Original Article

Emotions Of Users In Front Of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

Emoções de Usuários Frente a Vídeos Educacionais: um estudo de caso voltado à construção de orientações para a produção de moocs

Emociones De Los Usuarios Ante Videos Educativos: un estudio de caso enfocado en la construcción de directrizes para la producción de moocs

Autores: Vitória Vasconcellos da Luz, Sandra Dutra Piovesan e Valesca Brasil Irala.

Abstract

In recent decades, the growth of online education practices has been remarkable, aimed at promoting different types of knowledge and also professional training. In this scenario, massive and open online courses (known by the acronym MOOCs) gained greater visibility. In this modality, there is a centrality, in the pedagogical design, of the presence of educational videos. On the other hand, there are studies that characterize the role of emotions as fundamental, capable of accelerating or disturbing the learning process. Therefore, this study aims to analyze the perceptions and emotions of users when watching videos made available in a MOOC, aiming to define guidelines for those who intend to develop videos for MOOCs. The research took place in four stages, the first being a bibliographic review and definition of the method for inferring emotions. The method chosen was the analysis of facial expressions, through the use of the online tool EZ-MMLA ToolKit, which captures the image and defines the emotions, considering the basic

emotions and the application of a questionnaire called AEQ (Achievement Emotions Questionnaire) that aims to assess the emotions of achievement experienced during educational tasks, adapted for research. Subsequently, the users' emotions were evaluated when watching the videos. The third step was data analysis and, finally, the definition of guidelines on creating videos for MOOCs. In general terms, it was possible to define that the best accepted video, among those presented, was the one that shows a tutorial and the least accepted was the one that presents synthetic voice narration. Other comments from the participants regarding the analyzed videos emphasize the need for an expressive number of examples with images, less information (short duration videos), removing background music, and, as a positive point, the presence of an avatar and posture who passes the information.

Keywords: Emotions. Educational videos. MOOCs.

1. Introduction

In light of the notable technological growth, especially in recent years, it is natural for technology to permeate various human realms, from the personal to the academic and professional spheres. The rise of virtual education is a reality that allows students to access the most diverse learning experiences available today, making technology a significant anchor for education, particularly in non-face-to-face modalities. One of these modalities is digital courses, which provide recorded videos made available online, allowing students to access them from anywhere at any time (SANTOS, 2009).

The interest in online learning platforms became even greater in times of social distancing following Covid-19, as, for a long period (especially in the year 2020 and partially in 2021), formal education occurred exclusively through these platforms (SILVA; ACCORSI; MUNHOZ, 2020). This phenomenon did not go unnoticed in the current context, marked by the return of face-to-face learning in most educational scenarios, and it continues to reflect overt trends for contemporary education.

Looking back in time, there was already a demand in this field in 2008 with the creation of Massive Open Online Courses (MOOCs), courses designed to cater to a large number of students, in an online and open manner, that is, free of charge, allowing, in addition to training, the experimentation of new teaching methods. Thus, despite the challenges, institutions identified opportunities to ensure the engagement of a portion of society that no longer

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

2

fit, for various reasons, into the face-to-face teaching model, adopting free online courses as a continuity strategy for teaching (SICILIANI, 2016).

MOOC platforms have made significant advances in a relatively short period, and the trend for the coming years is for the offering to expand even further (FONTANA; LEFFA, 2018). Given this increase, it has also been possible to understand the main problem of the teaching model offered by this type of course: the high dropout rate of users (SILVA; ACCORSI; MUNHOZ, 2020). The main MOOC platforms suffer from this latent symptom: student attrition. Therefore, it is necessary to seek alternatives that can maintain users' interest in MOOCs and identify the reasons leading to the high dropout rate.

The use of videos is common in online teaching modalities. In this sense, it becomes relevant to understand students' emotional responses to this artifact. The audiovisual language responds to the sensitivity of a large part of the adult population in contemporary times. Videos are dynamic and affect emotion before reason, reaching the user in different ways and forms, starting with the emotional aspect and subsequently reaching reason (LAASER; TOLOZA, 2017).

Studies highlight that emotional aspects play an important role in learning and cognitive processes developed by students, as the construction of knowledge about different objects is interconnected with cognitive schemes and emotional dimensions (PIAGET; INHELDER, 2003). Emotions affect learning in different realms, as the greater the emotional involvement, the more cognitive functions such as attention, perception, and memory are mobilized (DA FONSECA, 2016).

One way to analyze emotions is through the inference of emotions via affective computing. Four channels are possible parameters for analysis: textual, visual, physiological, and multimodal. The use of the multimodal channel (integration between two channels) is notable, with the most commonly used integration being the union of the textual (questionnaires) and visual (facial expressions) channels (YADEGARIDEHKORDI et al., 2019). Facial expression is one of the best direct methods to accurately detect emotional states, especially in virtual learning environments (YADEGARIDEHKORD et aI., 2019), because the movement of facial muscles can express different emotions (BUSIN, 2013).

For the recognition of emotions, the existence of some primary or

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

3

basic fundamental emotions is considered. These emotions are considered universal in the aspects of facial expression and the possibility of recognition when manifested. These emotions can be defined as joy, sadness, fear, anger, surprise, and disgust (DAMÁSIO, 2000).

Despite the constant growth of MOOCs, the number of Brazilian studies addressing this theme is still small, which reinforces the importance of countries like Brazil making their presence felt in the universe of such research, considering the fact that Brazilian institutions already offer this mode of teaching, including on international platforms. Research investigating the experience with MOOCs can provide a greater understanding of this mode of education and user behavior in these courses (ZHENG *et al.*, 2015).

The development of research on MOOCs is justified as it is a topic with great expansion, due to the specificities of this mode of teaching, which stands out for promoting student autonomy and for offering quality knowledge for free. There is a strong tendency for platforms that offer these courses to expand more and more in the coming years, considering the growth that has occurred in a short space of time and substantial changes in the study profile of new generations (FONTANA; LEFFA, 2018).

With this, we propose a multimodal analysis that shows how the insertion of different characteristics in videos offered in a MOOC influences emotional states. In light of this, the present study, of an exploratory nature, aims to analyze users' perceptions and emotions when watching videos provided in a MOOC, with the practical goal of the research proposal being to define guidelines for those who intend to develop videos for this purpose. In the next section, we will detail the free computational tool used to conduct the research, in order to highlight its potential in the educational field.

2. Choice of Tools Used

For the execution of the research, specifically for the inference of emotions, the EZ-MMLA ToolKit was used, which allows multimodal emotion recognition through the implementation of an emotion detection application. This tool uses a Convolutional Neural Network (CNN) trained to assess an individual's emotional state through their facial expression.

> Human facial muscles are innervated by a single nerve: the facial nerve (cranial nerve VII) that resides within the

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

4

brainstem, where involuntary and unconscious expressions that occur spontaneously are controlled, while the motor cortex is involved in intentional and consciously controlled facial expressions. Because of this nerve relationship, the same regions of the brainstem that activate facial expressions also control emotional processing and regulation - which makes it possible to study a person's physiological state through their facial expressions (HASSAN; LEONG; SCHNEIDER, 2021, p. 583, own translation).

In light of this, a person's physiological states and emotions are often physically expressed through complex actions involving facial expressions, making it possible to evaluate a person's affective reaction through their facial expressions (HASSAN; LEONG; SCHNEIDER, 2021). The EZ-MMLA ToolKit receives a video sequence as input and then isolates the subject's face before using the CNN model to predict the emotion that the subject is expressing in each frame (HASSAN; LEONG; SCHNEIDER, 2021).

There are three main components behind the automatic analysis of facial expression (HASSAN; LEONG; SCHNEIDER, 2021):

- Face detection: the implementation carries out face detection to identify and isolate the subject's face for subsequent inference. The toolkit uses the OpenCV-python implementation of the Viola-Jones object detection framework. As the toolkit uses an existing implementation, all the training is done beforehand. This allows the toolkit to quickly identify faces in a given real-time video input. In short, the framework allows the toolkit to draw a box around a face (even when the subject's head moves or the lighting changes), identifying the location of the face and scaling it, as shown in Figure 1;

- Identification and recording of facial landmarks: once the person's face is detected, EZ-MMLA identifies facial landmarks, such as eyes, eyebrows, lips, and nose, to get a mapping of the face. Visually, the identification of facial landmarks is represented by the thin green contours around the subject's facial features, as illustrated in Figure 1. The pre-trained face model is resized based on the features detected in the facial mapping, allowing a refined adjustment of the pre-trained model to match the user's actual face structure.

