhance and improve the results.

After filtering (the data were band-pass filtered at 2–30 Hz), Vecchiato et al. segmented the data and analyzed their results using independent component analysis, wherein a common average reference was computed, and the individual alpha frequencies were acquired as suggested by Klimesch [67]. The authors found differences related to the way the two video clips were observed between men and women. Women had higher interest in some scenes of both clips, while men were interested in other scenes. This led to variegated patterns for the different cerebral variables.

Belle et al. [29] used two EEG channels (FP1 and FP2) and ECG measurements to assess cognitive attention. The main objective of the study was to investigate whether ECG signals can be used to measure an individual’s attention and focus, and EEG was used as a benchmark. In this case, ECG could be more practical than EEG because it is more portable and reliable during a variety of activities and executive tasks. However, EEG produces a larger amount of accurate data.

ECG data were pre-processed, and statistical features such as mean, sum, and mean of autocovariance were extracted. After pre-processing, wavelet transform was carried out, then, the data were decomposed, and features such as standard deviation, entropy, mean of frequencies, and variance of probability distribution were extracted (for extensive details of
the processing and decomposition procedures, see [29]). All the features from both channels were fed into three different classifiers: C4.5, random forest, and classification via regression to classify the data into two categories: attention and inattention. These classifiers were selected because the numbers of training and testing data points were not largely used. The highest accuracy was achieved by the random forest classifier and was 77% and 86% for ECG and EEG data, respectively. The above experiment demonstrated the superiority of EEG for the detection of cognitive attention (by almost 9%) when compared to ECG. However, the findings of the study also illustrated the feasibility of ECG in determining the attentional status of an individual and showed that fluctuations in attention have recognizable and considerable effects on cardiac rhythm.

One issue that was not addressed by the authors is the sequence of the presented stimulation. The stimulation comprised 20 min of interesting clips, including high-speed car chases and popular movie scenes, followed by 20 min of non-interesting or specifically boring clips containing repetitive video and still images. It is not clear whether the same accuracy values would be obtained if the stimulation sequence were reversed, i.e., starting with the boring clips and then moving on to the interesting clips. Furthermore, the manner in which the clips were rated as interesting or non-interesting was not described.

Another issue is the detection of attention. Considering that the stimulations used were categorized as interesting or non-interesting, along with the composition of the stimuli in each category, it is not clear why the authors used the term cognitive attention instead of interest. For example, the study in [37], used interesting stimuli that contained clips from movies to assess interest or the level of interest in subjects. Furthermore, non-interesting stimuli that contained clips with still images have also been used by researchers to measure the level of interest [33]. Since interest is defined as focused attention and, as described earlier by Dewey, as “being engaged, engrossed, or entirely taken up with” an activity, object or topic [14], is it therefore reasonable to refer to the state resulting from the viewing of the interesting stimuli as “interest” and to the attention resulting from interest in the presented stimulus as selective attention? We may not be able to accurately answer this question based on the limited information provided. Although this study clearly indicates that cognitive and physiological effects measured by EEG and ECG have common aspects while the subject views interesting and non-interesting stimuli.

Fukai et al. [56] used a laboratory-made device that utilized one frontal channel placed over the frontal region of the brain (FP1). The author selected the frontal region because EEG change in this region is clearly noticeable [56]. The FP1 is located over the left frontal lobe of the brain according to the international 10–20 system. Therefore, the above discussion regarding the definition of interest as an approach motivation also holds true for this work. The raw EEG data collected each second were processed using Fourier transform. After transforming the EEG data to their respective frequency components, a moving average filter was used to account for perception time, remaining time, and measuring error. Cluster analysis was then used to classify the subjects based on the amplitudes of EEG spectra with frequencies of 4–22 Hz. Finally, multiple regression analysis was used to quantitatively evaluate the degree of human interest. The estimation model was constructed based on training data using EEG features and the results of questionnaires. Based on the analysis results, the authors considered the combination of waves at 6 and 10 Hz to be associated with “favorite”, the combination of waves at 11 and 16 Hz with “dislikeable” for haptic sense, the combination of waves at 4 and 7 Hz with “favorite”, the combination of waves at 9 and 14 Hz with “dislikeable” and dozing was associated with waves at 5, 6, 9, and 16 Hz. While the study addressed the effects of television commercials, which are normally visual stimuli, the authors did not clarify the meaning or usage of the terms haptic and olfactory senses in the results section. Thus, it is not clear if the combination of waves with these frequencies is related to the type of television commercial product, e.g., those with odor, which triggers the olfactory organs, or if the authors are referring to other uncited studies that they had carried previously. Comparing the results of the questionnaire used by the
subjects to rate their interest levels to those obtained using the proposed system yielded an accuracy of 79.6%.

Nomura and Mitsukura [49] carried out their experiment using the same EEG device described in the above report. The purpose of their study, however, was to detect latent emotions experienced by subjects while watching television commercials. The authors specified these emotions as like, interest, concentration, drowsiness, and stress. The experiment included 100 Japanese television commercials falling into one of two categories: award-winning and non-award-winning. The commercials were randomly ordered and differed between subjects. Both categories contained television commercials of two different lengths: 15 s and 30 s. Questionnaires were used to evaluate the subjects’ emotions after they had watched 10 television commercials. One month after viewing the commercials, the subjects were asked to name television commercials that they found memorable. The EEG device “KANSEI” was used to calculate percentage values for the aforementioned five emotions (0%: low degree to 100%: high degree) by converting the raw EEG data to a frequency domain spectrum using fast Fourier transform.

This study is thus far the only one that investigated the effects of the stimulus length on the subject’s emotions while viewing a stimulus. Fifteen-second television commercials induced increases in the subjects’ stress levels. In contrast, 30-s commercials decreased the subjects’ stress. The authors argue that this is because the 15-s commercials are required to present sufficient information within a limited time, while the 30-s commercials have double the time to create and present emotional content, such as a storyline. The authors found significant correlations for only two emotions out of the five studied (including interest): concentration and stress. Therefore, the study focused on concentration and stress, and did not discuss interest or the other two states (like and drowsiness) in depth. Hence, the significant contribution of this study to the field could be the procedure and experimental design, but not the outcomes or the result. The importance of the correlation between stress and concentration, which led to the exclusion of the other non-significant correlations and their respective emotions, was not clarified. Moreover, since the device was used to perform all the processing, the reliability of the calculations and classifications need to be inspected.

Suzuki et al. [34] examined the applicability of another method used to assess the level of interest. P3 is known to reflect mental workload and attention paid to a task [68]. The authors used this information to extend the work of Rosenfeld et al. [69] by using a three-tone oddball task to assess the level of interest in the subjects. In this paradigm, when an interesting or non-interesting video is played (primary task), the subject is asked to simultaneously respond to an occasional presented auditory stimulus. This technique is based on the idea that there is a fixed capacity for attention. Thus, the greater the attention required by the task, the less the resources available to probe stimulus. This, in turn, will cause a reduction in P3 amplitude.

