Chapter - 7

Learner Engagement Analysis through Facial Expression Recognition in Mulsemedia-Synchronized Learning Environments

7.1 Introduction

Multimedia is a combination of audio and video, enhanced by technology to improve understanding or memorization [234]. It supports verbal learning through the use of static and dynamic images, incorporating visualization technology for better expression and comprehension. The software and hardware used to create multimedia-based applications are collectively known as multimedia technology [235][236]. Multimedia designed for learning refers to the process of building mental representations from words and images in different contexts. Such multimedia is designed to aid learning with tools that can be used in presentations, classroom or laboratory learning, simulations, e-learning, computer games, and virtual reality, allowing learners to process information both verbally and visually [237]. Learning-focused multimedia also requires understanding certain theories, such as the cognitive theory of multimedia learning, which posits three assumptions about how people learn from instructional multimedia materials. However, Multimedia presentations that lack interactivity can lead to a passive learning experience, where users simply watch or listen without actively engaging. This can reduce information retention and understanding.

In recent years, interactive multimedia has advanced significantly, enabling learners to engage in dynamic and immersive learning experiences [238]. This progress is achieved by integrating various media elements, such as audio, text, video, images, and animation. Technologies like virtual reality and augmented reality have notably transformed the educational landscape, offering learners immersive and captivating experiences [239]. This learning environment provides numerous advantages, including flexibility, cost-effectiveness, and scalability, which contribute to higher retention rates and improved learning outcomes. However, while interactivity can enhance learning, it can also create distractions. Users may focus too much on navigating or interacting with the material, potentially missing the actual content and purpose. This can result in ineffective learning or "game-like" behavior, where engagement with the material diminishes.

Despite their appeal, most multimedia applications only engage two senses: sight and hearing, making them *bisensory.* This approach contrasts with the understanding that around

60% of human communication is nonverbal, and that our perception of the world typically relies on a blend of five senses—sight, hearing, touch, taste, and smell. Consequently, current multimedia experiences are limited in scope, unable to replicate sensations like warmth, humidity, or the rich scents one would encounter while strolling through a spice market in India. To address these limitations, *Multiple sensorial media* (or *Multismedia*) has been introduced by Prof. George [6].

The remainder of this chapter is organized as follows: Section 7.2 focuses on the application of *mulsemedia* in learning environments and the methods used by researchers to measure its effectiveness. Section 7.3 introduces an innovative approach for designing an IoT-based, mulsemedia-synchronized web learning portal, detailing its various mulsemedia components. Section 7.4 covers the experimental findings related to subjective QoE measures in mulsemedia. Section 7.5 presents the results of learner engagement analysis through FER on mulsemedia-synchronized content, comparing outcomes across two distinct groups. Finally, Section 7.6 summarizes the chapter, highlighting the main findings and their implications.

7.2 Mulsemedia in Learning

Mulsemedia defines multimedia experiences that stimulate multiple human senses—beyond just sight and hearing, which are typical for traditional multimedia. Mulsemedia integrates additional sensory inputs, such as touch (haptics), smell (olfaction), and sometimes even taste, to create a more immersive and realistic experience [240]. By engaging multiple senses, mulsemedia aims to enhance user engagement, retention, and emotional response, making it particularly valuable for applications in education, entertainment, training, and virtual reality [241].

In recent years, numerous studies have highlighted the integration of various technologies in education to enhance student motivation, engagement, and academic achievement, particularly in science, technology, engineering, and mathematics fields [243]. However, many learners encounter difficulties in fully engaging with and comprehending the learning material. By incorporating a blend of audio and other sensory stimuli—such as olfactory, haptic, and gustatory inputs—interaction with digital content can be significantly enhanced. There is an increasing emphasis on integrating multisensory effects, which include visual and auditory information, haptic feedback, as well as olfactory and taste sensations, into the learning process. This method aims to utilize multiple senses to create immersive and enriched experiences for learners within educational environments. Figure 7.1 illustrates the components of mulsemedia.

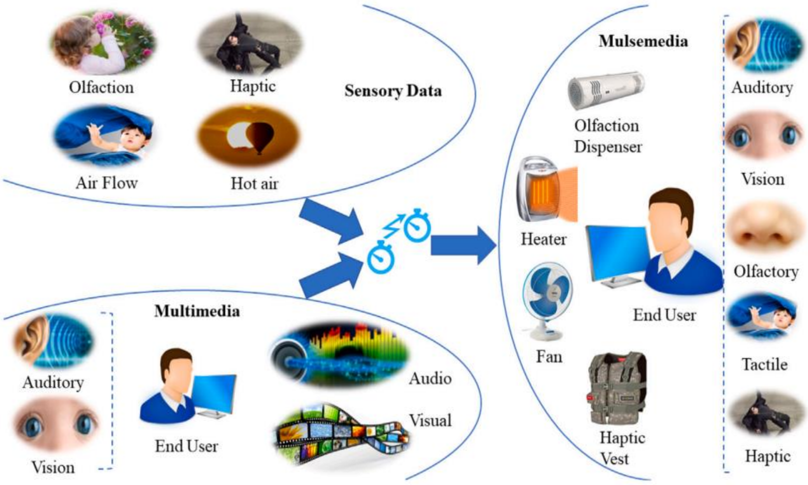


Figure 7.1 Mulsemedia Components, synchronization, and delivery to end users [242]

Current innovative technologies and approaches, such as multisensory augmented and virtual reality, along with game-based learning, provide an ideal environment for experimentation—an essential aspect of engineering education. Developing multisensory course content can accommodate diverse learning abilities and promote inclusivity in educational settings. Additionally, this approach can foster an embodied and enactive learning experience, where knowledge is not just conveyed through the senses but is gained through active participation. The rationale behind this perspective is that the human brain is inherently multisensory [244], and the processing of multimodal information enables effective interaction with the world. Therefore, it is argued that learning can be enhanced and enriched by engaging multiple sensory pathways.

Several significant studies have examined the role of mulsemedia in educational settings. For instance, Tal et al. [245] argue that incorporating multimedia elements—such as text, audio, and visual content—enhances student engagement and enriches the overall learning experience. They highlight the effectiveness of mulsemedia in promoting a deeper understanding of complex concepts, encouraging collaboration among students, and better preparing them for real-world challenges in the telecommunications sector. Similarly, other studies [246][11] emphasize that integrating various media formats can boost student motivation, facilitate deeper knowledge retention, and improve overall engagement. These findings indicate that incorporating mulsemedia into conventional learning environments can provide significant advantages by creating a more immersive and interactive educational experience. However, existing research has not thoroughly analyzed learner engagement in mulsemedia-synchronized learning environments.

7.2.1 Tools for Measuring Mulsemedia Effects in Learning Environments

In the context of learning environments, QoE in mulsemedia refers to the subjective satisfaction and overall effectiveness experienced by learners when engaging with multimedia content that incorporates multiple sensory elements—such as visuals, sound, touch, and even smell or taste. QoE in mulsemedia is critical as it influences learners' engagement, motivation, and retention of knowledge, making it a vital factor in assessing and optimizing educational technologies. According to [247], QoE can be defined as "*the degree of delight or annoyance of a person whose experience involves an application, service, or system. It results from the person's evaluation of the fulfillment of his or her expectations and needs concerning the utility and enjoyment in light of the person's context, personality, and current state.*" As has been reported, user QoE can be influenced by several factors, including technical, social, and psychological [248].

Figure 7.2 shows the type of QoE assessment includes subjective and objective metrics. Subjective methods involve learner feedback, surveys, and engagement ratings, while objective methods might include tracking user interactions, completion rates, and physiological responses (e.g., heart rate, skin conductance). Data from these metrics help refine mulsemedia content to better suit learner needs and maximize engagement and effectiveness.