- Emotion classification: Finally, the frames are passed through the CNN, a classification model that has been pre-trained to classify

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

5

emotions based on a large set of labeled data. The CNN is trained on a large number of faces to estimate the normative distributions of emotional features. The model produces a set of probabilities for the face identified in the provided video frame, indicating the likelihood of each emotion. In Figure 1, these probabilities are indicated on the left in blue text, as part of the 'Emotional Report'. Each emotion is associated with a value that indicates the probability of the subject experiencing that emotion based on their facial expression.
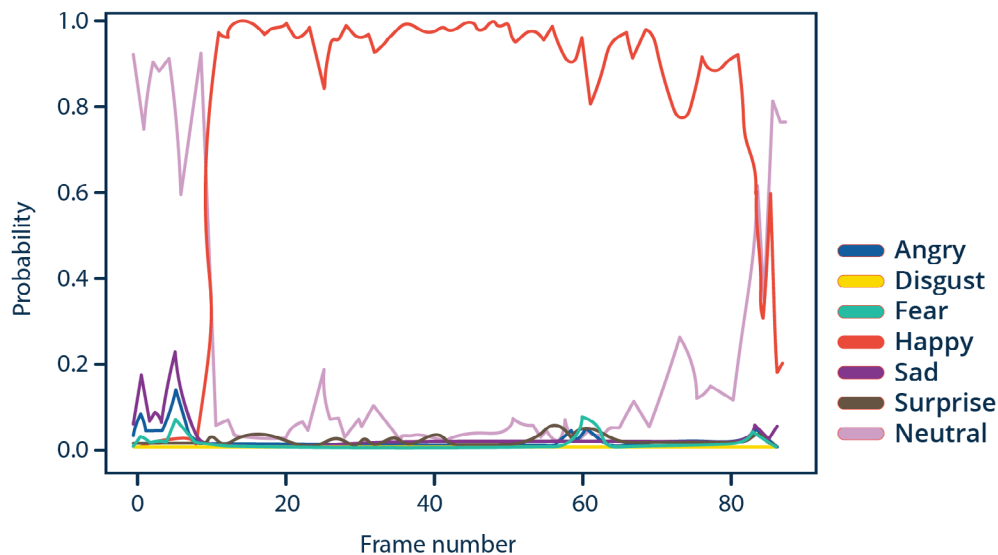
Figure 1 — A use case of the emotion detection function

Source: Hassan, Leong e Schneider (2021).

Figure 1 shows an example where the user presents a relatively neutral expression, and the multimodal toolkit successfully identified this. Observing the emotional report, the results seem to map well to intuitive expectations. The report suggests the possibility that the subject may be expressing fear or disgust, while the subject's expression likely conveys neutrality, sadness, or mild annoyance (HASSAN; LEONG; SCHNEIDER, 2021).

To use the emotion detection function, simply select the 'Emotion Detection' option. As soon as the page is loaded, the user only needs to activate their webcam, and the function will automatically detect the person's face and infer the emotions. After recording, it is possible to download the video and the probabilities of each test.

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

6

Figure 2 — Example of Results Visualization

Source: Hassan, Leong e Schneider (2021).

Subsequently, on the 'View Results' page, it is possible to observe the probabilities predicted by the model based on the input video, as shown in Figure 3. Additionally, on the same page, there is also the option to download the data in a CSV file format, as illustrated in Figure 3.

Figure 3 - Example of Results Output

Frame		Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
0	0	0.024298	0.000040	0.007412	0.002039	0.051032	0.000342	0.914837
1	1	0.075400	0.000234	0.023033	0.001283	0.157993	0.000625	0.741433
2	2	0.027957	0.000034	0.011739	0.000375	0.055193	0.000100	0.904604
3	3	0.037663	0.000053	0.013582	0.000446	0.074070	0.000194	0.873992
4	4	0.031092	0.000038	0.009945	0.000550	0.052167	0.000158	0.906051
5	5	0.058116	0.000139	0.020596	0.001673	0.116943	0.001106	0.801426
6	6	0.131211	0.000723	0.070025	0.004128	0.211424	0.008572	0.573917
7	7	0.065237	0.000348	0.057968	0.009963	0.102794	0.008120	0.755569
8	8	0.015605	0.000015	0.009772	0.006046	0.056028	0.000122	0.912411

Source: Hassan, Leong e Schneider (2021).

The EZ-MMLA ToolKit was combined with the application of a questionnaire in this research, aiming to compare the information obtained through the computational tool with the participants' reported perception. The questionnaire was chosen as the method because it is widely used in emotion assessments that employ textual channels. Several types of questionnaires

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

7

were found during the literature review, and the chosen model was the AEQ (Achievement Emotions Questionnaire), as it is the most popular in affective computing studies in the field of education (YADEGARIDEHKORDI *et al.*, 2019).

The AEQ was developed to assess various achievement emotions experienced by students in academic environments (PEKRUN *et al.*, 2011). The original instrument consists of 24 questions aimed at recognizing the emotional states experienced during classroom activities. In this research, the questionnaire was adapted for specific purposes, utilizing 12 questions adapted from the original instrument to specifically evaluate the emotional state during video viewing. Some questions from the original instrument were not used due to the impossibility of adapting them to the research context, prioritizing the use of a concise questionnaire with prioritized aspects for emotion evaluation.

To assess the users' emotions, four videos available in a MOOC were selected. The videos address the topic of educational video development and are directed towards professionals or aspiring professionals who wish to develop educational videos in their professional activities. Each video has different characteristics, maintaining only a standardized color palette, duration, and sound effects. The standardized color palette in the videos was chosen to consider visibility requirements, allowing users to pay greater attention due to contrast, and it is also in accordance with accessibility standards. Colors can produce significant impressions, sensations, and sensory reflections, as each color has a specific vibration in our senses and can act as a stimulant or disruptor of emotions, consciousness, and impulses and desires (FARINA; PEREZ; BASTOS, 2013).
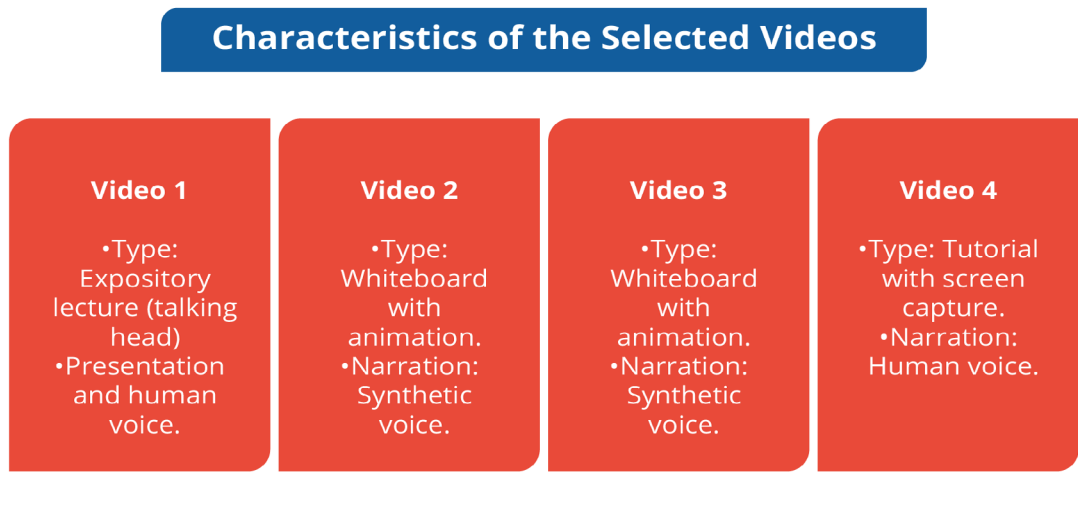
As for the duration, all the videos have an average length of 7 minutes. According to the literature, videos that are well-received by students typically range from 5 to 10 minutes and have a more dynamic presentation style, such as whiteboard videos or animations (STANLEY; ZHANG, 2018). The images used to exemplify the video content were selected based on their quality, allowing users to visualize the content in the best possible way and avoiding visual fatigue. Preference was also given to using images with varied color palettes and vibrant colors to capture the user's attention.