The authors of the study differentiated between interesting and neutral states using video clips that were interesting or neutral (neither non-interesting nor boring). The use of interesting vs. neutral stimuli is thought to allow investigation of the effects of interest on the brain in a more proper manner than using interesting vs. boring stimuli, because “boring” stimuli might be biased. EEG data were recorded using 21 channels covering most brain areas (frontal lobe, central, parietal lobe, and occipital lobe) besides recording EOG data. Since the experiment used probe stimuli, the correct responses were defined and used to obtaining event-related potential (ERP) waveforms. Scalp topographic maps containing P3 data were drawn after pre-processing and computation. Statistical analysis was then performed to confirm the finding that P3 amplitude decreased when the subjects viewed interesting video clips when compared to neutral stimuli. This finding is consistent with that of Rosenfeld et al. [69], who reported that P3 is a sensitive indicator of the subject’s interest in video clips.

The authors found a significant difference between the neutral and control conditions only for the deviant P3a, and not for the P3b. The authors, therefore, claimed the superiority
of the deviant P3 over the target P3 for the assessment of the amount of residual processing resources. This may be explained by the fact that the task demand of the target stimulus, which the subjects were required to respond to, was higher than that of the deviant target, which the subjects could neglect. This would have led to the allocation of more processing resources for the target stimuli. Therefore, it is hypothesized that the frontocentral P3a component might be more sensitive to resource availability than the parietal P3b component. Nevertheless, further research is required to investigate and confirm the above hypothesis, which opens a new direction for the investigation of interest.

According to the cognitive model discussed by Polich [70], the P3a originates during task processing by stimulus-driven frontal attention mechanisms, while the P3b originates due to temporal-parietal activity associated with attention and seems to be related to subsequent memory processing and storage. In some way, one may thus argue that prior knowledge, which requires memory processing and updating of operations to relate given information to stored information, is not necessary for the triggering of interest. While many psychological studies have reported a weak or no relationship between interest and prior knowledge e.g., [71–73] other studies have reported significant correlations between interest and knowledge, e.g., [74]. There is debate regarding the manner in which prior knowledge affects interest, although the effects of interest on knowledge are acknowledged. The four-phase model suggested by Hidi and Renninger [13] describes the first two phases of interest development as composed of focused attention and positive feeling, while the latter two phases consist of knowledge and stored values, along with positive feelings. Research on interest should carefully consider this issue, as low or high levels of knowledge are expected to decrease the level of interest, while a moderate level of knowledge is reported in some studies to generate interest [75,76] (for a review see [77]).

The study conducted by Shigemitsu and Nittono [33] suggested the utilization of N140 and P300 elicited by a vibratory probe stimulus to assess interest in subjects presented with particular content. In this study, the authors collected EEG and EOG data from 18 subjects while they watched interesting videos, boring videos, and still images. Unlike the Suzuki study, wherein auditory probe stimuli were used, the authors of this study used a somatosensory (vibratory) probe stimulus to account for the masking problem, which may occur when auditory probe stimuli are used. Specifically, it is likely that the auditory probe might be masked when other auditory materials are presented at the same time. After data pre-processing and filtering, ERPs were calculated separately for each participant and condition. The N140 and P300 were then defined and computed, and the peak amplitude and latency were measured (Cz for N140 and Pz for P300). The mean power in the alpha band (8–12.5 Hz) was computed at the Oz site (occipital midline).

As expected, the subjective ratings indicated that subjects allocated more attention to the interesting clips and less attention to the probe stimuli. This was not the case for boring clips, which drew less attention. The N140 was significantly smaller in amplitude when the subjects viewed interesting clips when compared to boring clips. As noted by Kida et al. [78], the amplitude of N140 is an indicator of a passive shift in attention. This to some extent explains the change in N140 during the viewing of interesting videos, which captured the subject’s attention, when compared to the boring videos or the still images. Interestingly, the P300 amplitude decreased when the subjects viewed interesting videos when compared to boring clips or still pictures. However, the difference was not significant between the two conditions. This suggests that the N140 elicited by vibratory stimuli might be more sensitive in measuring the attention resulting from the viewing of interesting content than the P300.

The authors further assessed the effects of interesting content on the alpha band. They found that alpha amplitude was attenuated in the interesting condition when compared to the boring condition or the still picture condition. This finding was in line with the previous findings of Smith et al. [79]. The authors proposed the use of N140 amplitude elicited by vibratory probe stimuli, along with the power of the alpha band recorded from the occipital lobe using EEG, as a useful index of the subject’s interest in audio-visual content.
Azcarraga et al. [31] used an Emotiv EPOC device with 14 channels. They achieved an accuracy of 73% in the detection of interest, excitement, frustration, and confidence when using EEG and mouse behavior data. In this study, the difference between the raw EEG value and the mean value of the baseline data was normalized and segmented into 2-s bins. Each segment was then treated as a single instance. Mouse data, including number of clicks, distance traveled, and click duration, were used in addition to the EEG data. The EEG data, data on mouse clicks, and self-report data were synchronized and merged. The EEG data and data on mouse clicks were labeled based on the self-report data. Data were then fed to two different classifiers: a support vector machine (SVM) and a multi-layer perceptron. The F-measures were higher when using brainwaves only than when only using mouse data. The F-measure was highest when both modalities were combined. The F-measure reached 73% for the average of all emotions when using the MLP classifier and was lower (51%) when using the SVM classifier. It should be noted that data from 16 out of the 25 subjects were valid for feeding into the classifiers. This is because the number of instances of the occurrence of the four emotions was required to be balanced. Interestingly, the F-measure was improved when the authors used “outliers”, which were defined as values deviating from the mean by at least one standard deviation. The F-measure reached 92% for the average of all emotions when using the MLP for combined modalities (brainwaves and mouse) data and reached 97% for “interest.” According to the authors, interest was likely the most frequently occurring emotion during the learning process in this particular study and was related to the subject’s engagement.

The authors hypothesized that the high accuracy values obtained when using the outliers were likely due to the smaller sample size because only the outliers were used in the calculations. However, this is probably inaccurate, as the MLP, SVM, and machine learning algorithms, in general, are known to produce better results when larger sample sizes are used. An explanation could be that the outliers were not, in fact, mistakes or false data from the devices, but rather real data that deviated from the mean. For instance, an interesting subject might be so engaged at some points of the presentation that he/she performs mouse click movements different from the earlier movements. The same effects may occur in the EEG data. The authors of the study suggested the utilization of outliers when assessing emotions, and specific interest, to achieve high accuracy.

The MLP was found to be superior to the SVM in cases where the data are known to have regular and unpredicted variations. In addition, mouse data alone were shown to be insufficient for the accurate detection of the emotional states of learners, which suggests the need for additional measures to assess brain activities. Using combinations of modalities was found to be a powerful strategy and led to more accurate results, as will be discussed in the following sections.

