Unpacking Academic Emotions in Asynchronous Video-based Learning: Focusing on Korean Learners’ Affective Experiences

Jihyang Lee1 · Hyo-Jeong So1 · Seunghye Ha1 · Eunyoung Kim1 · Kyudong Park2

Accepted: 25 March 2021 / Published online: 15 April 2021
© De La Salle University 2021

Abstract Recognizing learners’ motivational problems in online remote learning, this study examined the patterns of learner’s academic emotions and facial expressions detected in asynchronous video-based learning. The research questions of this are: (1) What are the patterns between facial expressions and negative emotions (i.e., boredom, confusion, and frustration) that learners experienced during online video-based learning? (2) What are the learner’s overall perceptions about their emotions during online video-based learning? This study was conducted with 26 Korean adult learners who took an online video lecture via an asynchronous self-directed mode. Their facial expressions during online learning were recorded and analyzed for detecting affective states. Two judges trained in Ekman’s Facial Action Coding System (FACS) analyzed 210 scenes segmented from the recorded video data using a retrospective affect judgment procedure. The analysis identified 101 incidences of boredom, 82 confusions, and 62 frustrations. Moreover, using Spearman’s rank correlation coefficients and hierarchical clustering, we identified some patterns of significant relationships between facial action units and affective states. The interview data reveal that the potential reasons for negative emotions include lack of concentration, uninteresting content, and one-way content delivery method. Given the massive transition to online video-based learning during and after COVID-19, the results of this study can provide implications concerning how to detect and intervene in potential issues associated with learner’s motivation and affective states.

Keywords Academic emotions · Affect detection · Online video-based learning · Facial expressions

Introduction

The COVID-19 pandemic has made radical changes in our lives across the globe, especially the way we teach and learn. According to the April 2020 statistics, 191 countries closed their schools due to the COVID-19 pandemic, which affected approximately 91% of students enrolled (UNESCO Institute of Statistics, 2020). Due to the school closure, the main teaching and learning environments have shifted from the traditional face-to-face mode to the online distance mode (Reimers & Schleich, 2020). The fact that societies are experimenting with nationwide distance learning at all levels and types of education is unprecedented in history. The current situation of online learning and school closure caused by the COVID-19 posed both challenges and opportunities. UNESCO (2020) describes the recent transition to online learning as “large real-life laboratories for experimenting with nationwide and system-wide distance learning practices and the possible models relevant to the future of education” (p. 4). Schools that adhered to the traditional classroom learning had to make a transition to fully remote learning methods via online platforms or a blended learning structure that combines face-to-face classroom and online learning. The combination of traditional classroom learning and online learning is predicted to be more pervasively adopted in various educational settings even after the COVID-19
pandemic (Bonk, 2020). This implies the necessity to help students to be more prepared in online learning environments.

The openness and flexibility of online learning require a learners’ high level of monitoring the learning process and maintaining motivation for a sustained duration (Onah et al., 2014). The existing literature, however, suggests that even adult and well-educated learners tend to experience difficulty in sustaining motivation in online learning environments (Anders, 2015). Indeed, meaningful experiences in online learning environments demand learners’ self-directedness and regulation in a physically and psychologically separated space without direct contact with instructors and peers. The self-directed and isolated nature of online learning tends to lead learners to experience various states of emotions, including some negative emotions that may inhibit their learning process and outcomes (Marchand & Gutierrez, 2012). How to mitigate such motivational and psychological barriers that learners experience is a critical challenge in the design of online learning environments.

Academic emotions are pertinent to examine learners’ affective states that occur during online learning. As a psychological construct, academic emotions refer to emotions experienced by students in various academic settings, which can be classified based on the positive (e.g., joy, hope, and pride) or negative (e.g., anxiety, boredom, and hopelessness) valence (Pekrun, 2016). Although academic emotions play a critical role in overall learning processes, little is known about learners’ emotions in online learning situations (Ashwin & Guddeti, 2020). Methodological challenges may explain the paucity of this line of research. The first challenge involves the difficulty of classifying academic emotions that occur during online learning. Several research studies on human emotion are based on Ekman’s pioneering work on the measurement of facial movements for detecting six basic emotions (i.e., sadness, happiness, anger, fear, disgust, and surprise) that were developed about a half-century ago (Ekman & Friesen, 1978). These basic emotions were found to easily appear and be detectable in a short time-frame of interaction, such as watching a commercial video. However, they have little relevance to most learning contexts that require more intensive interaction and longer engagement. Beyond basic emotions, educational researchers have found that a distinct set of emotions, such as boredom, frustration, flow, eureka, and confusion, are more prominent and important to examine in learning contexts (Pekrun et al., 2002). However, few theoretical and methodological frameworks have detected and classified this range of academic emotions occurring in the context of online learning.