In the videos where content presentation or narration by the instructor takes place, established techniques were used to promote greater student engagement. Narration was performed using a synthetic voice in one of the

videos to assess how this modification affects emotions. For the selection of background music used in the videos, upbeat tracks with low volume were chosen to attract the student's attention. Instrumental music was preferred to avoid the user focusing more on the lyrics of the songs rather than the video content. Figure 4 presents the main characteristics of each selected video.

Figure 4 — Characteristics of the Selected Videos

Characteristics of the Selected Videos

Video 1	**Video 2**	**Video 3**	**Video 4**
•Type: Expository lecture (talking head) •Presentation and human voice.	•Type: Whiteboard with animation. •Narration: Synthetic voice.	•Type: Whiteboard with animation. •Narration: Synthetic voice.	•Type: Tutorial with screen capture. •Narration: Human voice.

©VG Educacional

Source: Created by the authors.

The first video addresses the topic of engagement in videos. The content was presented in a theoretical manner through an expository lecture, with the instructor explaining the topic and appearing in the video. This video was selected to evaluate the level of acceptance of content presentation by the instructor, without the use of image exemplification.

The second video discusses the different ways of using educational videos. The topic is presented theoretically, with narration by a synthetic voice. The video utilizes a whiteboard format, with the display of a handwriting or dragging information, and includes some examples with images. The images used were chosen to be visually appealing in terms of colors, in order to capture participants' attention, and they were of high quality, allowing the content to be viewed in the best possible way. Particularly in the second video, the narration by a synthetic voice was used to assess how the presence of this type of narration affects users' perceptions and emotions, in order to make a comparison with findings from research that highlight a better learning experience when listening to humans rather than a synthetic voice (HILLAIRE; INIESTO; RIENTIES, 2019).

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

9

The third video addresses the elements of audiovisual content. The topic is presented theoretically, with narration by a human voice. The video also utilizes a whiteboard format, with the display of a handwriting or dragging information. Examples with images and an avatar are used during the explanation. The selection of this video is especially for evaluating the differences in participants' evaluations between videos 2 and 3, considering that they are models that have many similarities, with the main difference being the mode of narration, in this case, by a human voice.
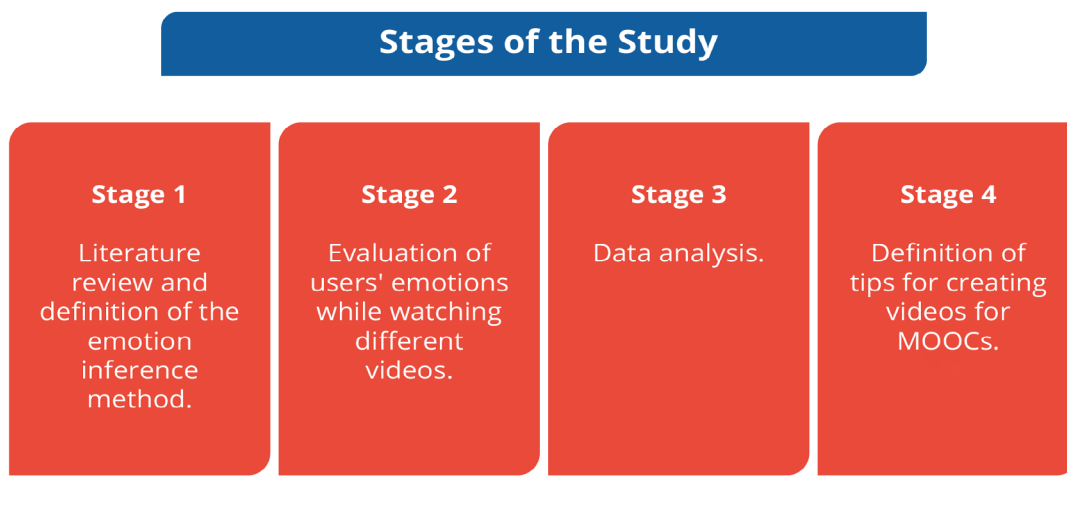
Video 4 presents a video creation and editing tool. The presentation is done through a tutorial video with screen capture and human narration. All the necessary steps for creating a video using the chosen tool, Animaker, are explained in the video. This video was selected considering the dynamic nature of tutorials, allowing us to confirm the high acceptance of these characteristics, as mentioned in several studies.

3. Methodological Procedures

The research is of a qualitative, descriptive, and exploratory nature, providing new perspectives on an already empirically known reality (TRIVIÑOS, 2011). A significant portion of studies in the field of education follows this approach. As pointed out by Denzin and Lincoln (2003, p. 32), "qualitative research can no longer be seen from a neutral or objective positivist perspective", but rather perceives research itself as a process, within a specific interpretive field, understood in a situated manner, giving centrality to lived experience, regardless of the number of participants involved.

To achieve this, a case study approach was adopted, which seeks to investigate a specific case that is well-defined and contextualized in time and space (VENTURA, 2007), although without the pretension of generalization as in quantitative research. For this study, four videos made available in a pilot MOOC developed at a public university in the interior of Rio Grande do Sul were used. The perceptions of eleven users interested in the proposed theme, i.e., individuals who have the potential to produce educational videos in the future, were observed.

Figure 5 presents the chronological order of the study's stages, aiming to achieve the central objective of analyzing emotions in response to the selected videos and, thus, reaching a practical goal of providing guidelines for video production for MOOCs.

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

10

Figure 5 — Stages of the Study

Stages of the Study

Stage 1

Literature review and definition of the emotion inference method.

Stage 2

Evaluation of users' emotions while watching different videos.

Stage 3

Data analysis.

Stage 4

Definition of tips for creating videos for MOOCs.

© VG Educacional

Source: Created by the authors.

In the initial stage, a literature review was conducted to investigate the main contributions in the field, based on recent literature (LUZ; PIOVESAN; IRALA, 2023). During the reading of selected publications, the available methods for emotion inference were analyzed, allowing the definition of the method to be used in the present study, considering its specificities. As mentioned earlier, emotion inference can be classified into channels, which include the textual channel (questionnaires or interviews), the visual channel (observable behavior or facial expressions), the physiological channel (variation in heart rate or pupil diameter), and the multimodal channel (combination of two channels) (YADEGARIDEHKORDI et al., 2019).
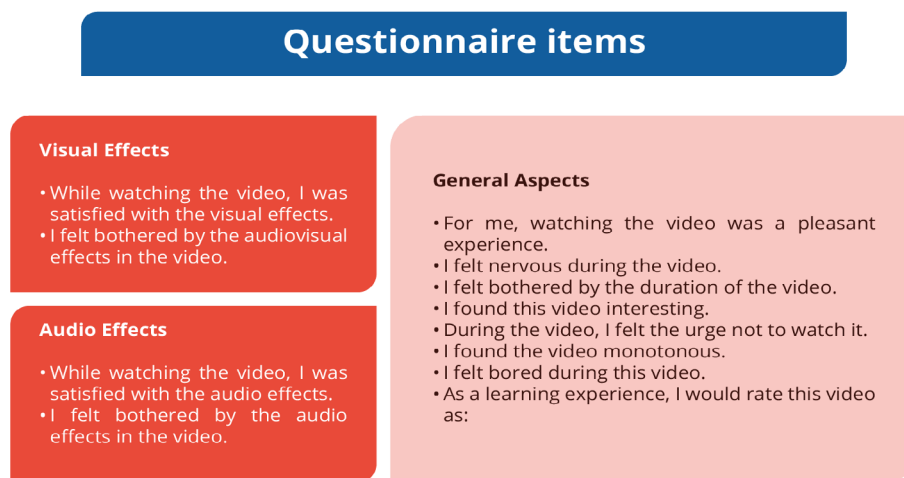
The chosen method for emotion inference was the multimodal channel, using the integration of the visual and textual channels. For facial expression recognition, the EZ-MMLA ToolKit was selected, which considers the basic emotions to define facial expressions, including a neutral expression when the detected facial expression does not fit into any of the basic emotions.