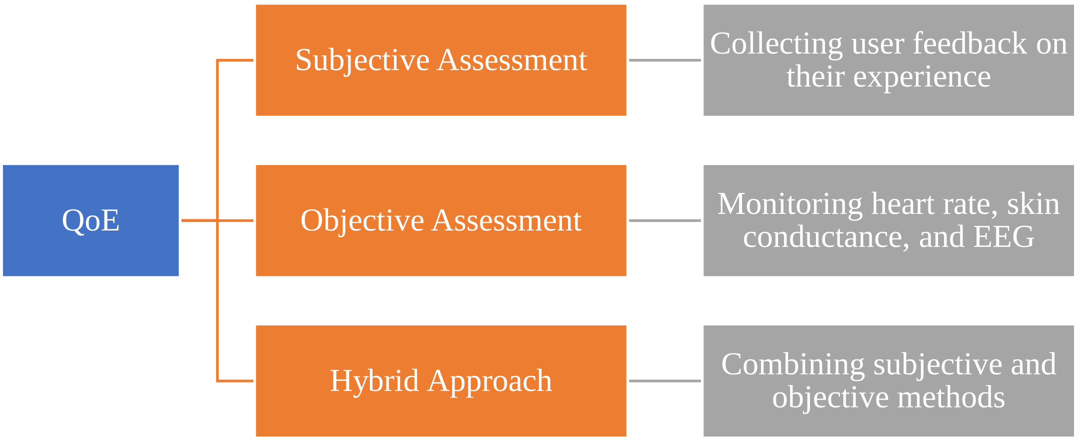


Figure 7.2 Methods to Measure QoE in Mulsemedia

The literature suggests that the accepted approach to measuring a user's perceived quality of experience has been based on self-reported measures via post-experience questionnaires. Such questionnaires are used to determine an overall Mean Opinion Score (MOS) based on feedback from users. However, several researchers have highlighted issues with MOS rating scales and post-test experience reporting; they are considered time-consuming and expensive to implement [249]. For these reasons, the research community has started to investigate methods that capture the user's QoE as the user experiences the content. Such methods have included analyzing facial expressions and physiological signals.

In this research study, we have employed both subjective and objective assessments to evaluate learner QoE in a learning environment utilizing mulsemedia. The subjective assessment involved using post-feedback questionnaires to gauge the learners' QoE. The objective assessment involved FER analysis to monitor the learners' emotions during the learning process.

7.3 Materials and Methods

This section examines the impact of employing mulsemedia—incorporating olfactory, airflow, and vibrotactile effects—on content within a mulsemedia learning environment, as well as its effect on learner QoE. Figure 7.3 shows the proposed mulsemedia-synchronised learning web portal.

7.3.1 Mulsemedia-Synchronised Learning Web Portal

Figure 7.3 Proposed Mulsemedia-Synchronised Learning Web Portal

A. Video Materials

There are two types of learning content − specifically two videos, one targeting various biological aspects of the rosemary plant such as origin, appearance, smell, and their medical application and benefits, the other focusing on the physics behind the phenomena of thunder and lightning. The choice of the subject matter of the two videos was done with a view towards the rosemary video being enhanced with (rosemary) aromas for users in the experimental group, whilst the thunder and lightning video would be accompanied by airflow and vibrotactile effects, again for users in the experimental group. Both videos lasted 5 minutes and 28 seconds and were encoded and rendered at 30 frames per second as shown in Figure 7.4.

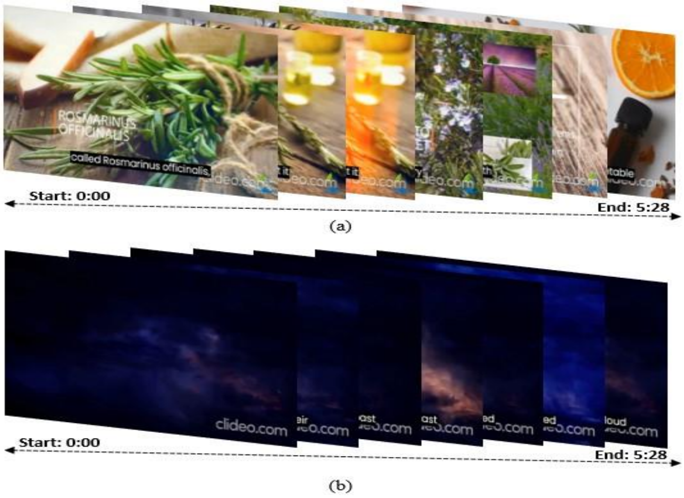


Figure 7.4 Video Material (a), Rosemary Plant (b) Thunder and Lightning

Figure 7.5 (a) Ultrasonic Humidifier DC 24 V 36 mm, (b) Arduino-based 5v DC fan, (c) Gaming Haptic Mouse: Rival 700

B. **Experimental Setup**

The two videos were telecast on a Dell desktop boasting a resolution of 1600 x 900 pixels, and a diagonal of 20 inches. JBL Quantum 100 wired over-ear gaming headphones were used to deliver noise-free high-quality audio for the learners. Sensory effects beyond the audio-visual are respectively delivered by a set of affordable devices, which shall now be described.

Accordingly, the aroma dispenser utilized in this study was an Ultrasonic Humidifier shown in Figure 7.5 (a). This is a compact and high-performance device that utilizes ultrasonic cool mist technology. It functions by vibrating a diaphragm to generate a mist of water droplets, thereby reintroducing moisture into the air. The device, constructed from plastic and zinc alloy, features a single atomizing head with a 36mm atomizing sheet. It boasts a production capacity of 440ml of fog per hour and operates at a working voltage of DC 24V, with a power dissipation ranging from 800mA to 1.5A. Its operational frequency is 1.7MHz.

Airflow effects were emitted by a DC 5V fan, based on Arduino, as shown in Figure 7.5 (b). This is a compact device powered by a brushless motor. It functions by receiving power from the Arduino board and can be controlled using an NPN transistor. The fan operates at a rated voltage of DC 5V and an operating voltage of 4.5-5V. It consumes a current of 0.16A and has an air volume of 5 m3/h. The fan operates at a speed of 13200rpm and produces a noise level of 18dB. It utilizes double micro-ultra-fine bearings for smooth operation.

For vibrotactile effects, our study employed the Rival 700 Gaming Haptic Mouse, depicted in Figure 7.5 (c). This is a high-performance device equipped with an optical sensor. The device features 7 buttons, a scroll wheel, and a wired connection. It weighs 159 g and has a height of 124.85 mm. It supports different grip styles including palm, claw, and fingertip. Last but by no means least, a purposefully designed IoT-based Mulsemedia Enhanced web portal (Figure 7.3) was implemented for the study using a combination of Hypertext Markup Language is the standard markup language (HTML), Cascading Style Sheets (CSS), and Hypertext Preprocessor (PHP). Arduino programming was used to control the timings and release (respectively switch off) of mulsemedia effects for experimental group learners. This portal integrated three key pieces of hardware, as depicted in Figures 7.3 and 7.5, each of which respectively delivered different sensory effects.