Morillo et al. [80] also used the Emotiv EPOC to assess whether the subjects liked certain television advertisements based on their interest in the content. The authors defined three levels of interest and distraction: High (high interest and low impact of distraction), Medium (high interest and noticeable impact of distraction), and Low (low interest and high impact of distraction). This information was added to the profiles of the subjects, which also contained the subjects’ ratings of how much they liked each advertisement on a scale of 1 to 5. The EEG data were pre-processed, filtered, and then divided into two categories: dislike (rated lower than 3), and like (rated higher than 3). The data were then fed into three different classifiers: neural network, decision tree-based approach, and Ameva-based approach (for more details, see [80]). The neural network was superior in classifying the data. The authors also proposed a new approach based on the Ameva algorithm that allowed them to discretize the system’s inputs. The proposed discrete classification technique based on the Ameva algorithm in this study required less computation time and had higher accuracy than C4.5 and comparable accuracy to neural the network.

Nor et al. [32] studied interest by investigating emotions. The authors suggested four emotions, namely happiness, sadness, calmness, and fear, to be used to determine student’s interest in science and math. The authors claimed that it might be possible to predict the
subject’s interest using neural network MLP to analyze the features of EEG activity recorded while answering mathematics and science questions. First, an effective space model (ASP) was generated for each subject based on EEG activity recorded while the subject viewed images from the International Affective Picture System (IAPS). These images depicted the emotions of happiness, fear, calmness, and sadness. The ASP information was then used to encode the subject’s emotion during the science and mathematics tests. Emotional responses were elicited using Bernard Bouchard’s synthesized musical clips and Gross and Levenson’s movie clips. Each clip lasted for one and one-half minutes, during which time EEG activities were recorded. As in the previous experiments, subjective responses and ratings were obtained after the viewing of each video clip. The students were then required to answer scientific and mathematical questions in two separate tests. Raw EEG data were pre-processed, filtered, and the features were extracted using mel-frequency cepstral coefficients, which then were fed into a neural network (MLP) to classify the data into two categories: valence and arousal. According to the authors, positive valence and negative arousal were identified as calmness, positive valence and positive arousal were identified as happiness, negative valence and negative arousal were identified as sadness, and negative valence and positive arousal were identified as fear. Because only data from one subject were analyzed, the results of this study are not discussed to avoid bias and perhaps the presentation of inaccurate information. Even though the importance of a learner’s emotional state to effective learning is highlighted in recent research, the authors did not elaborate on or explain the reasons for choosing the above four emotions and not other emotions, e.g., hopefulness, confusion, or frustration, which are well known to affect the learning process and are referred to as learner’s emotions in several studies, e.g., [33,81].

The discussed studies in this section showed the feasibility of using EEG to detect interest. The number of EEG channels, in this case, ranged between 1 and 33 channels. The majority of these channels were placed in the frontal brain regions which are claimed to play a significant role in detecting interest. Several methods and classifiers were employed for analyzing and classifying EEG data and are summarized in Table 1 with their respective accuracy. ECG modality was also employed successfully to detect interest, but it had lower performance than EEG. Differences in experimental designs should be considered in interpreting the result because certain stimulations might be more effective than others, e.g., using video clips from movies compared to using tests and educational content.

3.2. Eye-Tracking Patterns

Eye-tracking has been used in many applications because of improved system capabilities, low cost compared to other devices, and the possibility of designing algorithms according to one’s needs. Eye-tracking has been used to detect the point of interest that the participant is looking at, interesting texts read by the participant, and interesting information presented to the participant while surfing the web, among other features. Based on the selection criteria, the studies included here were those that directly detected the interesting feeling or cognitive state and/or its levels and contributed to the characterization of interest. Table 2 lists the studies presented here, wherein the authors attempted to detect interest in two different areas: tourism (commercial) and learning.
Table 2. Detecting interest using eye tracker.

| Study                      | Purpose                                                                 | Participants                  | Emotion                              | Technique                      | Stimuli                          | Measurement              | Features                      | Classifier                  | Remarks                                                                 |
|----------------------------|-------------------------------------------------------------------------|------------------------------|--------------------------------------|-------------------------------|---------------------------------|---------------------------|-----------------------------|-----------------------------|--------------------------------------------------------------------------|
| Qvarfordt and Zahi 2005    | To examine the possibility of using a computer to converse with the user based on eye-gaze patterns | 10 men, 2 women. Age range: 27 to 41 years | Interest and interest level          | -Eye tracker                   | iTourist contains information about 36 places shown on a map, including hotels, restaurants, etc. Objects have a number of spoken utterances. | -Post-test questionnaire   | Eye gaze intensity          | All functions in iTourist are implemented in C++. | -Subject’s interest level was measured based on 2 variables: IScore (determining the object’s arousal level) and FIScore (measuring how users maintain their interest in the object). -Adapting to an individual’s preferences is a critical factor in eye-gaze multimodal information systems. |
| Asteriadis et al., 2009    | Estimation of user behavior based on eye gaze and head pose in an e-learning environment | 20 students. Age range: 8 to 10 years | Frustrated/struggling to read, distracted, tired/sleepy, not paying attention, attentive and full of interest | -Eye tracker                   | -Head pose                     | 20 video segments, between 800 and 1200 frames each | Dyslexia experts labeled the data. | Position and shape of mouth, eyes, eyebrows, and hands, and the features related to these parameters | Neuro-fuzzy network         | Accuracy of detecting the states of the users was 100% for attentive and 72% for inattentive users, with an average absolute error between the output and annotation of 0.11. |
| Sungkur et al., 2015       | Enhancing learning experience using eye tracking to determine interest and behavior of users | 13 subjects. Age range: Not reported | Interest and attentiveness           | Eye tracker                    | Not applicable                  | Not applicable             | Not applicable              | Not applicable              | -To keep the user’s attention, the system produces an alarm when the eyes are not detected. -True acceptance rate for detecting eyes and pupils was 75%. -The system needs to be tested in real settings for the detection of the learner’s interest and attentiveness. |
In [83], the authors developed a system to detect learner’s eyes in order to enhance the learning experience. The system uses OpenCV tools to first detect the eyes and then the eye pupils. It then uses this information to detect and interpret the learner’s behavior, interest, and attentiveness. This system detected interest by recording the direction toward which the learner gazed and by counting the number of fixations by the learner in a particular region. To keep the learner’s attention, the system sounded an alarm when the eyes were not detected within a specified time period. In addition, the area between the user and the web camera was defined and recorded, so that continued fluctuations in the data collected from this region were indicative of the inattentiveness of the learner. The time period between the loss of eye detection and the re-detection of the eyes was also recorded. If this time period was frequently longer than 1 s, is the system assumed that the individual was probably not focusing or that his/her attention was distracted or diverted. However, the system sometimes lost detection of the learner’s eyes even when the learner was still in front of the web camera. This loss was identified by its duration, which was shorter than one second. The advantages of this system were: the achievement of a true acceptance rate of 75% for detecting eyes and pupils, low cost, and compatibility with any computer equipped with a web camera. The system nevertheless had some limitations, e.g., it could not detect the eyes if more than one head was shown in the image, it was affected by luminance, and was inoperative when the background had high illumination. The irises of the eyes could not be detected when they were located at the corners of the eyes. Moreover, although the system was tested for the detection of eyes and pupils, it also requires testing in real learning settings to assess the interest and behavior of a learner.