This method of emotion inference, using the visual and textual channels, was chosen for its higher accuracy, allowing for a comparison between the data obtained from each channel. Furthermore, preference was given to less invasive methods for emotion inference, considering the available resources. Methods involving physiological aspects such as heart rate and pupil variation, although potentially more reliable, require more advanced technological equipment.

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

11

For participant selection, the convenience sampling technique was employed, where the researcher seeks participants who are accessible and available to collaborate (FREITAG, 2018). Participants were invited based on their interest in developing educational videos in their professional practices. The majority of them belong to the same educational institution and are mostly students of graduate programs. It is important to note that during the data collection period in 2022, the COVID-19 pandemic was still ongoing, which posed challenges in data generation and accessing a larger number of potential participants.

After conducting affectivity inference using the EZ-MMLA ToolKit, an adaptation of the AEQ questionnaire was administered to record participants' perceptions regarding the watched video, as illustrated in Figure 6. The AEQ was adapted according to the research objectives. The response options were based on the Likert scale, where participants select their level of agreement with one of the five presented items: Strongly Agree, Agree, Neutral/Indifferent, Disagree, Strongly Disagree.

Figure 6 — Questionnaire items

Questionnaire items

Visual Effects
- While watching the video, I was satisfied with the visual effects.
- I felt bothered by the audiovisual effects in the video.

Audio Effects
- While watching the video, I was satisfied with the audio effects.
- I felt bothered by the audio effects in the video.

General Aspects
- For me, watching the video was a pleasant experience.
- I felt nervous during the video.
- I felt bothered by the duration of the video.
- I found this video interesting.
- During the video, I felt the urge not to watch it.
- I found the video monotonous.
- I felt bored during this video.
- As a learning experience, I would rate this video as:

© VG Educacional

Source: Created by the authors.

In addition to the presented items, there were three open-ended questions in which participants could report what they found most and least enjoyable in the videos and provide suggestions, criticisms, or comments. Regarding the questionnaire, participants were asked to provide not only responses about their perception of the watched video but also some sociodemographic

and academic information. Age, gender, and undergraduate course were identified. Participants' identification data were not obtained to provide greater freedom in their responses.

For the analysis of the emotions inferred using the EZ-MMLA Toolkit, the obtained results were first translated since the tool presents the results in English. The translation was done with regard to the predominant emotion in each of the watched videos. For the analysis of the questionnaire data, the participants' profiles were initially outlined, followed by an analysis of each participant's overall perception of each of the four videos watched. Lastly, an analysis of the responses obtained from the open-ended questions was conducted.

Furthermore, a comparison was made between the data obtained from the facial expression-based emotion inference and the data obtained from the questionnaire, considering the different characteristics of the displayed videos and the individual profiles of each participant. Finally, based on the results obtained, certain characteristics were identified as being better accepted in the videos, while specific aspects were identified as requiring more attention when creating educational videos for MOOCs.

As a result, several tips were defined for those wishing to offer such courses. These tips were devised based on the analysis conducted, using appropriate language and, whenever possible, providing examples for each item presented. In addition to a technical approach, these tips and points of attention aim to assist and guide instructors, following the main recent literature on the subject.

4. Results And Discussion

To obtain the research results, the mentioned stages were conducted, involving the participation of 11 participants who contributed to data generation. Initially, sociodemographic and academic background data were collected. The collected data included age, gender, and undergraduate course. Figure 7 presents the data regarding the participants' ages.

Figure 7 — Participants' Ages
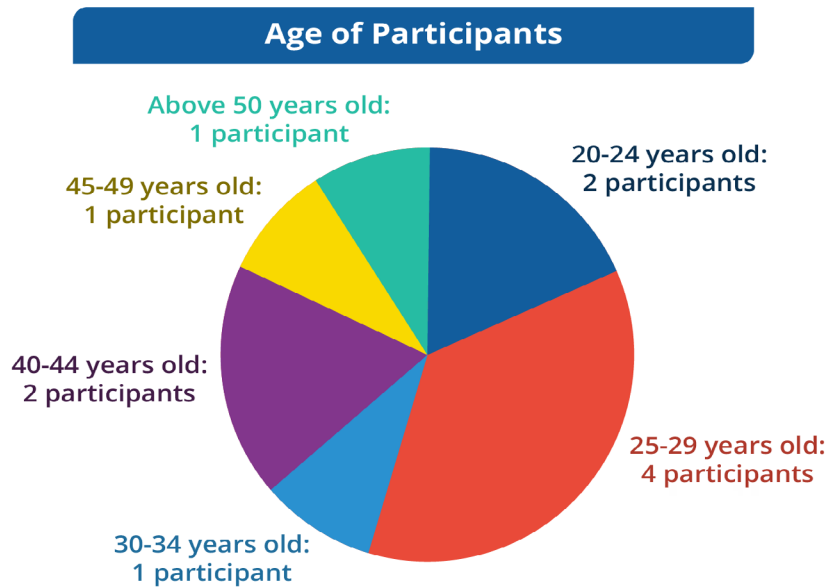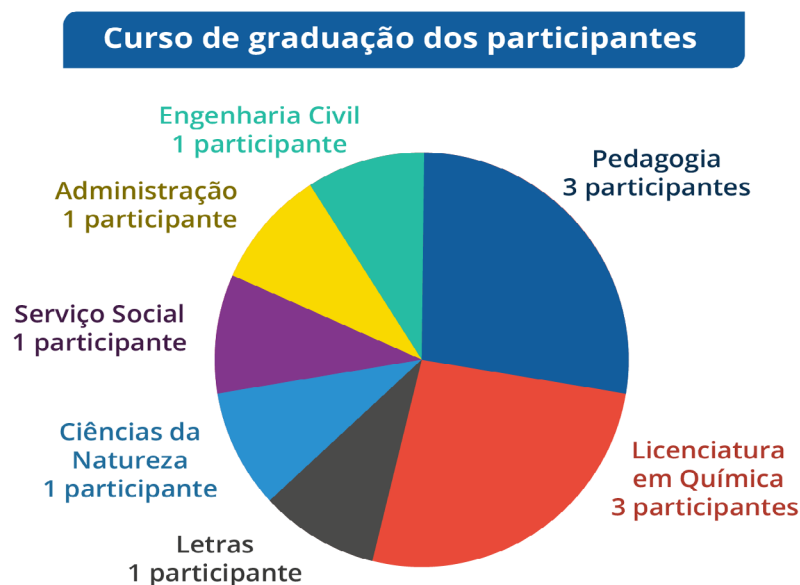

Source: Created by the authors.

Through the analysis of the data related to the participants' ages, it was possible to identify a heterogeneous audience, with participants from different age groups. Regarding the participants' gender, Figure 8 displays the obtained data.

Figure 8 — Participants' Gender
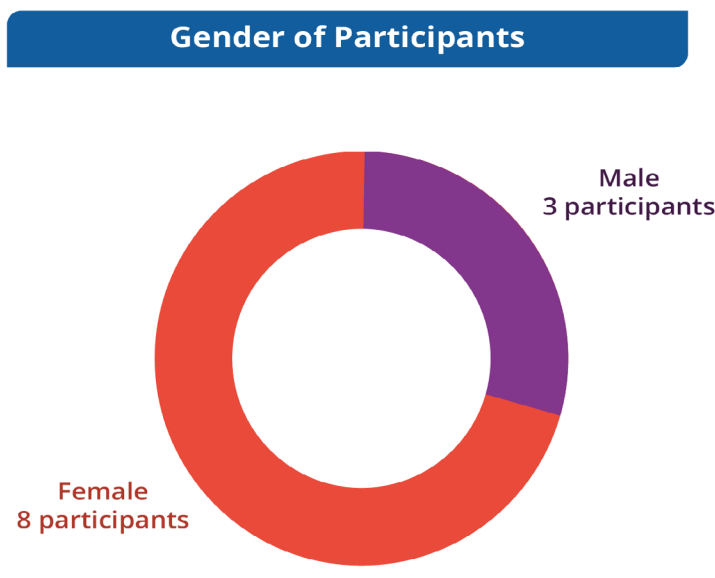
Source: Created by the authors.