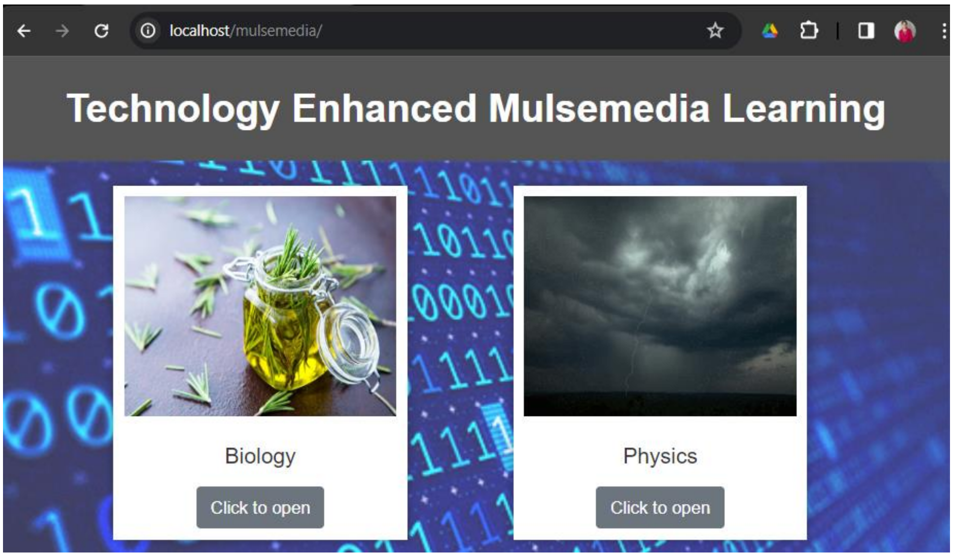


Figure 7.6 Home Page of Mulsemedia Synchronized Web Portal

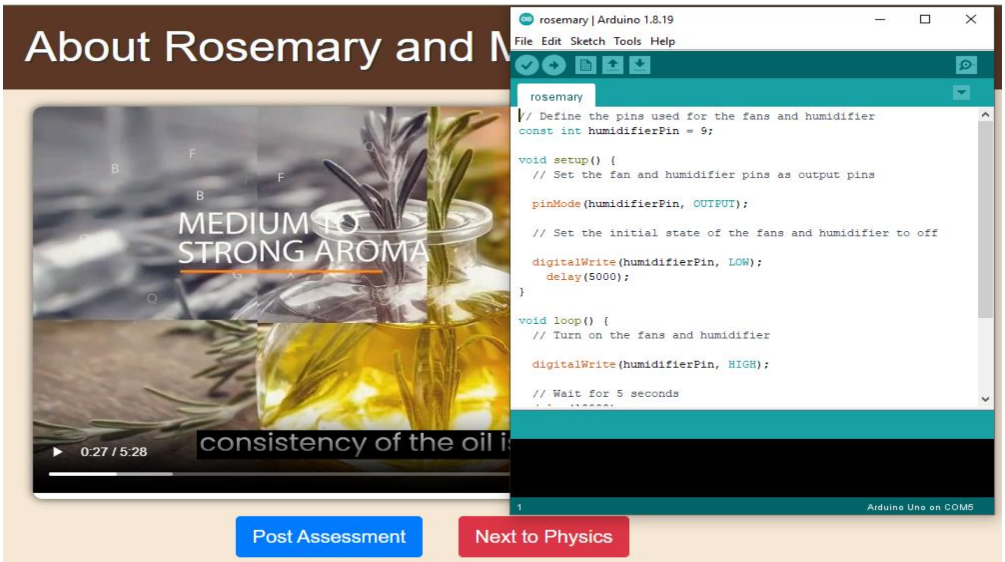


Figure 7.7 Mulsemedia effects emitted by programmed Arduino IDE

In the designed web portal, sensory effects delivered by the above devices have been programmed to synchronize with audiovisual content with the help of an Arduino time-based controller. Accordingly, 5s before the start of the botany (rosemary) video, the humidifier was activated to emit rosemary scent aromas which lasted the length of the video every 5s (aromas were emitted before the video started playing to give time for the olfactory effects to dissipate and reach the user, which was seated 50cm away [250]). Similarly, airflow and vibrotactile begin 5s into the video and are synchronized with the video content when thunder and lightning appear (Figure 7.4). The synchronization effects were tested with five users before initiating experiments with actual participants. Based on their feedback, multisensory media effects were configured to emit before the video started, aiming to enhance sensory engagement.

7.4 Experimental Results and Discussion

7.4.1 Participants

The research study was conducted at the Centre for Machine Learning and Intelligence Laboratory, Avinashilingam University, India, with a group of 70 (35 control group, 35 experimental) participants, all aged between 21 and 30 years. About 40% of these participants were undergraduates in the age group of 21-23, another 40% were postgraduates aged between 23-25, and the remaining 20% were researchers aged between 25-30, pursuing science degrees. This study was conducted after receiving ethical approval from the Ethical Committee of Avinashilingam University to include students, Approval number: AUW/IHEC/CS-22-23/XPD-07.

Each participant was randomly assigned to one of two equal-sized experimental and control groups. Experimental group (EG) learners were exposed to learning content enhanced with mulsemedia (incorporating olfactory, airflow, or vibrotactile effects). In contrast, control group (CG) learners experienced the same content without mulsemedia effects.

7.4.2 Experimental Protocol

All participants were briefly explained the purpose of the experiments and requested to sign the consent form, allowing the use of their feedback for analysis. Initially, content and web portals were designed according to the target participants.

The experimental protocol is outlined in Figure 7.8. Participants were randomly divided into two equal groups: experimental and control groups. The experimental group experienced mulsemedia, while the control group did not. Before watching the selected video, participants took a pre-assessment test. The designed questionnaire was stored in Google Forms and integrated into the web portal before video playback. Following the video, users completed the post-assessment questionnaire (Appendix I). The same process was repeated for the thunder and lightning content as well. Figure 7.7 shows a screenshot of the Arduino IDE trigger for rosemary olfactory effects for the experimental group. After viewing both videos, participants were asked to evaluate their QoE questionnaire. Overall, the process took roughly 18 minutes to complete for control group participants and 20 minutes for learners in the experimental group, as these had to answer more questions concerning the QoE, as described below.