The learning system proposed by Asteriadis et al. [82] is a full extension of the LTSA IEEE learning system initially described in [84]. This system consisted of 3 different entities: processes, stores, and flows. It performed all the analyses and computations of the learner’s states and the durations of the states online. This information was then stored in the learner’s records for future reference. One advantage of this system was its ability to build and update different profiles for different learners based on both their current behavior and their stored profile. In other words, the system considered the individual differences of the learners and their different modes of behavior and learning at different times. Another advantage of the system was providing feedback to enhance the learning process. For instance, a highlighted text modified based on the subject’s state is shown to an inattentive subject.

The architecture of the model consisted of head pose estimation, eye gaze estimation, positions of hands during the video, a fraction of the inter-ocular distance between frames (by monitoring the variation in the distance between the user and the computer screen), and facial features (eye blinking, frowning, mouth opening, eyes wide open, or not moving at all). These features were fed into a neuro-fuzzy inference system, which then decides the status of the user as frustrated, struggling to read, distracted, tired/sleepy, not paying attention, attentive, or full of interest. In this experiment, the subjects had difficulty decoding written text (dyslexia). As a result, dyslexia experts annotated the video clips used as “ground truth”. The actual features of the testing set were the region around the eyes, eyebrows, mouth, and/or features related to hand location and movement. Only two states were detected and reported. The accuracy was 100% for attentive individuals and 72% for inattentive individuals, with an average absolute error between the output and annotation of 0.117. The state of the user was noted as attentive if the output value was greater than 0.5 and inattentive otherwise. However, the algorithm was still incapable of detecting some features, such as “frozen lips”. Therefore, some states were not accurately detected or not detected at all.

The work by Qvarfordt and Zhai [38] is interesting because it utilized a real-life application. The interest of subjects was detected based on eye-gaze patterns in an interactive system called iTourist, which is used for city trip planning. Two types of eye-gaze patterns associated with a person’s interest in an object on the screen were identified. The first pattern comprised accumulated time and high intensity of a participant’s gaze toward
an object on the screen. The second pattern comprised the gaze of the participant on the
surrounding objects related to the object of interest. The authors also found that frequent
eye-gaze switching between two objects indicated an interest in the relationship between
the two objects, e.g., participants were found to switch eye gaze between a hotel and a
museum when they were interested in the museum and wished to book a nearby hotel.
The system, which was implemented fully in C++, provided information to the users
regarding an interesting place based solely on eye-gaze information. However, it had some
limitations. For example, it falsely provided information that the participants were not
interested and provided this information for a while before stopping. In addition, the
system often missed providing information regarding places of interest. Nevertheless, the
overall reaction of participants was quite positive.

The studies discussed above showed the potential of using an eye-tracking system to
detect subject's interest. The availability of such equipment and the high obtained accuracy
especially in educational and tourism perspectives makes it promising and opens the door
for more creative ideas. The relatively few numbers of studies compared to EEG modality
suggest the need of further research in this area. However, the studies also showed that
eye-tracking needs to be combined with other modalities to improve its performance. This
is because the eye movement is an indicator of several conditions, and using it separately
could result in an inaccurate result.

3.3. Patterns Obtained Using Other Sensors

Several other types of sensors are used to detect interest in subjects while they view
stimuli. These sensors may include pressure sensors that are placed on the subject’s
seat, conductance bracelets, and pressure mouse devices. The authors of the following
studies, which are listed in Table 3, evaluated the interest of subjects while they played an
educational game, watched movies and used a tutoring system. More details regarding
these studies are described below.

We discuss the use of multimodality systems in this section. The studies discussed
previously in this review used telemetry biopotential measures such as EEG and ECG,
which measure the electrical potential between two electrodes, or eye-tracking. Here,
pressure mouse devices, pressure seats, cameras, and conductance bracelets are used to
measure variables related to interest. There is thus a diversity of physiological sensors used
to detect interest.

Kapoor et al. [85] proposed the use of a multi-sensor affect recognition system to
classify interest and disinterest in children while solving an educational puzzle on a
computer. The approach was based on using a mixture of Gaussian processes to analyze
facial features, current posture, and level of activity, as determined using a camera and
a pressure seat, respectively. Information regarding the level of difficulty and the state
of the game were obtained from the computer. The postures were labeled by teachers
who viewed the camera and the chair details before the experiment. Posture was the most
discriminating feature and had an accuracy of 81.97%. The accuracy for discrimination
between interesting and uninteresting stimuli increased to 86.55% after considering all of
the other features. Only 50 samples out of the 136 had features from all the modalities. The
remaining samples lacked details from the face channel. The number of samples classified
as “interested” was 65 and the number of samples considered “uninterested” was 71. Only
9 children were involved in this study. This highlights the need for studies with a larger
sample size.
To describe the use of physiological sensors in intelligent tutors to detect the student’s emotional states and provide emotional support

| Study | Purpose | Participants | Emotion | Technique | Stimuli | Measurement | Features | Classifier | Remarks |
|-------|---------|--------------|---------|-----------|---------|-------------|----------|------------|---------|
| Mota and Picard, 2003 [35] | To analyze postures to detect learner’s interest | 5 boys and 5 girls. Age range: 8 to 11 years | High interest, medium interest, low interest, taking a break, and boredom | -Two pressure sensors placed on the seat-pan of a chair and on the backrest | Solved a constraint satisfaction game called Fripples Place. Length: 20 min. | -Short interview | Mixture of Gaussian procedures. Gaussians parameters were prior probability, mean, and variance. | For static posture classification: 3-layer feed-forward neural network. For recognizing interest: HMM modeling | -Average accuracy of 82.25% for the 3 affect-related categories (high and low interest and taking a break). -For 2 new subjects, accuracy dropped to 76.5% because of the decline in recognition of “taking a break”. |
| Kapoor et al., 2005 [85] | To investigate the feasibility of combining different modalities to detect interest | 9 children. Age range: 8 to 11 years | Interest and disinterest | -A chair with two sensor sheets placed on the backrest and on the seat -Camera to record facial features -Game state information | Playing a game called Fripples Place, which had a number of puzzles. Length: 20 min. | -Several teachers annotated the data segments for the affective states. -FACS coders coded video faces. | -Eyebrow and eye shape, likelihood of nod, shake, and blink, probability of fidget, and smile. -Current posture and level of activity -Game state and the played level | After combining different features from face, posture, and game, GP classification was applied. | -Posture activity channel was the most discriminating, with 81.97% recognition accuracy. After application of the classifier combination method, the average accuracy for all features was 86.53% when using the multimodal Gaussian process approach. |
| M. Yeasin et al., 2006 [86] | To recognize the six universal facial expressions from visual data and use them to derive the level of interest using psychological classification of images | -Database of 488 video sequences for 97 subjects -Twenty-one subjects watched the movie. -108 sequences of television broadcast | -Sadness, happiness, fear, surprise, disgust, and anger. -Levels of interest | -Cohn-Kanade database [87] -Camera -Collected sequences of facial expressions from television broadcast | -Movie clips to arouse spontaneous natural emotions | Three persons unaware of stimuli content classified the sequences into the 6 emotions. | Selected components after applying PCA were used to form a feature vector with horizontal and vertical components of the optical flow. | Two-step classification technique: KNN followed by discrete HMMs | -Using the Cohn-Kanade database, the accuracy achieved was 90%, with significant failure in recognizing fear and disgust when compared to the rest of the 6 emotions. -In the 2nd experiment, accuracy decreased to 82%, the emotions fear and happiness were misclassified, and disgust was often recognized as sadness. -When using the television broadcast, the accuracy based on the level of interest classifier was 72%. |
| Arroyo, et al. 2009 [88] | To describe the use of physiological sensors in intelligent tutors to detect the student’s emotional states and provide emotional support | -1st study: 38 students. Age range: 15 to 17 years-2nd study: 29 females UG | Interest, frustration, excitement, and confidence | -Facial expression sensor, conductance bracelet, pressure mouse, and posture analysis seat | A multimedia adaptive tutoring system for geometry called “Wayang Outspot” during regular math class for 4-5 days | -Self-report | N/A | -Stepwise linear regression -Linear regression model | -Sensors may account for 60% of the variance in the student’s emotional states. -Moderate knowledge is essential for inducing interest in students. |