The next challenge is related to what extent emotions are universal or culture-specific. Although facial expressions of emotions have been considered universal (Ekman, 1982), an increasing number of researchers have suggested the need to examine the culture-specific nature of expressing emotions. For instance, D’Mello et al. (2009) argue that “emotions are notoriously difficult to study due to inherent variations across personalities, experience, age, gender, culture, and time” (p. 167). In a similar vein, Baker et al. (2019) contend that macro-theories that tend to characterize universal patterns are limited in explaining cultural and regional differences, even in affect detection. Overall, these methodological challenges imply the need to accumulate more research findings that better reflect the unique characteristics of online learning environments with a more diverse range of learners’ cultural profiles.

Given that learner’s emotions will become increasingly important in the context of self-directed and isolated online learning, learners’ emotions must be analyzed in more depth and emotion-based scaffolding must be provided in the design of online learning environments. With this backdrop, this study investigates Korean learners’ emotions experienced in online learning. This study put two foci on the research questions among various issues on academic emotions and online learning. First, this study focuses on video-based learning delivered in an asynchronous mode, given that this mode of online learning has been reported as the most commonly used in remote education during the COVID-19 (BRIC View, 2020). Second, this study particularly focuses on detecting three types of negative emotions, namely, boredom, confusion, and frustration, which have been reported as important affective states likely to cause off-task behaviors and low motivation for learning (e.g., Baker et al., 2010; D’Mello & Graesser, 2009). Ultimately, this study aims to draw implications that can inform the design of online video-based learning that is affect-sensitive and provides affective interventions to help online learners experiencing motivational difficulties. The following research questions guided this study: (1) What are the patterns between facial expressions and negative emotions (i.e., boredom, confusion, and frustration) that learners experienced during online video-based learning? (2) What are the learner’s overall perceptions about their emotions during online video-based learning?

**Theoretical Background**

**The Complexity of Affect and Cognition**

The definition of affect varies among scholars, with no single agreed-upon definition. Various terms such as affect, emotion, mood, and feeling have been used to indicate or
explain one’s emotional experiences. Depending on the context of the study, affect and emotion can have the same meanings and be used interchangeably (Ainley, 2006). The existing literature suggests an inextricable relationship between cognition and emotion in learning situations (Pekrun, 2006). Learning is not a purely cognitive process but a complex process bound to emotional states. Notably, the recent studies went beyond dichotomizing cognition and emotions but adopted an integrative perspective that acknowledges the intricate relationship between cognition and emotion in learning situations (D’Mello et al., 2009). The integrative perspective is supported by evidence of brain science, suggesting that the distinction between emotional and cognitive brain is fuzzy; the brain territories associated with cognition are closely connected to the emotional processing (Do & Schallert, 2004). In addition, Meyer and Turner (2002) argue that learner’s problem-solving process in complex tasks is driven by an emotional response; hence, the role of emotions should not be overlooked when analyzing learners’ goal achievement, self-efficacy, and learning strategies.

Emotions can be classified in multiple ways. In general, emotions that learners experience in learning situations are largely divided into positive and negative emotions. Positive emotions include enjoyment, confidence, satisfaction, hope, and joy, whereas negative emotions include boredom, frustration, anxiety, anger, sadness, confusion, and shame (Pekrun et al., 2002). Based on the control-value theory, Pekrun et al. (2007) further suggests the two types of academic emotions: activity-related and outcome-related emotions. Activity-related emotions are experienced with ongoing learning activities such as the enjoyment of discovering new knowledge, frustration caused by difficult tasks, and boredom experienced in listening to lectures. Meanwhile, outcome-related emotions are emotions attached to the achievement outcomes such as success and failure (e.g., joy, shame, and pride).