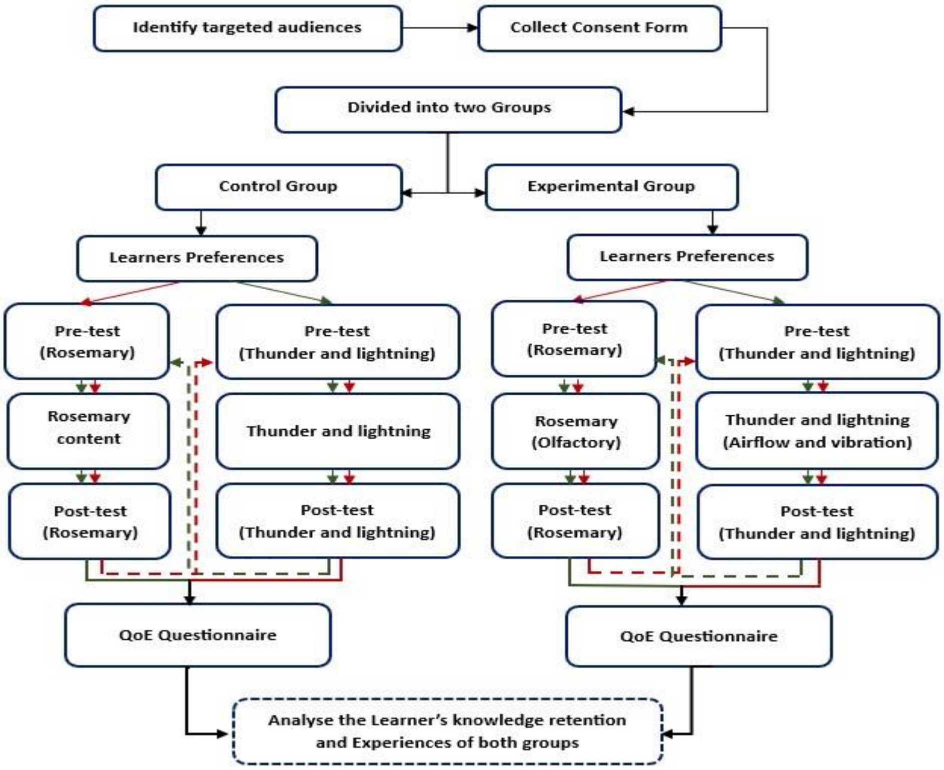


Figure 7.8 Experimental Protocol

7.4.3 Research Instruments

A. **QoE Questionnaire:** This consists of 22 questions designed to assess various aspects of the participant's experience during the mulsemedia-synchronised learning process (See Table 7.1). The first 12 questions were only answered by experimental group participants, as they targeted the QoE in a mulsemedia context. Accordingly, the first four questions (Q1-Q4) focus on the impact of olfactory effects on the learning experience, including their enhancement of reality, the potential for distraction, and overall involvement. Q5 to Q8 explore the influence of airflow effects on the sense of reality, distraction, annoyance, and enjoyment in learning content. Q9 to Q12 delve into the effects of vibration on the learning experience, including its enhancement of reality, potential for distraction, annoyance, and enjoyment. The remaining questions (Q13-Q22) evaluate the overall learning approach. They assess the ease of use, potential cumbersomeness, understanding of concepts, knowledge acquisition, satisfaction of learning needs, improvement of the learning experience, practical engagement in the learning process, enjoyment, recommendation to friends, and desire to continue learning with the experienced approach. Participants answered each QoE question on a 5-point Likert Scale ranging from Strongly Disagree to Strongly Agree, which for analysis purposes were coded with numerical values ranging from 1 to 5 [251]. The feedback gathered from these questions is used to shed light on the efficacy of the multimedia-enhanced learning system and its influence on learners' QoE.

Table 7.1 QoE Questionnaire

No.	Questions
Q1	The olfactory effects enhance the sense of reality in e-learning.
Q2	The olfactory effects are distracting.
Q3	The olfactory effects are annoying
Q4	I enjoy this learning with olfactory effects.
Q5	The airflow effects enhance the sense of reality in e-learning
Q6	The airflow effects are distracting
Q7	The airflow effects are annoying
Q8	I enjoy this learning with airflow effects
Q9	The vibration effects enhance the sense of reality in e-learning.
Q10	The vibration effects are distracting
Q11	The vibration effects are annoying
Q12	I enjoy this learning with vibration effects
Q13	The experienced approach was easy to use
Q14	The experienced approach was cumbersome
Q15	The experienced approach helped me better understand the concepts explained
Q16	The experienced approach did not help me learn knowledge
Q17	The experienced approach can satisfy my learning needs
Q18	The experienced approach did not improve my learning experience
Q19	The experienced approach did not help me be more practically engaged in the learning process.
Q20	I enjoyed the experienced approach.
Q21	I would not recommend the experienced approach to my friends.
Q22	I would like to learn more with the experienced approach

B. **Learner Questionnaires**: Learners answered a set of questionnaires designed based on the video content experienced and different levels of Bloom's taxonomy of learning [252]. This classification system comprises six levels used to study the learner's behavior and infer cognitive memory achievement (i.e., remembering, understanding, applying, analyzing, evaluating, and creating). In this study, the questionnaire was designed to encompass three different levels of Bloom's taxonomy [253]. It has been noted as a potential tool for student's self-assessment of learning experiences by Athanassiou et al. [254].

In this study, the questionnaire was designed to encompass three levels of Bloom's taxonomy [253]. For this analysis, pre- and post-questionnaires were designed to assess learners' memory retention before and after a learning experience to gauge the effectiveness of video instruction. Accordingly, the rosemary questionnaire focuses on the rosemary plant, information relevant to which is presented in the corresponding video (Ros). The first eight questions were created to assess memory (remembering), while the remaining two questions were designed to evaluate the judgment of information acquired (evaluating). Questions range from basic identification, such as the scientific name of rosemary and the family it belongs to, to more detailed inquiries about its growth, medicinal benefits, and the origins of its name. We also delved into the properties of rosemary oil; a product derived from the plant that is often used in aromatherapy. The physics questionnaire shifts to the atmospheric phenomena of thunderstorms and lightning as contained in the corresponding video clip (TL). The breakdown of questions is identical to the case of the biology questionnaire, with eight questions targeting remembering and a further two focused on and another two questions were created focused on understanding, and applying content related to the respective video. Here, questions covered matters such as the temperatures reached by the air around a lightning strike, the distance from which the sound of thunder can be heard, and the factors that affect the speed of thunder. For both questionnaires, learners were marked on a 1 to 100 scale, with the mark reflecting the percentage of correct answers given.

7.4.4 Results Analysis

Statistical analysis was performed using the IBM Statistical Package for the Social Sciences (SPSS) for Windows version 29.0. Paired sample t-tests were utilized to compare the means of two variables within the same group, and an independent sample t-test was employed to identify statistically significant differences between two distinct groups of QoE questionnaires. A significance level of $p < 0.05$ was adopted for this study. Typically, subjective test results are assessed about internal reliability, indicating the degree to which participants maintain consistency in their ratings across various questions (See Table 7.3). This is commonly measured using metrics such as Cronbach's alpha and McDonald's omega. The reliability of mulsemedia effects and the QoE questionnaire within the experimental group has been validated. The Cronbach's alpha value is 0.900, indicating a high level of internal consistency for the scale. The McDonald's omega value is 0.915, further confirming the reliability of the scale. (Table 7.2 shows the standard Cronbach's Alpha reliability level).

Table 7.2 Cronbach's Alpha Reliability Level [254]

Cronbach's Alpha Value	Level of Reliability
0.0 >0.20	Less Reliable
>0.20 – 0.40	Somewhat Reliable
>0.40 – 0.60	Reliable enough
>0.60 – 0.80	Reliable
>0.80 – 1.00	Very Reliable

Table 7.3 Reliability Analysis Score of Proposed QoE Questionnaire

	Mean	SD	Cronbach's α
scale	4.16	0.493	0.900

7.4.4.1 Analysis of Learner Performance between Groups

Comparative analysis of learning performance between the Control Group (CG) and the Experimental Group (EG) reveals some interesting insights. Accordingly, as shown in Table 7.4 for both CG and EG, there was an increase in the mean learning performance from the pre-test to the post-test, irrespective of the visualised content (Ros or TL). Moreover, a paired samples t-test revealed that, whilst there were no statistically significant differences in pre-test scores between CG and EG participants for either biology or physics content (Pair 1 and Pair 3), the learning performance for EG participants in post-tests was higher – and statistically significant − than that of their CG counterparts, for both types of subject matter (Pair 2 and Pair 4). Thus, in our study, exposure to mulsemedia has resulted in statistically significant higher learning performance. In addition, Figures 7.9 and 7.10 show the participants involved in this study, both with and without mulsemedia.

Table 7.4 Paired Sample T-Test

		Mean	t	df	Sig. (2-tailed)
Pair 1	CG_Ros_Pre - EG_Ros_Pre	3.429	1.046	34	.303
Pair 2	CG_Ros_post - EG_Ros_Pos	-20.286	-7.875	34	.000
Pair 3	CG_TL_Pre -EG_TL_Pre	2.857	.935	34	.356
Pair 4	CG_TL_Pos - EG_TL_Pos	-25.143	-7.209	34	.000

7.4.4.2 Analyzing the Impact of Multimedia on Information Retention

The impact of employing mulsemedia (including olfactory, airflow, and vibrotactile effects) on learner QoE in this study has been assessed, and Table 7.5 highlights that learners displayed highly positive opinions with respect to the use of mulsemedia in the two-science content. Moreover, these opinions are all statistically significant. Specifically, learners perceived a heightened sense of reality when olfactory effects were integrated into the learning environment but did not find olfactory effects to be either distracting or annoying. Importantly, participants reported elevated enjoyment levels when olfactory effects were integrated into the learning. Encouragingly, the same pattern of highly positive and statistically significant responses was observed with respect of the use of airflow and vibrotactile effects. These findings underscore the subjective nature of sensory effects in learning. The majority of participants experience an enhanced sense of reality and enjoyment with the incorporation of olfactory, airflow, and vibrotactile effects, without perceiving any of them as distracting or annoying.