Regarding the participants' gender, it is evident that the majority are female, with only 3 male participants. Finally, concerning their undergraduate education, Figure 9 presents the participants' courses, including both teaching and non-teaching programs. It is important to note that regardless of their specific undergraduate backgrounds, all participants share a common interest in developing educational videos, which is the central theme of the analyzed videos.
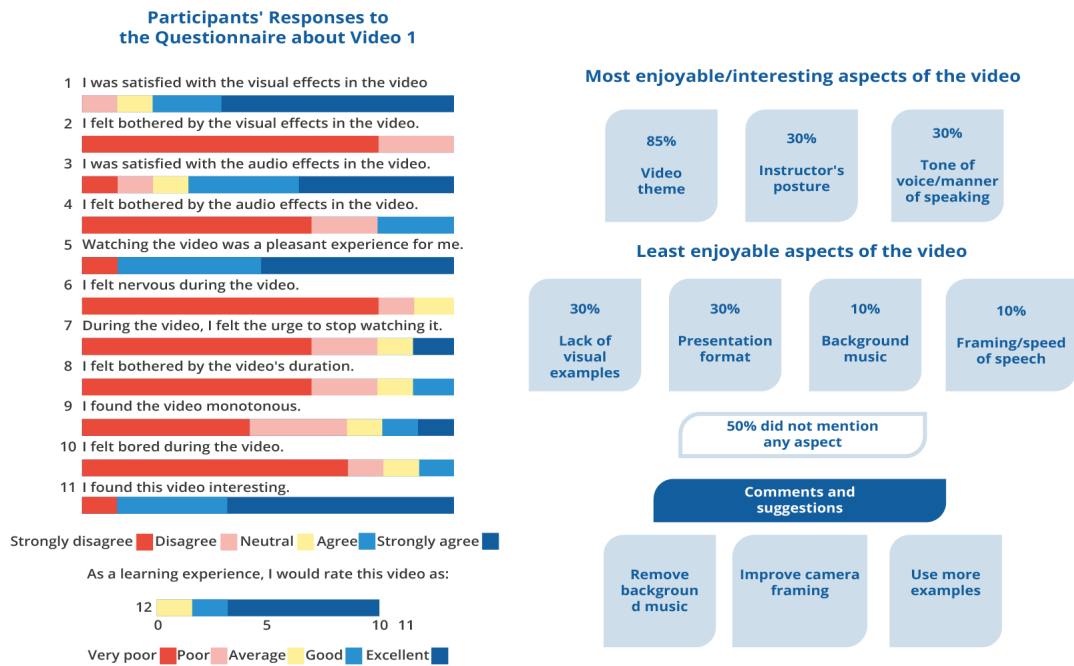
Figure 9 — Participants' Undergraduate Courses

Source: Created by the authors.

An analysis was conducted on the data obtained from the evaluation of participants' emotions, both through the inference of emotions from facial expressions and the data obtained from questionnaire responses. During the analysis of participants' facial expressions while watching the first video, the predominant emotion identified was neutrality. Regarding the questionnaire, participants' responses are presented in Figure 10.

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

15

Figure 10 — Participants' Responses to the Questionnaire about Video 1

Source: Created by the authors.

Based on the participants' responses to the first open-ended question, it is notable that the interest in the presented theme in the video was frequently mentioned, along with aspects related to the instructor and her manner of presenting the content. Some participants mentioned multiple items from the provided list. Among their responses, the following remarks can be highlighted: "The posture of the presenter; Pleasant tone of voice; There are no elements that could distract attention" (Participant 11).

Regarding the less enjoyable aspects, it is important to note that around 50% of the participants did not mention any negative aspects. The two most cited aspects were the lack of visual examples and the presentation format. It is worth mentioning that the video format, a lecture-style video, was intentionally chosen to analyze participants' perceptions of this type of video, and that is why it was not extensively edited with additional items and images. Among the responses indicating less enjoyable aspects, the following remarks stand out: "I believe that other visual elements would be interesting (to avoid relying solely on the spoken content)" (Participant 8); "The background audio can be distracting during the presentation" (Participant 11).

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

16

Lastly, participants had the option to provide comments or suggestions they deemed relevant. Since this was an optional item, not all participants provided suggestions or comments. The received comments included suggestions to remove the background music, improve camera framing, and incorporate more visual examples during the content presentation. Among the mentioned comments, the following remarks can be highlighted: "The camera should be more centered" (Participant 5); "There could be more illustrative images during the explanation" (Participant 3).

Overall, it is noticeable that the theme of the video was frequently mentioned by the participants, and according to their responses, participants with an interest in the topic can apply the content covered in their professional and academic endeavors, resulting in more positive perceptions of the video. This relationship is significant as it influences the presence of positive emotions in users. Additionally, the lack of visual exemplification throughout the presentation was a topic that was frequently addressed.

Based on the analysis of user responses regarding Video 1, some important tips for video creation have already been identified, especially regarding the use of visual examples, presentation style (instructor's posture, manner of speaking, etc.), and camera framing.

Regarding Video 2, an analysis of participants' facial expressions during its viewing identified sadness as the predominant emotion. The participants' responses to the questionnaire are presented in Figure 11.

Figure 11 — Participants' Responses to the Questionnaire about Video 2
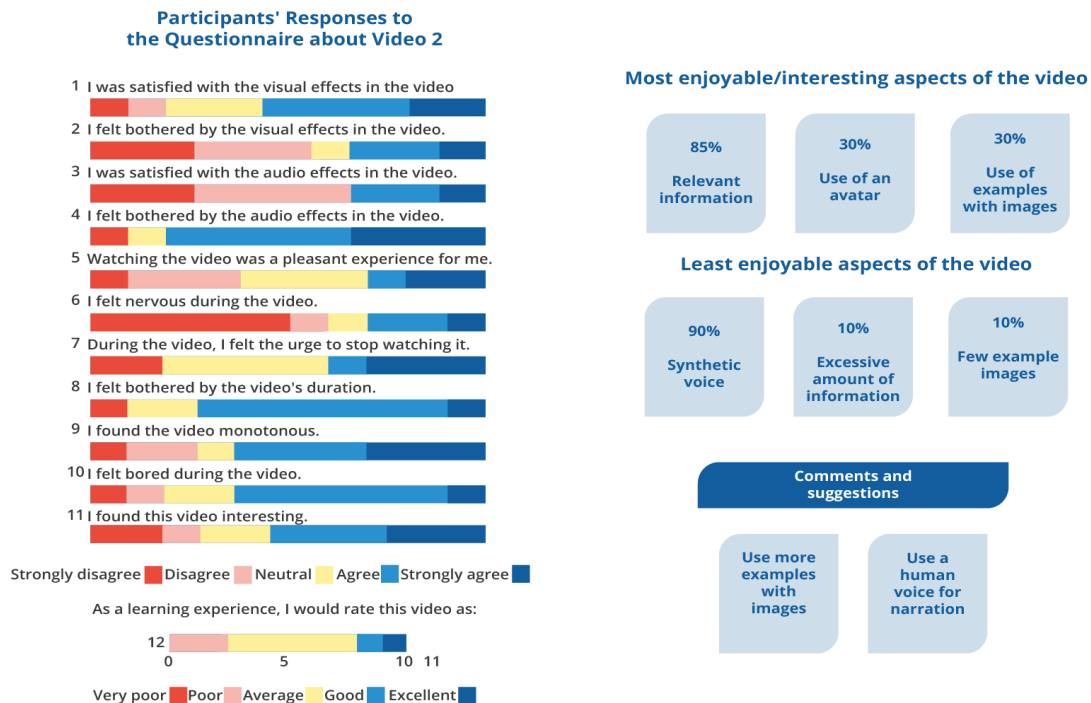
Source: Created by the authors.

Based on the participants' responses regarding the most enjoyable/ interesting aspects of the second video, the presented information and visual elements stood out, such as the use of an avatar and images for exemplification. Among the responses given to this question, the following statements can be highlighted: "The information was highly relevant. I also liked the images incorporated throughout the video" (Participant 8); "I found the avatar interesting" (Participant 5); "The information conveyed in the video is highly relevant and interesting for education, and the images used were well employed. However, I felt that the display time for each image was excessively long" (Participant 11).