### Table 3. Detecting interest using other physiological sensors.

| Study | Purpose | Participants | Emotion | Technique | Stimuli | Measurement | Features | Classifier | Remarks |
|-------|---------|--------------|---------|-----------|---------|-------------|----------|------------|---------|
| Mota and Picard, 2003 [35] | To analyze postures to detect learner’s interest | 5 boys and 5 girls. Age range: 8 to 11 years | High interest, medium interest, low interest, taking a break, and boredom | -Two pressure sensors placed on the seat-pan of a chair and on the backrest | Solved a constraint satisfaction game called Fripples Place. Length: 20 min. | -Short interview | Mixture of Gaussian procedures. Gaussians parameters were prior probability, mean, and variance. | For static posture classification: 3-layer feed-forward neural network. For recognizing interest: HMM modeling | -Average accuracy of 82.25% for the 3 affect-related categories (high and low interest and taking a break). -For 2 new subjects, accuracy dropped to 76.5% because of the decline in recognition of “taking a break”. |
| Kapoor et al., 2005 [85] | To investigate the feasibility of combining different modalities to detect interest | 9 children. Age range: 8 to 11 years | Interest and disinterest | -A chair with two sensor sheets placed on the backrest and on the seat -Camera to record facial features -Game state information | Playing a game called Fripples Place, which had a number of puzzles. Length: 20 min. | -Several teachers annotated the data segments for the affective states. -FACS coders coded video faces. | -Eyebrow and eye shape, likelihood of nod, shake, and blink, probability of fidget, and smile. -Current posture and level of activity -Game state and the played level | After combining different features from face, posture, and game, GP classification was applied. | -Posture activity channel was the most discriminating, with 81.97% recognition accuracy. After application of the classifier combination method, the average accuracy for all features was 86.53% when using the multimodal Gaussian process approach. |
| M. Yeasin et al., 2006 [86] | To recognize the six universal facial expressions from visual data and use them to derive the level of interest using psychological classification of images | -Database of 488 video sequences for 97 subjects -Twenty-one subjects watched the movie. -108 sequences of television broadcast | -Sadness, happiness, fear, surprise, disgust, and anger. -Levels of interest | -Cohn-Kanade database [87] -Camera -Collected sequences of facial expressions from television broadcast | -Movie clips to arouse spontaneous natural emotions | Three persons unaware of stimuli content classified the sequences into the 6 emotions. | Selected components after applying PCA were used to form a feature vector with horizontal and vertical components of the optical flow. | Two-step classification technique: KNN followed by discrete HMMs | -Using the Cohn-Kanade database, the accuracy achieved was 90%, with significant failure in recognizing fear and disgust when compared to the rest of the 6 emotions. -In the 2nd experiment, accuracy decreased to 82%, the emotions fear and happiness were misclassified, and disgust was often recognized as sadness. -When using the television broadcast, the accuracy based on the level of interest classifier was 72%. |
| Arroyo, et al. 2009 [88] | To describe the use of physiological sensors in intelligent tutors to detect the student’s emotional states and provide emotional support | -1st study: 38 students. Age range: 15 to 17 years-2nd study: 29 females UG | Interest, frustration, excitement, and confidence | -Facial expression sensor, conductance bracelet, pressure mouse, and posture analysis seat | A multimedia adaptive tutoring system for geometry called “Wayang Outspot” during regular math class for 4-5 days | -Self-report | N/A | -Stepwise linear regression -Linear regression model | -Sensors may account for 60% of the variance in the student’s emotional states. -Moderate knowledge is essential for inducing interest in students. |
The study by Mota et al. [35] is the oldest of the studies discussed in this paper. This study is similar to the previous experiments outlined in Table 3 in terms of subject age and the sensors used, but differs in its purpose and classification algorithm. This study is also slightly different from the other studies in terms of features extraction. The purpose of this study was to measure the level of interest in learners while they played a constraint satisfaction game called Fripples Place over a period of 20 min. After acquiring the data using a camera and a pressure chair, three teachers separately labeled the data based on videos recorded by the cameras. These classifications had an average agreement of 78.57%, and this strategy was used because children are less likely to provide accurate statements about their feelings than adults are. After data pre-processing and noise removal, the data were modeled using a mixture of Gaussian processes. The Gaussian parameters were then fed into a 3-layer feed-forward neural network, which achieved an average static posture classification accuracy of 87.64%. Hidden Markov models (HMMs) were then used to model posture sequences for the affective states identified earlier by the teachers. These affective states were high interest, low interest, and taking a break. Medium interest was eliminated as an affective state because the model always confused this state with the high- or low-interest states, which resulted in poor performance. The final model had an average accuracy of 82.25% for classifying high interest, low interest, and taking a break. The study provided evidence that the observed patterns in the dynamics of the student's postures contained significant information related to the affective states of high interest, low interest, or taking a break. However, higher sample sizes are required in physiological studies to enhance model robustness and to obtain more validated results. The authors of this study also did not explain how they measured or quantified high vs. low interest or the theoretical or experimental basis for this measurement.

Arroyo et al. [88] studied the applicability of physiological sensors in detecting student emotions during intelligent tutoring classes. As shown in Table 3, the authors performed two experiments: one with school students and another with undergraduate students. They used self-reports with several physiological sensors, including a camera for facial expression recognition, pressure-sensitive seat cushions and back pads with audiometers, pressure mouse devices, and wireless conductance bracelets. These sensors were used to detect four emotional conditions: interested, confident, excited, and frustrated. The aim of the research was to find an alternative to the use of classical contextual variables such as time spent to solve a problem, number of hints given by the tutor, and number of attempts. The authors found a significant correlation between pre-test scores and interest. In particular, students with better mathematical knowledge reported higher interest than those with less extensive mathematical knowledge.