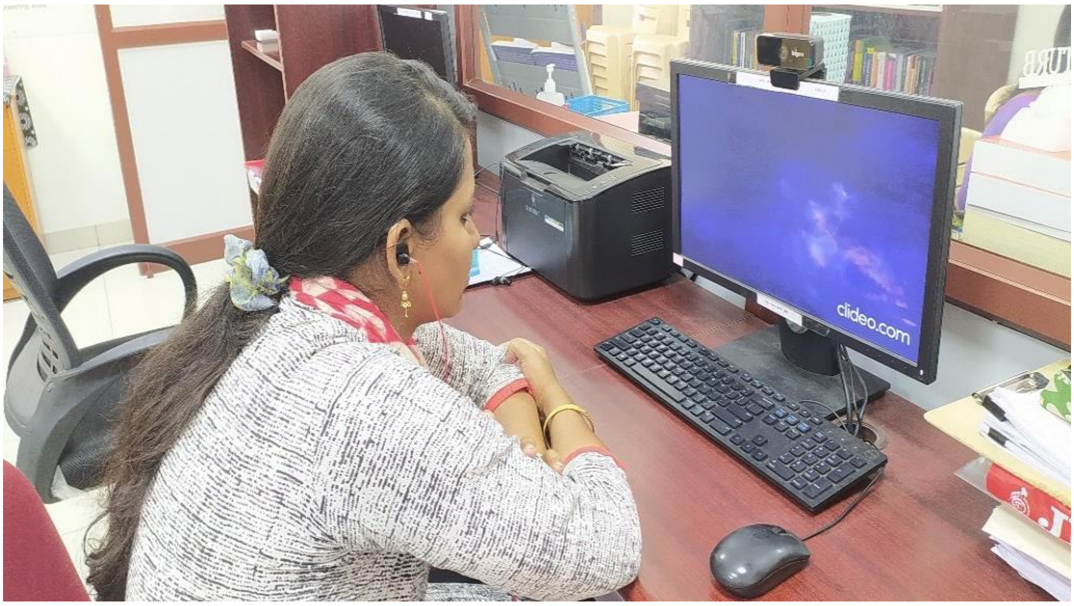


Figure 7.9 Participants without Mulsemedia

Figure 7.10 Participants with Mulsemedia effects

Table 7.5 One-Sample T-Test

	Test Value = 3							
	N	t	df	Significance		Standard Error Mean	95% Confidence Interval of the Difference	
				Two-tailed P	Mean Difference		Lower	Upper
Q1	35	18.392	34	.000	1.743	.095	1.55	1.94
Q2	35	-18.181	34	.000	-1.657	.091	-1.84	-1.47
Q3	35	-22.127	34	.000	-1.714	.077	-1.87	-1.56
Q4	35	30.946	34	.000	1.857	.060	1.74	1.98
Q5	35	15.197	34	.000	1.714	.113	1.49	1.94
Q6	35	-8.341	34	.000	-1.314	.158	-1.63	-.99
Q7	35	-9.731	34	.000	-1.457	.150	-1.76	-1.15
Q8	35	21.377	34	.000	1.771	.083	1.60	1.94
Q9	35	10.204	34	.000	1.400	.137	1.12	1.68
Q10	35	-7.948	34	.000	-1.286	.162	-1.61	-.96
Q11	35	-9.942	34	.000	-1.429	.144	-1.72	-1.14
Q12	35	12.579	34	.000	1.571	.125	1.32	1.83

7.4.4.3 Analysis of Self-Reported Mulsemedia Quality of Experience

Table 7.6 Comparison of Control Group and Experimental Group

Items	Group	Mean (Std. Dev.)	P-value (t-value)
Q13	CG	3.94 (.482)	.000
	EG	4.86 (.355)	(-9.040)
Q14	CG	3.20 (.797)	.000
	EG	1.51 (.612)	(9.923)
Q15	CG	3.91 (.507)	.000
	EG	4.86 (.355)	(-9.011)
Q16	CG	2.17 (.707)	.000
	EG	1.29 (.622)	(5.568)
Q17	CG	3.71 (.710)	.000
	EG	4.86 (.355)	(-8.517)
Q18	CG	2.40 (.812)	.000
	EG	1.23 (.426)	(7.560)
Q19	CG	3.37 (.910)	.000
	EG	1.17 (.382)	(13.183)
Q20	CG	3.71 (.622)	.000
	EG	4.91 (.284)	(-10.386)
Q21	CG	2.57 (.850)	.000
	EG	1.17 (.382)	(8.885)
Q22	CG	3.66 (.684)	.000
	EG	4.89 (.323)	(-9.615)

The comparison between the CG and the EG indicates substantial differences in how participants perceived and experienced the educational approach (See Table 7.6). The EG consistently rated the approach more positively across various measures. Participants in the EG found the approach notably easier to use (4.86) compared to the CG (3.94), indicating a higher level of user-friendliness. Additionally, the EG perceived the approach as significantly less cumbersome (1.51) in contrast to the CG (3.20), suggesting a stark contrast in perceived difficulty levels. In terms of understanding concepts and learning outcomes, the EG gave higher ratings (4.86 and 1.29, respectively) compared to the CG (3.91 and 2.17, respectively). This implies that the approach might have been more effective in aiding understanding and improving learning outcomes for participants in the EG. Moreover, the EG consistently reported higher satisfaction, better engagement, enjoyment, and a greater likelihood to recommend the approach to friends. This suggests that the intervention positively impacted various facets of the learning experience for the EG compared to the CG. Notably, the EG expressed stronger interest in further learning (4.89) compared to the CG (3.66), indicating that the approach had a more significant motivational influence on continued learning for participants in the Experimental Group. In summary, the EG consistently rated the experienced approach more positively across various dimensions compared to the CG. These findings suggest that the intervention or the method applied to the EG had a more favorable impact on user experience, perceived learning outcomes, engagement, and motivation compared to the CG.

Figure 7.11 presents the QoE results for both group experiences on a mean scale ranging from 1 to 5. The findings indicate that mulsemedia-synchronized learning significantly enhanced the learning experience compared to the conventional multimedia approach. Furthermore, participants recommended the adoption of mulsemedia-based learning for future educational experiences.

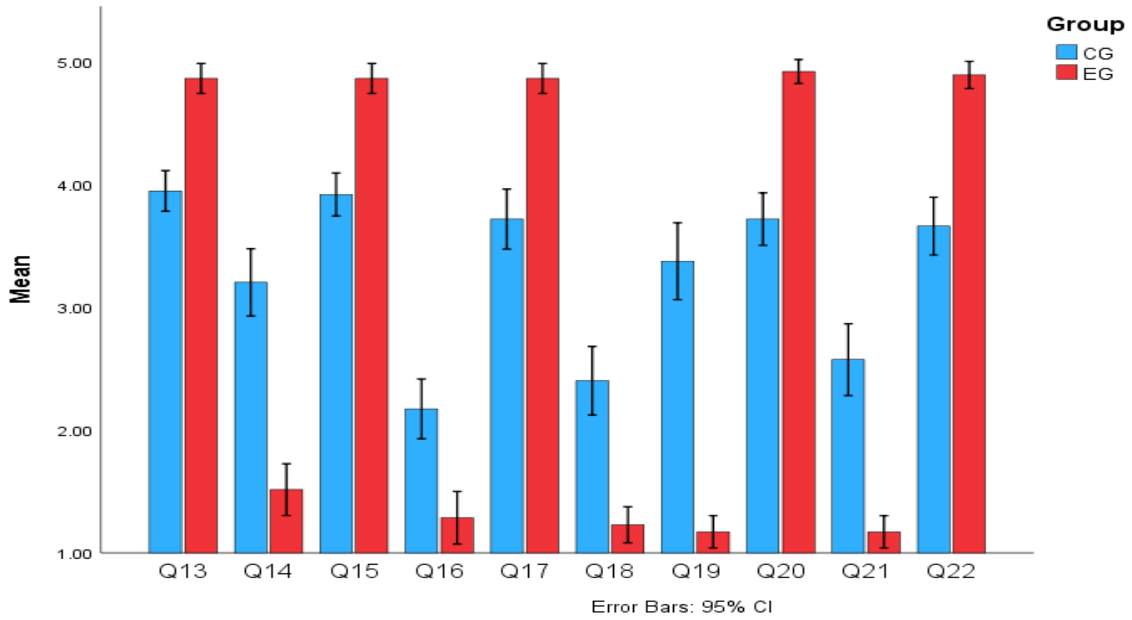


Figure 7.11 QoE Analysis of Both Groups

7.5 Learner Engagement Analysis in Mulsemedia-synchronized Learning through FER

7.5.1 Participants

For this assessment, 20 participants were included, divided equally between two groups: an experimental group and a control group. Each group consisted of 10 participants. All participants were postgraduate and PhD students with science backgrounds. In addition, ethical clearance was obtained from the Institutional Human Ethics Committee of Avinashilingam University for mulsemedia in affective computing (Approval number: AUW/IHEC/CS-22-23/XPD-07).