Regarding the least enjoyable aspects, the use of a synthetic voice was the most frequently mentioned item. It is important to note that this choice was made specifically for this test in order to observe the participants' perception of this narration tool. Additionally, participants mentioned the limited use of images and the excessive amount of information added to the video. Among the comments provided by the participants regarding the least enjoyable aspects, the following statements stand out: "More images could be included to make the video more enjoyable and interesting. The synthetic

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

18

voice is also not very appealing; it would be better if someone narrated the information about the images" (Participant 11); "The time and the excessive amount of information" (Participant 2); "The synthetic voice made the video a bit confusing and dull to listen to because it does not express any emotion regardless of the sentence" (Participant 1).

As it was an optional item, not all participants provided suggestions and comments. The majority of the comments received suggested the use of a human voice for narration. Additionally, one participant suggested the use of more images as examples. Some of the mentioned comments include: "For this specific video, I would include more images and a real narrator's voice, as the synthetic voice seems too mechanical, making it difficult to understand and pay attention" (Participant 11); "The voice could be from the teacher in the first video, or there could be subtitles to accompany the narration" (Participant 3).

In general, it is evident that the synthetic voice was a problem for the participants. Some mentioned that it hindered comprehension and did not help maintain attention. It is noteworthy that the use of human narration was a widely mentioned suggestion, along with the use of more images as examples and the use of the avatar.

After analyzing Video 2, more important tips were identified, such as the importance of short videos (the most accepted videos are between 5 and 10 minutes in duration), the use of subtitles in educational videos (providing greater accessibility), and the use of the avatar. It is also necessary to avoid using synthetic voice narration, which was the most criticized aspect in the second video. As for Video 3, the analysis of participants' facial expressions during its screening identified neutrality as the predominant emotion. The participants' responses to the questionnaire are shown in Figure 12.

Figure 12 — Participants' Responses to the Questionnaire about Video 3
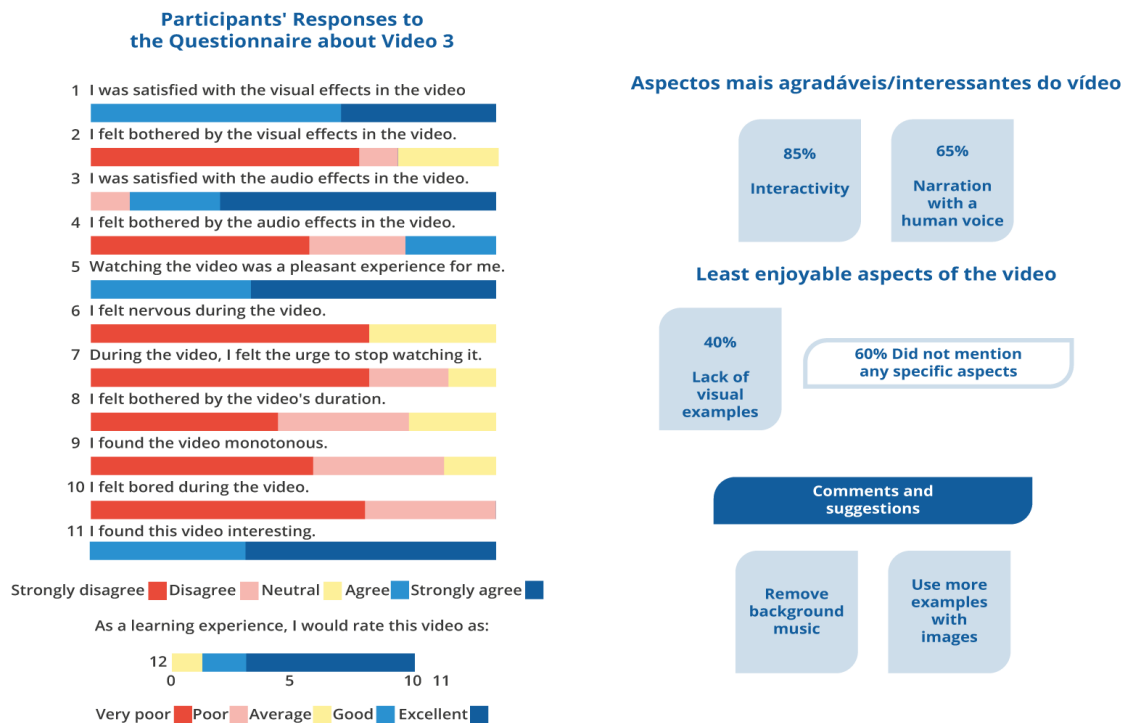
Source: Created by the authors.

Regarding the most enjoyable/interesting aspects, it is noteworthy that narration with a human voice and interactivity through the use of the avatar and whiteboard-style video were not mentioned as positive aspects by the participants. Some participants mentioned more than one item from the list, such as: "I found the animated video with human voice narration more enjoyable and it captured my attention, in addition to the content being personally interesting to me" (Participant 7); "I found the video more interactive, the speaker's speech became more pleasant with the images. The words being added to the whiteboard were very interesting" (Participant 11).

As for the least enjoyable aspects, about 60% of the participants did not mention any negative aspects, and the only aspect mentioned was the lack of visual examples during the video explanation. Among the items listed by the participants as less enjoyable, the following statements stand out: "I believe there is nothing unpleasant in this video" (Participant 3); "Few images were used to illustrate what was said" (Participant 11).

The received comments mostly suggest the use of more examples with images. Additionally, one participant suggested removing the background

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

20

music. Among the comments received, one participant stated: "I found it very interesting how the words appeared on the screen as if they were being written, but I still think the background music is unnecessary. I believe it distracts from the focus of the content being explained in the video. I also suggest adding images to illustrate the text being spoken during the video" (Participant 11).

Overall, it can be observed that the use of narration with a human voice was one of the most positively mentioned aspects of this video, along with visual features such as the avatar and the type of video presentation (whiteboard or "handwriting").

After analyzing the responses regarding Video 3, suggestions related to the use of background music in videos were identified, which had been mentioned before in relation to other videos. Additionally, the importance of choosing the type of video presentation, in this case, the whiteboard or "handwriting" style, was highlighted.

Finally, based on the analysis of participants' facial expressions during the fourth video, the predominant emotion identified was neutrality. The questionnaire responses are presented in Figure 13.

Figure 13 — Participants' responses to the questionnaire about Video 4
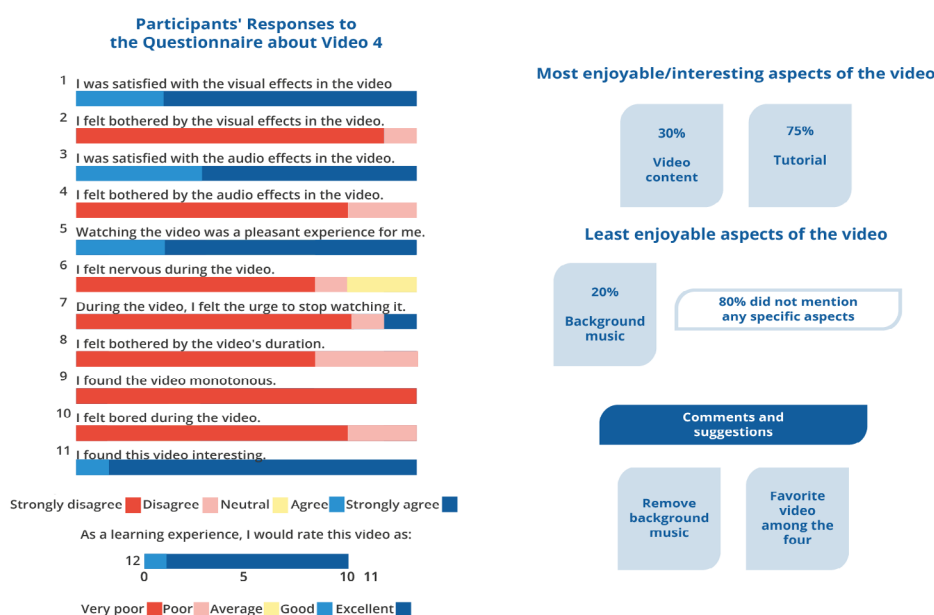
Source: Created by the authors..