The authors showed that the feeling reported during each self-report question was related to what the subjects faced just before the question was presented. Therefore, when a subject self-reported, “I feel frustrated”, it was likely because they had one or more incomplete attempts in solving the previous problem. Similarly, when a subject reported, “I feel confident”, it was because he/she had just solved a problem. This was verified using stepwise regression, which indicated that interest (variance of 14%) could be predicted based on incorrect attempts and the gender of the learning companion. One limitation of this proposal was the unavailability of all the sensors used at the same time for all students due to several practical problems. The contribution from the camera sensor alone accounted for 52% of the variance in detecting confidence. The camera had similar contributions to the detection of other emotions. Using these sensors can help the tutor predict more than 60% of the variance in the student’s emotional state. The R values for the fit models used for the four emotions when using all the sensors and the tutor were 0.82, 0.72, and 0.70 for the confident, frustrated, and excited states, respectively. However, no fit model was found for the interested state. As mentioned earlier, the sensor data were only analyzed for 50% of the students due to difficulties in assessing all of the subjects using all the sensors simultaneously.
The study by Yeasin et al. [86] represents one of several studies that uses facial expressions to detect emotions and interest e.g., [88,89] (for a review, see [90]). However, only the Yeasin et al. study is discussed here because it used information from six emotions (surprise, sadness, fear, anger, happiness, and disgust) to measure the level of interest of subjects while they watch different movie clips. This study thus proposes a new indirect way of measuring the level of interest of an individual while he/she interacts with a machine. To measure the level of interest, the authors used a three-dimensional affect space (discussed in [91]). Each affective state was described using three parameters: arousal, valence, and stance (discussed in the level of interest section in this paper). A weight $W$, which was calculated by summing the contributions of each parameter, was assigned to each expression and then multiplied by the intensity ($I$) of that expression to compute the level of interest ($L$) using $L = W \times I$.

The proposed algorithm for facial expression recognition achieved an average accuracy of 90.9% using a Cohn–Kanade database. The algorithm had significantly higher accuracy in recognizing happiness, sadness, anger, and surprise when compared to fear and disgust. This study suggests that a two-step classification system is more efficient than a direct classification system, as the authors obtained better results when using both K-nearest neighbors and discrete HMM algorithms compared to continuous HMM. To validate their analysis of expressions and level of interest, the authors conducted an experiment with 21 subjects watching movie clips. The data were labeled with six emotions by three individuals who did not know the content of the videos. The recognition accuracy of the proposed algorithm was 82.1%.

Furthermore, in order to use real data instead of laboratory data, the authors collected 108 sequences from television broadcasts, including different facial expressions used worldwide. Three individuals labeled these data. The accuracy achieved was 72% due to the misclassification of surprise, fear, anger, and disgust, in addition to the unbalanced distribution of subject sex, age, and ethnicity between the test and training sets. The algorithm used had difficulty classifying the emotions correctly when the subject was talking. It also tended to misclassify expressions with similar patterns, such as fear and surprise, and sadness and disgust.

The above studies discuss the feasibility of using sensors that depend on the pose or the dynamic of the subject's posture to detect interest. There was a need to use more than one sensor at a time to improve classification accuracy. Compared to the EEG, ECG, and eye-tracking, these sensors had less accuracy and required more time and resources to handle and synchronize the data. However, it provided an insight into the effect of interest on physiological measures. Moreover, the subjective measures used especially by the academic staff had increased the reliability of the physiological measure. That could be a significant contribution to improve the educational environment by strengthening the student-educator relationship.

3.4. Correlation between Self-Report and Physiological Sensors

The correlation between self-report and physiological measures has been discussed widely from different perspectives (for a review, see [93–95]). In this section, this correlation will be addressed with regard to interest.

The reviewed studies showed great potential for detecting interest using physiological sensors. However, subjective measures, especially the self-report method, was employed in the majority of these studies as presented in the measurement column of Tables 1–3. The self-report method is accepted as the gold standard for determining one’s status and is often used with physiological measures to enhance and support their result. This is because of the long history and the solid basic research that used the self-report method and because of the lack of adequate physiological equipment at earlier times. The question would be to what extent can we trust the self-report method, especially if it conflicted with the result of physiological measures? On the one hand, there have been serious concerns raised about the credibility of the self-report method in certain conditions such as classroom
sessions or for certain groups such as children. On the other hand, concerns have been raised as well about the result of physiological measures that could be affected by the individual’s health condition or by faulty equipment. This implies that each method will have its pros and cons and that combining methods or modalities could be a good way to avoid such conflict in the first place. Recent studies employed different methods such as cameras [48], interviews [38], mouse behavior [31], etc. This was applicable for two reasons: (1) interest can be reported subjectively and / or objectively unlike pain, for example, which is considered a subjective phenomenon [96] and (2) the advances in technology that made it easy to incorporate two or more modalities.

Hence, it is recommended to combine subjective and objective measures when applicable to measure interest in all of its richness. Subjective measures report how one’s feel and think consciously about certain condition while the objective measures are necessary to tell about the unconscious mechanisms underlying this feeling or thought. This will offer complementary information about the association between interest, environmental condition, behavior, and brain and will have a valuable contribution to the understanding of the nature of interest. it is suggested to combine both measures until finding a credible and reliable physiological index that will overcome self-report limitations.

4. Effects of the Experimental Design on Experiment Output

Some of the physiological studies discussed here aimed to provide an index for the state or the process of interest. Other studies aimed to identify behavioral phenomena in response to the presence of a physiological event or response suggested being associated with interest. Before discussing the factors that affect the outcomes of physiological experiments, some issues should be highlighted for future consideration. Many of the experimental designs discussed in this review are limited by what is known as hypothetic-deductive logic. In other words, these studies relied on the identification of a physiological response differentiating the presence vs. absence of “interest”. These studies may be considered to have weak experimental design, as stated by Cacioppo and Tassinary [97]. This is because associating the presence of interest with certain physiological responses does not imply that the converse is true. Thus, there is a need for considerable effort to make accurate inferences regarding the presence or absence of certain physiological effects or responses based on a psychological phenomenon. Another consideration is the signal acquisition, cleaning, and analysis processes, which were intentionally discussed in this paper in order to give the reader an opportunity to visualize the effects of these processes on the outcomes of experiments. Cacioppo and L. G. Tassinary [97] stated that “the appearance of unreliable psychophysiological relations can also stem from imprecision during signal acquisition on the psychological side of the equation”. For example, the occurrence of one or more emotions at the same time as the targeted emotion will lead to inappropriate detection or identification of that targeted emotion. In addition, some analysis methods may mask important identification parameters or introduce falsified parameters.

4.1. Task and Stimulation

The stimulation is a major component of the experiment because if it fails to induce the respective emotion in the participant with the required intensity, the results and findings will be inaccurate or non-significant. As seen in the previous sections, the stimuli differed widely in length, type, and intensity depending on the experimental setting and the purpose of the study. To validate the stimuli, some authors used a pre-selected stimulus that was well-categorized [86]. Alternatively, the authors used pre-tests [34], wherein they used multiple comparisons analyzed using Bonferroni \(t\)-tests to show that clips A and B were rated significantly more interesting that clips C and D. These data were later tested to determine whether they were in agreement with the outcomes of the physiological sensors used in this study to differentiate between interesting and neutral or less-interesting clips.
Moreover, the studies included here did not use standard stimuli, which makes it difficult to perform valid comparisons between the methods used and the findings. The majority of the studies used paradigms to compare interesting vs. boring stimulations. These studies indicate clearly that interesting information is processed differently than boring information, although further investigation is required. This allows us to gain better insights into the phenomenon of interest itself rather than considering interest and lack of interest as two different states, whereby the presence of one state is necessary for the absence of the other one. It is difficult to separate the feelings of an individual when viewing certain stimulations, but one can control the stimulation.