7.5.2 Experimental Procedure for Evaluating Mulsemedia

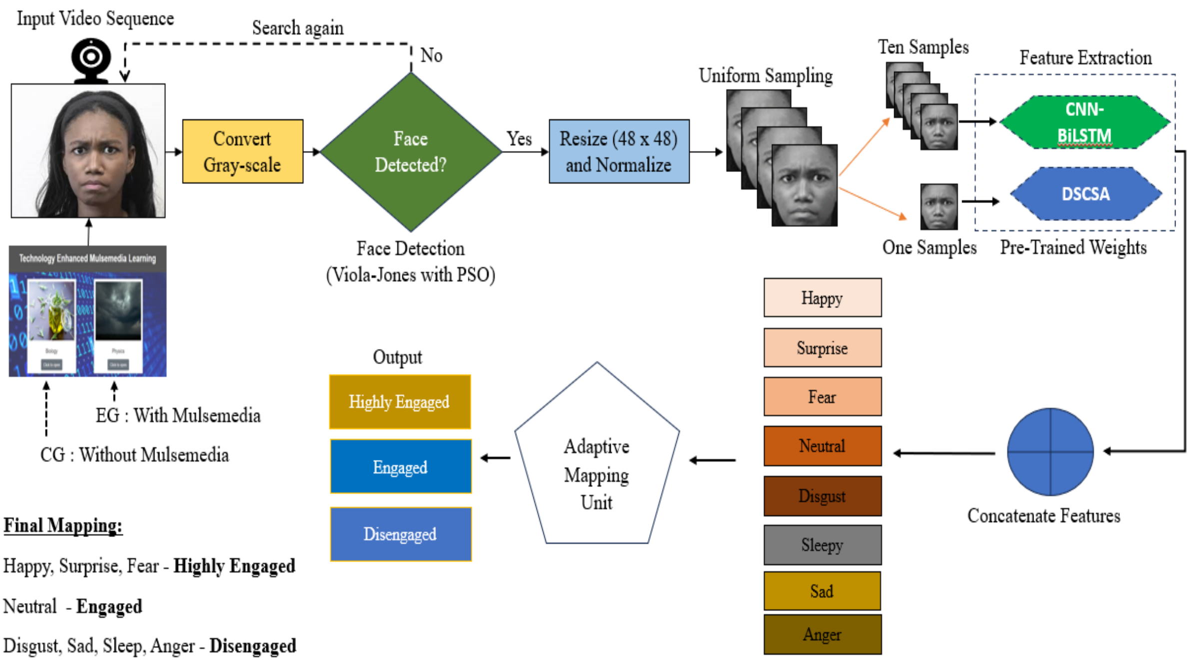


Figure 7.12 Overall Architecture of Learner Engagement Analysis through FER

Consent forms were initially provided to each participant prior to their involvement in the study. As previously mentioned, the study comprised two groups. The experimental group received mulsemedia-synchronized learning content, and participants' facial expressions were recorded using a laptop's web camera. Simultaneously, predicted engagement results for these learners were recorded individually and stored in CSV files for the overall final analysis. In contrast, the control group followed the same procedure but without the mulsemedia effects. A detailed description of the overall architecture for learners' engagement analysis through FER is provided in Chapter 3.

7.4.3 Learner Engagement Detection from Facial Expression

In this section, the proposed FER system is used to assess learner engagement levels during mulsemedia-synchronised learning.

Table 7.7 Leaners Engagement Classification

Emotion	Engagement
Happy (EP ≥ 50)	Highly Engaged
Surprise (EP≥ 80)	Highly Engaged
Neutral (EP ≥ 80)	Highly Engaged
Happy (EP < 50)	Engaged
Surprise (EP > 60 and EP < 80)	Engaged
Neutral (EP > 60 and EP < 80)	Engaged
Neutral (EP > 60) + Angry (EP <20) + Sad (EP > 30)	Engaged
Angry (EP > 20)	Disengaged
Sad, Sleep	Disengaged
Fear (EP > 30)	Disengaged
Angry (EP > 20) + Sad (EP > 30)	Disengaged
Angry (EP > 20) + Sad (EP > 30) + Fear (EP > 30)	Disengaged
Sad (EP > 30) + Fear (EP > 30)	Disengaged
Fear (EP > 30) + Surprise (EP < 60)	Disengaged
Neutral (EP < 60)	Disengaged

From Figure 7.12, the web camera captures the real-time facial expressions of learners, which are then used to determine the Engagement Index (EI) by the FER system. This index values categorizes engagement into three states: highly engaged, engaged, or disengaged, and also provides an engagement level in percentage terms (Table 7.7). The system recognizes a total of eight facial emotion classes, trained using three different datasets: CK+, JAFFE, and an In-house dataset.

The emotional states and their corresponding Engagement Probability (EP) thresholds are used to classify the level of engagement of participants. Highly engaged learners are identified through specific emotional patterns, such as Happy with an EP of 50 or above, Surprise with an EP of 80 or above, or Neutral expressions with an EP of 80 or more. On the other hand, learners are categorized as engaged (but not highly engaged) under conditions like Happy with an EP below 50, Surprise with an EP between 60 and 80, or Neutral with an EP within the same range. Engagement is also observed when Neutral expressions (EP > 60) are combined with Angry (EP < 20) and Sad (EP > 30). Disengaged states are characterized by emotions like Angry with an EP greater than 20, Fear with an EP above 30, or combinations such as Sad (EP > 30) and Fear (EP > 30), or Neutral expressions with an EP below 60. Additionally, extreme combinations of negative emotions, such as Angry (EP > 20), Sad (EP > 30), and Fear (EP > 30), also signify disengagement.

The predicted EP score is used as input to determine the engagement state. The EI is calculated using Equation (7.1), based on the predicted emotion's value to determine the engagement states.

$$EI = \sum(EP_i \; X \; WE_i) \qquad (7.1)$$

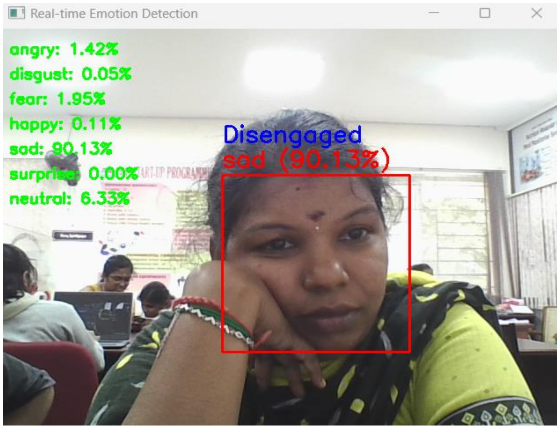
- where EP = Emotion Probability for emotion i (Emotion = Neutral, Angry, Sad, Happy, Surprised, Sleepy, Disgust, and Fear)
- WE = Weight of the corresponding Emotion i

The EP (emotion probability) score is generated by a fusion classifier consisting of CNN-BiLSTM and DSCSA models, along with the corresponding emotion weight. Emotion Weight describes the emotional state's value that reflects the learner's engagement at that moment. The engagement percentage is then calculated based on the EI given in Equation (7.1). The weights for corresponding emotions are scaled and represented in Table 7.8.