Based on the participants' responses regarding the most enjoyable/ interesting aspects, it can be highlighted that the tutorial format and the content of the video were the most mentioned items by the participants. Some participant quotes include: "The step-by-step approach in the video has great didactics and helps in understanding how to use the platform" (Participant 4); "The explanations were excellent. I will definitely produce videos using this application" (Participant 8); "The content was enriching. I was not familiar with Animaker, and I found the presentation explanation to be very important. In addition to discussing the program, the tutorial made the video more engaging to watch" (Participant 1).

Regarding the less enjoyable aspects, it is important to note that approximately 80% of the participants did not mention any negative aspects, making this the video with the highest number of participants who had no negative points to mention. The only aspect mentioned was the background music used in the video. Among the items mentioned by participants as less enjoyable, the following stand out: "The background music" (Participant 2); "Actually, there is nothing unpleasant in this video, as it contributes a lot to the learning process and is presented clearly and cohesively" (Participant 4).

As it was an optional item, not all participants provided suggestions and comments about the video. The majority of received comments praised the video, with some even highlighting it as their favorite among the ones watched. There were also suggestions to remove the background music. It's important to note that such suggestions were observed in most of the videos, and the use of this audio effect was done to gauge participants' perception of it.

Among the received comments, the following can be highlighted: "The video was very interesting, not monotonous at all, as each frame was explained and exemplified by the instructor. I still believe that the background sound is unnecessary, as it can sometimes distract from the learning focus" (Participant 11); "It's my favorite video out of the four" (Participant 3).

In general terms, in Video 4, the same relationship between the topic and participants' interest can be observed, with several comments indicating that they will use the content for their personal and academic productions. Additionally, the fact that it was a tutorial video (screen capture) was mentioned as a positive aspect, making the video more dynamic and

interesting. The lack of exemplification was not mentioned in this video.

Based on the analysis of Video 4, suggestions were identified regarding the dynamic nature of the video, exemplification of the presented content, and narration by a human voice. After analyzing the participants' emotions and perceptions of each video, a comparative general analysis was conducted. An overall analysis of the emotions displayed by the users while watching the videos was performed, linking the predominant emotions observed in facial expressions with the perceptions noted in the questionnaire responses.

Figura 14 — Emotions and perceptions of each video
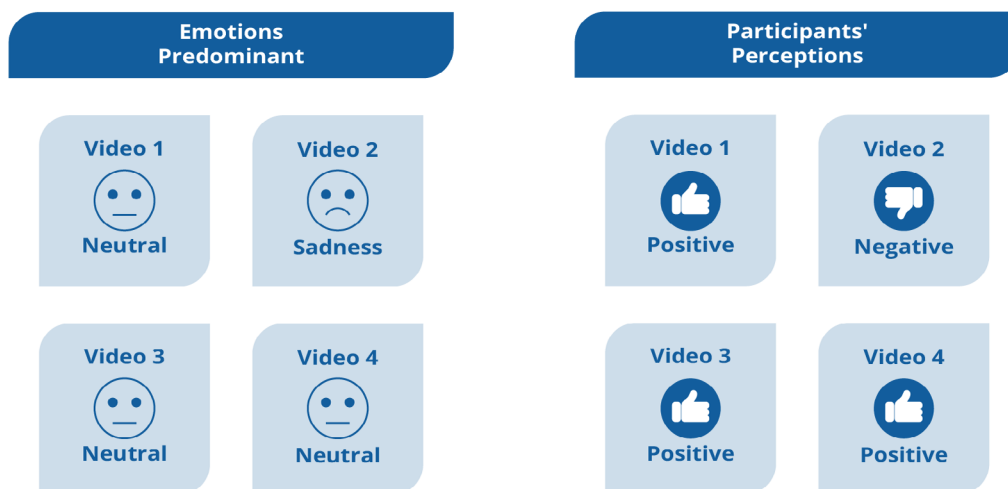
Source: Created by the authors.

Based on this, a differentiation can be observed between the information obtained from the emotions inferred from facial expressions and the questionnaire responses. In all the watched videos, the predominant emotion for the majority of participants was neutrality, meaning their facial expressions did not align with any of the basic emotions analyzed by the EZ-MMLA ToolKit. However, in terms of the perceptions expressed in the questionnaire responses, most participants demonstrated positive perceptions. It is important to note that the only video where a negative perception predominated, according to the questionnaire responses, was Video 2, which aligns with the predominant emotion observed in facial expressions, which was sadness.

The high acceptance of Video 4 can be attributed to its dynamic nature, practical presentation of the tool, and exemplification of its functionalities. Additionally, participants showed personal interest in the topic and the

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

23

presented tool. The more engaged participants are with the content, the more likely positive emotions will be evoked, stimulating cognitive functions and improving attention and memory (DA FONSECA, 2016).

In relation to the video with the least acceptance, the second video, it is observed that despite the participants recognizing the topic and content as interesting and important, they made several criticisms regarding the use of synthetic voice narration. These criticisms confirm what was found in the literature review about the difficulty most people have in absorbing content when narrated by a synthetic voice. The experience of listening to humans reading a text for learning tasks is superior to listening to synthetic voices reading the same text (HILLAIRE; INIESTO; RIENTIES, 2019). This perception is evidenced both by the results obtained from the analysis of participants' facial expressions and the questionnaire responses, as the video with synthetic voice narration is the only one where the predominant emotion, in general, was sadness.

The motivational factor had a direct impact on the perceptions reported by the participants, as according to the responses, the most well-received video among the four presented was the one that had a more dynamic nature and presented the video creation and editing tool, aspects that the majority of participants stated they would use in their academic or professional journey. Motivation to expand and acquire knowledge serves as a driving force for learning, creating favorable conditions for it to occur in the best possible way. Interest is a sensation or feeling that leads to a tendency to act towards something. Generally, students will be interested in what brings benefits to them (OLIVEIRA, 2008). Activities that spark the student's interest, regardless of age, are essential to ensure the effectiveness of learning and provoke positive emotions.

Based on the information obtained from the analysis of the results, important guidelines were developed for those who wish to develop educational videos, providing a practical contribution from this research. Based on the defined tips, a guide with suggestions for producing educational videos for MOOCs was created. The guide was designed and written in appropriate and easily understandable language, being concise and containing references. During the definition of tips and items, the use of illustrative images and clear and direct language was prioritized, presenting observations about each item, as well as aspects that require the reader's attention during video production. The guide can be accessed through the QR Code present in Figure 12.

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

24

Figure 15 — QR Code for accessing the produced guide

Source: Created by the authors.

The guide's contents were defined based on the research conducted in the literature review, in line with the analysis of the data obtained during the evaluation with the participants. The following are the contents of the guide.

- Introduction: This section provides an introduction to the guide and the research that guided its production, along with a brief presentation of the authors.

- Introduction: In this section, a brief introduction is given about the guide's topic.

- Classification of Educational Videos: This section presents the classifications of videos based on their purpose and design.

- Duration of Educational Videos: This section presents the ideal duration for videos, considering the attention span of users. It emphasizes the need to be cautious in not presenting the content too quickly and suggests dividing the content into parts if it is too extensive.

- Presentation Mode: This section covers aspects related to the mode of content presentation.

- Visual Effects in Educational Videos: This section presents the concepts of framing and shot types, along with different types of visual effects. It also emphasizes the use of examples with images and avatars.

- Audio Effects in Educational Videos: This section presents options for audio effects such as background music and narration. It highlights the low acceptance of synthetic voice narration and provides guidance on the careful use of background music.

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

25

- Subtitles in Educational Videos: This section discusses the possibility of using subtitles in videos, both to improve user comprehension and enhance video accessibility. It also introduces tools for automatically including subtitles and provides access to a step-by-step tutorial video on how to add subtitles using these tools.

- Tools for Video Creation and Editing: In this section, several tools for video creation and editing are presented, with a focus on online tools. Each tool is briefly introduced, and a tutorial video for their usage is provided.

- Supplementary Material: As supplementary material, a suggested MOOC course titled "Educational Videos: Everything You Need to Know" is recommended. This course contains the videos watched by the participants and the tutorial videos provided in the guide. Additionally, a book resulting from research conducted in 2014 by the Instituto Federal de Educação, Ciência e Tecnologia Sul Rio-grandense, titled "Parameters for the Production of Educational Videos", is suggested as supplementary material.