Many of the studies that addressed learning content did not consider the variability of the learning content or the impact of this content on the subjects. It is obvious that the subjects had individual differences in learning strategies. Differences in the learning materials should not be overlooked when studying interest. To control for such variability, researchers may wish to control for the age of participants, examine intelligence quotients (IQs) or emotional quotients (EQs) of the subjects, assess the individual and situational interest of the participants, and investigate pre vs. post knowledge. Using these strategies, one can obtain a deeper understanding of the relationship or association between the “interestingness” of the stimulation and the physiological results obtained. Furthermore, the length of the stimulation was not addressed in these studies. A common stimulus length was 20 min. However, the shortest stimulus lasted for 15 s, and the longest stimulus lasted for several sessions over 4 to 5 days. Thus, the information provided here is not sufficient for the discussion of this factor.

4.2. Number and Age of Subjects

The ages of participants in the included studies ranged from 8 to 63 years. None of the studies reported an effect of age on the status or feeling of interest. However, this does not mean that such an effect is negligible. There is a need to examine the different conceptualizations of interest and whether they are affected by age. This may be performed by controlling for all experimental factors while allowing only changes in the participant’s age. Moreover, the sample size is an important factor in physiological experiments because physiological activity fluctuates and tends to vary from person to person, as well as within the same person following exposure to different environmental factors, such as room temperature, noise, or experimental setting in the presence of a stimulus.

All the experiments discussed above were carried out using normal subjects, with the exception of one study, wherein the subjects had difficulty decoding written text (dyslexia). This is likely because further investigation and validation is still required to identify and describe the nature of interest and the methods used to detect interest before differentiating its effects on normal vs. abnormal subjects. Nevertheless, including other subjects, such as those with autism or depression, would broaden this research area and lead to further insight into interest. This would benefit the above subjects and contribute positively to the development and burgeoning of “interest” research. Further tests are required to identify the status of subjects in terms of IQ and EQ in order to determine whether any correlations exist between these parameters.

5. Conclusions

This review is a modest attempt to assist in the evaluation of current studies related to interest detection using physiological sensors. To the author’s knowledge, all the studies concerning physiological measures of interest that contain the search keywords and imply the inclusion criteria were included. We believe that the impact of possible missing studies—if any—would be negligible due to the inclusion of the variety of articles presented here, which served the aim of this review. We found that interest has a wide range of definitions due to the diversity of research fields. Interest has different levels and types; therefore, it is beneficial in the fields of education, economics, learning of life skills, entertainment, and others. Detecting interest using physiological sensors is advantageous because it is an
efficient, real-time, accurate (hard to manipulate voluntarily), less distracting, and practical method when compared to classical self-report methods.

EEG methods used to measure physiological reflections of mental phenomena associated with interest from the frontal cortex of the brain have produced promising results. The frontal cortex is primarily responsible for attention and higher-order functions, including working memory, language, planning, judgement, and decision-making. However, interest may also be detected using information collected from parietal cortex. Using ERPs to extract the two types of P3, as well as the N140, is extensively discussed in this review. Eye-tracking is a well-known method for the detection of the participant’s interest in an object and is used in many applications. However, researchers have shown that this method cannot be used separately and requires support from other types of data. This is because environmental effects or large movements cannot be detected using eye-tracking algorithms. Using other sensors to provide assisting features, e.g., posture was found convenient and led to high accuracy comparable to that of EEG results. The use of multiple modalities was proven to be efficient and improved detection accuracy.

Interest is individually variable but is a powerful determinant of attention and perhaps recall [7]. The journey to find a reliable and accurate interest detection method is no doubt long but is highly promising and rewarding. Interest is considered by many recently proposed research perspectives as a motivation variable and an element for accomplishment and goal achievement. The most important factor to consider is good experimental practices, starting with the recruitment of eligible subjects, valid stimulations, and proper data acquisition, and ending with the proper choice of analysis method, feature extraction, as well as classification techniques. Another factor to consider is collaboration between researchers from different fields because little is known about the phenomenon of interest. For example, establishing a psychological hypothesis and designing the experiment with physiological signals would strongly enhance and verify the outcome of the proposed study and provide proper and comprehensive analysis for the phenomenon of interest from a different perspective. Future studies could also benefit from introducing a standard dataset to improve the integrity of the studies and enable replication if required. Furthermore, the number of subjects should be considered when validating a study. An appropriate large number of subjects is essential in physiological studies because of the fluctuating nature and variability in physiological parameters. Unfortunately, the self-report approach is still used to annotate and label the data obtained from the physiological sensors, which is then fed to the classifiers. However, it ought to use self-report measurements along with the physiological sensors to provide a perception of how the conscious mood is associated with cognitive functioning and to find a more credible source of validation than self-reporting. This may be accomplished by repeating the experiment using different subjects placed in the same scenario, i.e., the same physiological sensors and identical experimental settings. Alternatively, the same subjects can be tested using different valid, interesting stimulations. Behavioral measures performed by experts are also recommended.

Author Contributions: A.B. collected the data, analyzed them, and wrote the first draft of the paper. Y.B. reviewed and commented on the paper drafts, added new items, and assisted in improving the paper. A.A.A. and I.F. guided, reviewed and commented on this paper and helped in improving it. G.A. provided a discussion on effect of brain activities in different brain locations on EEG data and preprocessing methods. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The APC was funded by Universiti Tenaga Nasional (UNITEN) under BOLD2025 (RJO10517844/114) and the Ministry of Education Malaysia (MoE) under HICoE Scheme, awarded to CISIR, Universiti Teknologi PETRONAS (UTP).