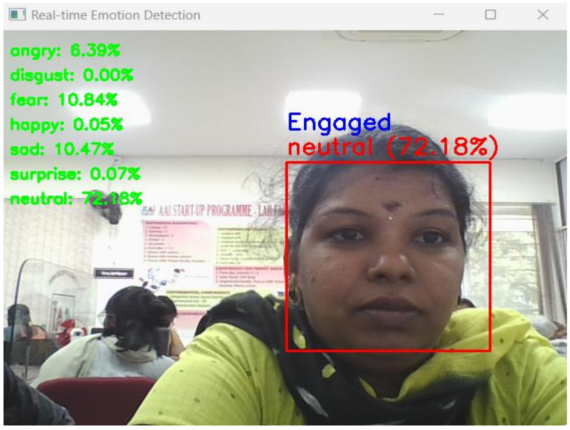
Table 7.8 Weight for the corresponding Label [23]

Emotion	Weight Value
Happy	0.8
Suprise	0.8
Fear	0.6
Neutral	0.5
Angry	0.3
Sad	0.3
Sleepy	0.2
Disgust	0.3

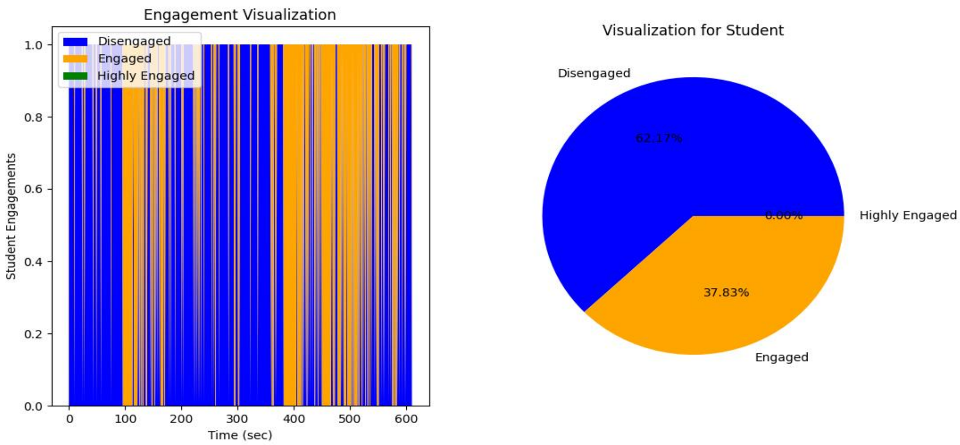
Figure 7.13 shows that learners without mulsemedia effects felt somewhat bored while watching the learning content, which may explain the fluctuations between engagement and disengagement. Ultimately, the results indicated that learners were more disengaged than engaged with the content.



(a)



(b)



(c)

Figure 7.13 (a) and (b) show participant engagement and disengagement detection without mulsemedia effects. (c) presents the overall results of learner engagement without mulsemedia.

Figure 7.13 shows two visualizations related to student engagement. The left graph, titled "Engagement Visualization," depicts engagement levels over time for individual students, using a color-coded scheme where blue represents disengaged, orange represents engaged, and green represents highly engaged students. The x-axis represents time in seconds, while the y-axis represents student engagement as a proportion. This graph shows frequent shifts between different engagement levels throughout the observed period. The right pie chart, titled "Visualization for Student," illustrates the overall engagement distribution, revealing that 62.17% of students are disengaged, 37.83% are engaged, and none are highly engaged. This suggests a predominance of disengagement among students with some engagement but no highly engaged students when multimedia learning.

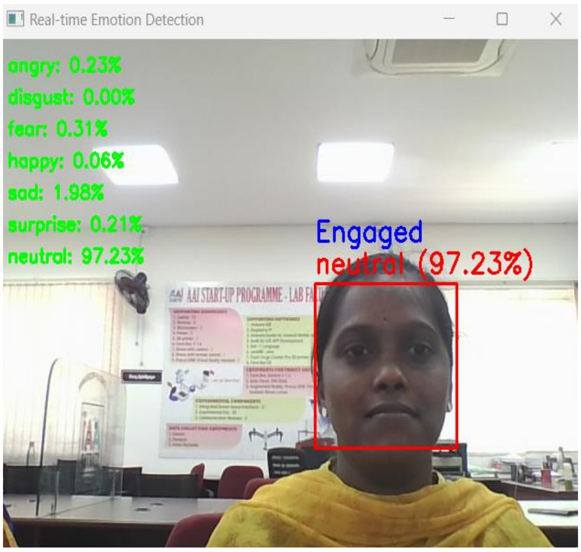
Control Group

Participants	Time Step	Dominant Emotion	Emotion Percentage	Engagement
1	0	sad	66.53683186	Disengaged
	1	sad	80.87893724	Disengaged
	2	sad	49.4366467	Disengaged
	3	sad	91.76254819	Disengaged
	6	sad	68.43494177	Disengaged
	7	sad	58.52235657	Disengaged
	8	sad	43.22674057	Disengaged
	9	neutral	47.66984852	Disengaged
	10	neutral	67.3910737	Engaged
	11	neutral	48.61036539	Disengaged
	12	angry	37.92746663	Disengaged
	13	neutral	40.63746929	Disengaged
	14	sad	37.72041426	Disengaged
	15	neutral	46.78665698	Disengaged
	16	neutral	38.5661602	Disengaged
	17	neutral	37.64521778	Disengaged
	18	sad	38.41333251	Disengaged
	19	angry	39.72476125	Disengaged
	20	sad	38.31894835	Disengaged
	21	sad	38.30789854	Disengaged
	22	sad	67.20539537	Disengaged
	23	neutral	38.08687627	Disengaged
	24	neutral	57.39126205	Engaged
	25	sad	58.29239239	Disengaged

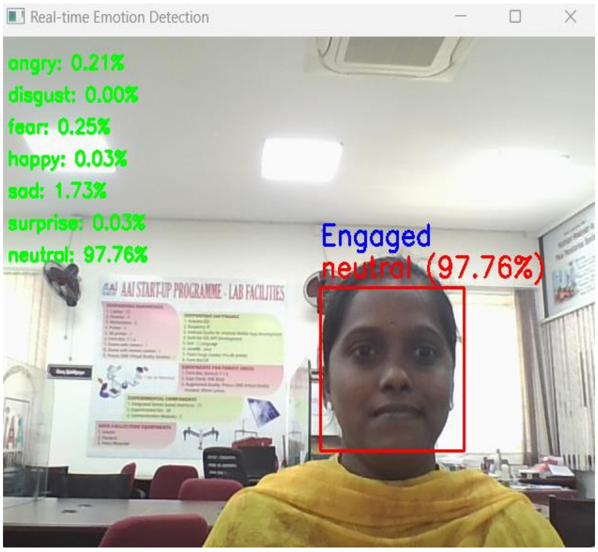
Figure 7.14 Emotion probability and engagement detection results were recorded during the assessment without mulsemedia effects

During the assessment of individuals' emotional reactions and their likelihood of being highly engaged, engaged, or disengaged, the data was stored in a CSV (Comma-Separated Values) file, as shown in Figure 7.14, for the final evaluation.