- References: This section includes all the references that guided the writing of the guide.

- Team: Lastly, the research team that conducted the study that led to the guide is introduced.

During the production of the guide, emphasis was placed on using attention-grabbing color images. Additionally, clear language was prioritized. The topics were approached through contextualization, providing tips for each item, as well as highlighting aspects that require the reader's attention during video production.

5. Final Conclusions

To investigate the topic, we sought to examine the studies conducted on the subject and understand the methodologies for inferring emotions, noting the limited number of research that specifically address emotions related to MOOC courses. During the data analysis, a discrepancy was identified between the data obtained from participants' facial expressions and the data obtained through questionnaire responses. It was found that

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

26

the predominant emotion during video viewing was neutrality, while the questionnaire responses varied depending on the presented video.

The discrepancy between the data obtained from the two data generation methods can be attributed to the main limitation of the method used to infer emotions from facial expressions, which only considered the basic emotions (joy, sadness, anger, fear, surprise, disgust), and assigned a neutral result when a facial expression did not fit into any of these emotions. This limitation arises because basic emotions are more exaggerated and considerably less evident in everyday situations.

By comparing the emotions of users in videos with different characteristics, it was possible to identify the most accepted and least accepted videos by users. The least accepted video was the one with synthetic voice narration, which aligns with the findings of the literature review, while the most accepted video was the tutorial video. This acceptance was also influenced by the personal and professional interest of the participants in the video's topic. These findings, along with the data obtained during the literature review, were crucial in formulating the suggestions presented.

After the analysis, it can be observed that in addition to the audiovisual characteristics of the videos, the presentation/development of content and motivational factors have a significant influence on the effectiveness of learning (OLIVEIRA, 2008). Activities and content that spark interest have a positive impact on students' emotions and learning effectiveness, while activities that generate negative emotions, even momentarily, can lead to temporary lapses in attention and concentration, requiring more time for the student to regain focus and acquire the necessary knowledge (DA FONSECA, 2016).

As a future perspective, there is a need for continuous updating of the produced guide due to technological advancements and the rapid growth of MOOCs. Additionally, the analysis of users' emotions and perceptions while watching videos was conducted with a small sample size, and it would be valuable to replicate the study with a larger number of participants, encompassing different profiles, to obtain a broader overview of preferences regarding this resource. It is important to note that the number of participants and the selection method were defined for exploratory analysis purposes, with the prospect of future research involving a larger number of participants from varied profiles and addressing other topics. Future research can also include investigating a method for inferring emotions that considers

secondary emotions, which may provide more accurate results.

Finally, it is essential to emphasize the importance of conducting further research on this subject in the Brazilian context. The growth of courses in this format is irreversible, and both formal and informal education have benefited from this model to provide access to knowledge in a wide range of areas. There is a vast and promising field for work in this regard, and it is crucial for educational science to embrace and develop investigations that can support and enhance the strategies adopted in the production of audiovisual content, which are often created intuitively.

REFERÊNCIAS

BUSIN, Y. **Análise de emoções em expressões faciais:** veracidade das emoções e rastreio ocular. 2013. Dissertação (Mestrado em Psicologia) — Universidade Presbiteriana Mackenzie, São Paulo, 2013.

DA FONSECA, V. Importância das emoções na aprendizagem: uma abordagem neuropsicopedagógica. **Revista Psicopedagogia**, São Paulo, v. 33, n. 102, p. 365–384, 2016.

DAMÁSIO, A. **O mistério da consciência**: do corpo e das emoções do conhecimento de si. São Paulo: Companhia das Letras, 2000.

DENZIN, N.; LINCOLN, Y. **O planejamento da pesquisa qualitativa**: teorias e abordagens. 2 ed. Porto Alegre: Artmed, 2006.

FARINA, M.; PEREZ, C.; BASTOS, D. **Psicodinâmica das cores em comunicação**. São Paulo: Edgard Blucher, 2013.

FONTANA, M. V. L.; LEFFA, V. J. MOOCs para o ensino de línguas: um estudo em call desde uma perspectiva conectivista. **Alfa: Revista de Linguística**, São José do Rio Preto, v. 62, n. 1, p. 75–89, 2018.

FREITAG, R. M. K. Amostras sociolinguísticas: probabilísticas ou por conveniência? **Revista de estudos da linguagem**, Belo Horizonte, v. 26, n. 2, p. 667-686, 2018.

HASSAN, J.; LEONG, J.; SCHNEIDER, B. Multimodal Data Collection Made Easy: The EZ-MMLA Toolkit. LAK21: 11th International Learning

Analytics and Knowledge Conference. **Association for Computing Machinery**, Nova York, p. 579–585, 2021.

HILLAIRE, G.; INIESTO, F.; RIENTIES, B. Humanising Text-to-Speech Through Emotional Expression in Online Courses. **Journal of Interactive Media in Education**, n. 1, p. 12, 2019.

LAASER, W.; TOLOZA, E. A. The Changing Role of the Educational Video in Higher Distance Education. **The International Review of Research in Open and Distributed Learning**, v. 18, n. 2, 2017.

LUZ, V. V.; PIOVESAN, S.; IRALA, V. O uso de computação afetiva em moocs: um mapeamento sistemático. **DESAFIOS - Revista Interdisciplinar da Universidade Federal do Tocantins**, Tocantins, v. 1, n. 1, p. 236-248, 2023. Disponível em: https://sistemas.uft.edu.br/periodicos/index.php/desafios/article/view/13656. Acesso em: 13 abr. 2023.

OLIVEIRA, J. B. A. **Aprender e ensinar**: a aprendizagem do ensino. 9. ed. Belo Horizonte: Instituto Alfa e Beto, 2008.

PEKRUN, R. *et al*. Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). **Contemporary Educational Psychology**, v. 36, n. 1, p. 36–48, 2011.

PIAGET, J.; INHELDER, B. **A psicologia da criança**. 16. ed. São Paulo: Bertrand Brasil, 2003.

SANTOS, E. Educação online para além da EaD: um fenômeno da cibercultura. *In*: CONGRESSO INTERNACIONAL GALEGO-PORTUGUÊS DE PSICOPEDAGOGIA, 10., 2009, Braga. **Anais** [...] Braga: Universidade do Minho, 2009. p. 5658-5671.

SICILIANI, I. D. S. **Elaboração, aplicação e avaliação de um Curso Online Aberto e Massivo (MOOC) interdisciplinar entre Física e Matemática**. 2016. Dissertação (Mestrado profissional em Ensino de Física) — Universidade Federal de Santa Catarina, Florianópolis, 2016.

SILVA, J. M. C. da; ACCORSI, M. I.; MUNHOZ, E. M. B. O impacto do distanciamento social nos cursos abertos e massivos sob a perspectiva da procura e oferta. *In*: CONGRESSO BRASILEIRO DE ENSINO SUPERIOR A DIST NCIA - ESUD, 17., 2020, Goiânia. **Anais eletrônicos** [...]. Goiânia: Cegraf UFG, 2020.

STANLEY, D.; ZHANG, J. Do Student-Produced Videos Enhance Engagement and Learning in the Online Environment. **Online Learning**, v. 22, n. 2, 2018.

TRIVIÑOS, A. N. S. **Introdução à pesquisa em ciências sociais**: a pesquisa qualitativa em educação. São Paulo: Atlas, 2011.

VENTURA, M. M. O estudo de caso como modalidade de pesquisa. **Revista SoCERJ**, Rio de Janeiro, v. 20, n. 5, p. 383–386, 2007.

YADEGARIDEHKORDI, E. *et al*. Affective computing in education: A systematic review and future research. **Computers & Education**, v. 142, p. 103649, 2019.

ZHENG, S. *et al*. Understanding student motivation, behaviors, and perceptions in MOOCs. In: ACM INTERNATIONAL CONFERENCE ON COMPUTER-SUPPORTED COOPERATIVE WORK AND SOCIAL COMPUTING, 18., 2015, New York. **Anais** [...]. New York: Association for Computing Machinery, 2015. p. 1882–1895.

Emotions of Users in Front of Educational Videos: a case study focused on the construction of guidelines for the production of moocs

30