Conflicts of Interest: The authors declare no conflict of interest.
60. Peterson, C.K.; Shackman, A.J.; Harmon-Jones, E. The role of asymmetrical frontal cortical activity in aggression. Psychophysiology 2008, 45, 86–92. [CrossRef]
61. Sutton, S.K.; Davidson, R.J. Prefrontal Brain Asymmetry: A Biological Substrate of the Behavioral Approach and Inhibition Systems. Psychol. Sci. 1997, 8, 204–210. [CrossRef]
62. Harmon-Jones, E. Clarifying the emotive functions of asymmetrical frontal cortical activity. In Psychophysiology; Willey Online Library: Hoboken, NJ, USA, 2003; Volume 40, pp. 838–848.
63. Davidson, R.J.; Pizzagalli, D.; Nitschke, J.B.; Kalin, N.H. Parsing the Subcomponents of Emotion and Disorders of Emotion: Perspectives from Affective Neuroscience. In Handbook of the Affective Sciences; Oxford University Press: Oxford, UK, 2003; pp. 8–24.
64. Davidson, R.J. What does the prefrontal cortex ‘do’ in affect: Perspectives on frontal EEG asymmetry research. Biol. Psychol. 2004, 67, 219–233. [CrossRef]
65. Harmon-Jones, A.E.; Lueck, L.; Fearn, M.; Harmon-Jones, C. The Effect of Personal Relevance Action and Approach-Related Expectation on Relative Left Frontal Cortical Activity. Psychol. Sci. 2006, 17, 434–440. [CrossRef]
66. De Pascalis, V.; Cozzuto, G.; Caprara, G.V.; Alessandri, G. Relations among EEG-alpha asymmetry, BIS/BAS, and dispositional optimism. Biol. Psychol. 2013, 94, 198–209. [CrossRef] [PubMed]
67. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. Brain Res. Rev. 1999, 29, 169–195. [CrossRef]
68. Kramer, A.F.; Weber, T. Application of psychophysiology to human factors. In Handbook of Psychophysiology; Cambridge University Press: Cambridge, UK, 2000; pp. 794–814.
69. Rosenfeld, J.P.; Bhat, K.; Miltenberger, A.; Johnson, E. Event-related potentials in the dual task paradigm: P300 discriminates engaging and non-engaging films when film-viewing is the primary task. Int. J. Psychophysiol. 1992, 12, 221–232. [CrossRef]
70. Polich, J. Updating P300: An integrative theory of P3a and P3b. Clin. Neurophysiol. 2007, 118, 2128–2148. [CrossRef] [PubMed]
71. Ercetin, G. Effects of topic interest and prior knowledge on text recall and annotation use in reading a hypermedia text in the L2. Recall 2010, 22, 228–246. [CrossRef]
72. Schiefele, U.; Krapp, A. The Effects of Topic Interest and Cognitive Characteristics on Different Indicators of Free Recall of Expositor. In Proceedings of the Annual Meeting of the American Educational Research Association, Chicago, IL, USA, 3–7 April 1991.
73. Schiefele, U. Content Related Motivation and Learning from Text. In Proceedings of the Annual Meeting of the American Educational Research Association, San Francisco, CA, USA, 20–24 April 1992.
74. Schneider, W.; Bjorklund, D.F. Interest and domain specific knowledge. In Proceedings of the Annual Meeting of the American Educational Research Association, San Francisco, CA, USA, 20–24 April 1992.
75. Kintsch, W. Learning from text, levels of comprehension, or: Why anyone would read a story anyway. Poetics 1980, 9, 87–98. [CrossRef]
76. Hidi, S.; McLaren, J. The effect of topic and theme interestingness on the production of school expositions. In Learning and Instruction in an International Con; Mandl, H., De Corte, E., Bennett, N., Friedrich, H.F., Eds.; Pergamon: Oxford, UK, 1990; pp. 295–308.
77. Tobias, S. Interest, Prior Knowledge, and Learning. Rev. Educ. Res. 1994, 64, 37. [CrossRef]
78. Kida, T.; Nishihira, Y.; Wasaka, T.; Nakata, H.; Sakamoto, M. Passive enhancement of the somatosensory P100 and N140 in an active attention task using deviant alone condition. Clin. Neurophysiol. 2004, 115, 871–879. [CrossRef]
79. Smith, M.E.; Gevins, A.; Brown, H.; Karnik, A.; Du, R. Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. Hum. Factors 2001, 43, 366–380. [CrossRef]
80. Morillo, L.M.S.; García, J.A.A.; Gonzalez-Abril, L.; Ramirez, J.A.O. Advertising liking recognition technique applied to neuromarketing by using low-cost EEG headset. Biomed. Eng. Online 2016, 5, 197–218.
81. Kort, B.; Reilly, R.; Picard, R.W. An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In Proceedings of the IEEE International Conference on Advanced Learning Technologies, ICALT 2001, Madison, WI, USA, 6–8 August 2001; pp. 43–46.
82. Asteriadis, S.; Tzovouli, P.; Karpouzis, K.; Kollias, S. Estimation of behavioral user state based on eye gaze and head pose-application in an e-learning environment. Multimed. Tools Appl. 2009, 41, 469–493. [CrossRef]
83. Sungkur, R.K.; Antoaaro, M.A.; Beeharry, A. Eye tracking system for enhanced learning experiences. Educ. Inf. Technol. 2016, 21, 1785–1806. [CrossRef]
84. Tzovouli, P.; Mylonas, P.; Kollias, S. An intelligent e-learning system based on learner profiling and learning resources adaptation. Comput. Educ. 2008, 51, 224–238. [CrossRef]
85. Kapoor, A.; Picard, R.W. Multimodal affect recognition in learning environments. In Proceedings of the 13th Annual ACM International Conference on Multimedia, Singapore, 6–11 November 2005; pp. 677–682.
86. Yeasin, M.; Bullot, B.; Sharma, R. Recognition of facial expressions and measurement of levels of interest from video. IEEE Trans. Multimed. 2006, 8, 500–507. [CrossRef]
87. Kanade, T.; Cohn, J.F. Comprehensive database for facial expression analysis. In Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France, 28–30 March 2000; pp. 46–53.
88. Arroyo, I.; Cooper, D.G.; Burleson, W.; Woolf, B.P.; Muldner, K.; Christopherson, R. Emotion sensors go to school. *Front. Artif. Intell. Appl.* **2009**, *200*, 17–24.

89. El Kaliouby, R.; Robinson, P. Real-time inference of complex mental states from facial expressions and head gestures. In *Real-Time Vision for Human-Computer Interaction*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 181–200.

90. Baltrušaitis, T.; McDuff, D.; Banda, N.; Mahmoud, M.; El Kaliouby, R.; Robinson, P.; Picard, R. Real-time inference of mental states from facial expressions and upper body gestures. In Proceedings of the 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011, Santa Barbara, CA, USA, 21–25 March 2011; pp. 909–914.

91. Pantic, M.; Rothkrantz, L.J.M. Toward an affect-sensitive multimodal human-computer interaction. *Proc. IEEE* **2003**, *91*, 1370–1390. [CrossRef]

92. Breazeal, C. Sociable Machines: Expressive Social Exchange between Humans and Robots. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2000.

93. Hadinejad, A.; Moyle, B.; Kralj, A.; Scott, N. Physiological and self-report methods to the measurement of emotion in tourism. *Tour. Recreat. Res.* **2019**, *44*, 466–478. [CrossRef]

94. Grassini, S.; Laumann, K. Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review. *Front. Psychol.* **2020**, *11*, 349. [CrossRef]

95. Spector, P.E. Using Self-Report Questionnaires in OB Research: A Comment on the Use of a Controversial Method. *J. Organ. Behav.* **1994**, *15*, 385–392. [CrossRef]

96. Twycross, A.; Voepel-Lewis, T.; Vincent, C.; Franck, L.; Baeyer, C. A Debate on the Proposition that Self-Report is the Gold Standard in Assessment of Pediatric Pain Intensity. *Clin. J. Pain* **2014**, *31*, 707–712. [CrossRef] [PubMed]

97. Cacioppo, J.T.; Tassinary, L.G. Inferring psychological significance from physiological signals. *Am. Psychol.* **1990**, *45*, 16–28. [CrossRef] [PubMed]