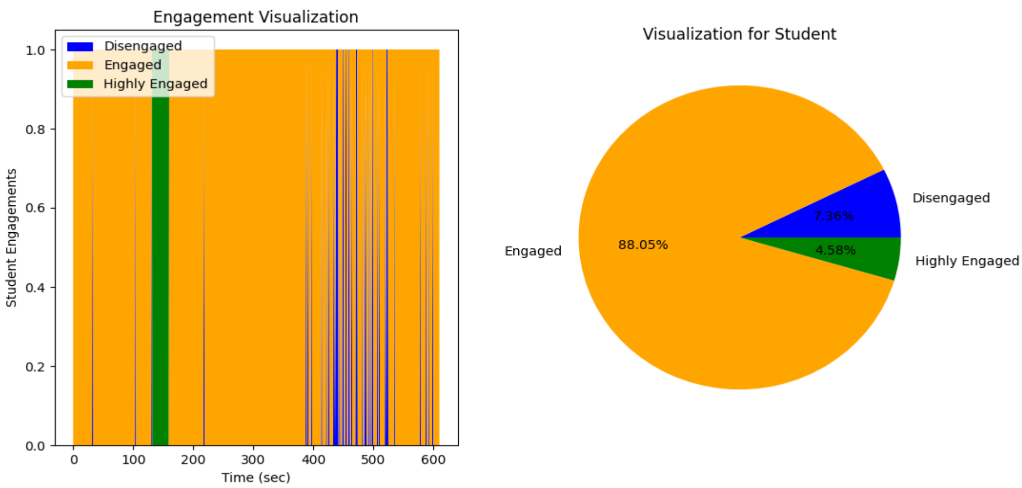
Figure 7.15 illustrates the engagement levels of a student over 600 seconds with mulsemedia effects. In the left panel, a bar chart titled "Engagement Visualization," shows the temporal distribution of engagement states. The majority of the time, indicated by the orange bars, the student was engaged. There is a brief period of high engagement, shown by the green bar, around the 100-second mark. Blue bars, representing disengagement, appear intermittently, especially towards the latter part of the timeline. The right panel, a pie chart titled "Visualization for Student," summarizes these engagement states, revealing that the student was engaged 88.05% of the time, disengaged 7.36% of the time, and highly engaged 4.58% of the time. These visualizations collectively highlight that while the student was predominantly engaged, there were notable instances of both disengagement and high engagement during the assessment period with mulsemedia effects.



(a)



(b)



(c)

Figure 7.15 (a) and (b) show participant engagement detection with Mulsemedia effects. (c) presents the overall results of learner engagement with Mulsemedia

In a similar way, 10 participants were involved in the control group, and another 10 participants were involved in the experimental group. Each participant's emotion probability has been stored individually as a CSV file, as shown in Figures 7.14 and 7.16. Afterward, the results were combined and analyzed to find the overall difference in satisfaction and engagement levels between the two groups in the e-learning environment. The combined results are shown in Figure 7.17.

Experimental Group

Participants	Time Step	Dominant Emotion	Emotion Percentage	Engagement
5	0	neutral	66.69196884	Engaged
	1	neutral	79.11374818	Engaged
	2	neutral	83.37296049	Engaged
	3	neutral	91.28611684	Engaged
	4	neutral	86.61549687	Engaged
	5	neutral	71.30287715	Engaged
	6	neutral	95.13882457	Engaged
	7	neutral	93.85132863	Engaged
	8	neutral	82.66880512	Engaged
	9	neutral	93.9279549	Engaged
	10	neutral	78.50936651	Engaged
	11	neutral	89.30434585	Engaged
	12	neutral	89.66869054	Engaged
	13	neutral	72.27879763	Engaged
	14	neutral	91.82698678	Engaged
	15	neutral	88.27546379	Engaged
	16	neutral	80.30006651	Engaged
	17	neutral	91.76896811	Engaged
	18	neutral	80.50628304	Engaged
	19	neutral	56.66629672	Engaged
	20	neutral	74.30566992	Engaged
	21	neutral	86.16896785	Engaged
	22	neutral	82.60566592	Engaged
	23	neutral	75.06996094	Engaged

Figure 7.16 Emotion probability and engagement detection results were recorded during the assessment with mulsemedia effects.

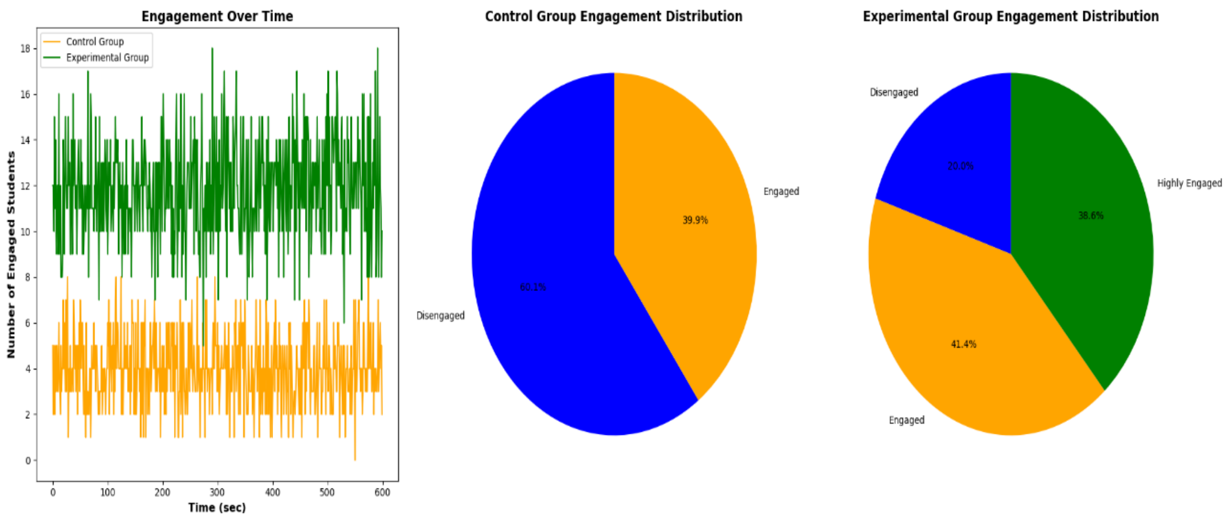


Figure 7.17 Plotting overall real-time EI of both group results

Figure 7.17 presents data on student engagement for both the control and experimental groups through three distinct visualizations. The left plot illustrates the number of engaged students over time. The control group (green) demonstrates a higher and relatively stable engagement level, whereas the experimental group (orange) exhibits lower engagement throughout the observed period. The middle pie chart represents the engagement distribution in the control group, revealing that 60.1% of students were disengaged, while 39.9% were engaged. In contrast, the right pie chart shows the engagement distribution in the experimental group, where 20.0% of students were disengaged, 41.4% engaged, and 38.6% highly engaged. These visualizations indicate a more varied engagement pattern in the experimental group compared to the control group. This suggests that mulsemedia-synchronized learning provided a better learning experience, fostering greater attention and focus among participants compared to conventional approaches. Additional results from participants are provided in Appendix II.

7.5 Chapter Summary

This chapter highlights the challenges associated with conventional multimedia approaches in learning environments and introduces innovative solutions using mulsemedia to create immersive learning experiences. A mulsemedia-enhanced web portal was developed, integrating haptic, olfactory, and vibration effects through IoT components. The effectiveness of these mulsemedia features was evaluated through QoE assessments, incorporating both subjective and objective measurements.

In the subjective assessment, a QoE questionnaire was designed to gauge learners' satisfaction levels with and without mulsemedia. The results revealed that learners expressed higher satisfaction and strongly recommended mulsemedia-based learning environments over conventional multimedia-based approaches.

In the objective measurement, learners' facial expressions were recorded and analyzed to predict their engagement levels during interactions with learning content. The findings demonstrated that learners were significantly more engaged with mulsemedia-synchronized content than with conventional multimedia content. Participants also suggested implementing this type of learning environment in future classroom settings.