**Academic Emotions and Motivation**

The literature increasingly suggests that emotions not only influence behaviors and cognition but also play an integrative role in mediating and sustaining motivation. Motivation is a desire and willingness to act (Pintrich & Shunk, 2002). Academic emotions and motivation have bidirectional relationships. When learners feel positive emotions, such as the joy of discovering new knowledge, they are motivated to pursue a goal. Conversely, learners who feel shamed or hopeless about their performance outcomes (e.g., test scores) tend to lose motivation for learning.

Interest as an immediate action to a new learning task is an affective state central to the motivation for learning. Ainley (2006) suggests that “the feeling of interest involves positive activation (affect), directed attention and impulses to action” (pp. 398–399). When learners have a high interest in learning tasks, they tend to be more concentrated and invest more effort in learning. In a study by Ainley et al. (2005), young adolescents were given reading materials taken from popular science and culture magazines on an interactive computer screen. They were then asked to indicate their feeling with text reading with icons (sad, interested, embarrassed, sorry, neutral, angry, surprised, happy, scared, bored, and disgusted). Ainley et al. (2005) found that the emotional response indicated as “interested” has strong intensity with the decision to continue or stop reading the text.

Anxiety is another important construct indicating the complexity of affect, motivation, and cognition. Among various academic emotions, anxiety has been extensively researched as an emotion that has a great influence on motivation and achievement. Anxiety is an emotion holding both positive and negative influences. High anxiety, such as test anxiety, tends to influence learners to lose internal motivation for learning. Simultaneously, anxiety can induce learners’ external motivation to invest more efforts to avoid failures (Pekrun, 2006). Some levels of anxiety, if not excessive, can be beneficial for sustaining motivation and may lead to positive achievement outcomes.

**Affective Experiences in Online Learning**

The existing body of knowledge about academic emotions has been mainly built on the research findings drawn from face-to-face learning situations. Academic emotions in face-to-face learning and online learning are rather different. Learners experience more diverse emotions in online learning than in face-to-face learning environments (Moneta & Kekkonen-Moneta, 2007). The difficulty in regulating and sustaining motivation occurs more frequently in online learning than in face-to-face learning due to the lack of human interaction, social presence, and emotional support (Hartnett, 2016). Technical aspects of online learning may affect students’ emotions, causing anxiety and frustration.

Moreover, the sense of isolation that students are faced with during online learning may cause feelings of fear and isolation (Martin & Ford, 2018). Fewer sources of reinforcement and prompts from instructors or peers to keep learners on-task also explain the differences in affective experiences between online and offline learning environments (Delen et al., 2014). Concerning emotions in online learning and potential causes, O’Regan (2003) conducted a
qualitative study to examine the types of emotions expressed in the online learning process and factors that induced such emotions, using data from 11 college students who took an e-learning course. The students who participated in the interview shared how various emotional experiences (e.g., frustration, fear, shame, and interest) triggered in specific situations facilitated or inhibited their learning. This study suggests that emotional experiences should not be overlooked in the design of online learning environments.

Recently, academic emotions related to online learning have been extensively examined in the context of Massive Online Open Courses (MOOCs). High dropout and low retention rates have been constant issues in MOOCs. There are increasing interests in understanding emotional learning experiences to explain these issues of dropout and retention. Dillon et al. (2016) reported that MOOC learners experienced various emotions based on the large-scale analysis of self-reported data. The study found that hope, enjoyment, and contentment were the most frequently reported emotions and that positive correlations exist between dropout and negative emotions, such as frustration, confusion, and anxiety. Xing et al. (2019) contend that learners’ affective experiences in MOOCs are complex and multi-dimensional, beyond the dichotomic descriptions of positive and negative emotions. They examined MOOC learners’ academic emotions expressed in online forums. Through the automatic detection modeling of emotion-related keywords in forum postings, they found that students’ exposure to negative emotions influenced their commitment to continue participating in MOOCs. Emotions were also found to be contagious in that students’ retention was more affected by the emotional exposures to their peers’ postings than the emotions expressed in their postings.

### Emotions and Facial Expressions

Methodologically, detecting emotional states is challenging and complex. Since the seminal work by Ekman’s model of basic emotions, facial expressions have been used to derive academic emotions. The main assumption is that certain facial expressions in learning situations are related to specific affective states (D’Mello et al., 2009). The Facial Action Coding System (FACS) by Ekman and Friesen (1978) is the most frequently used mechanism for analyzing facial expressions. As shown in Table 1, FACS specifies how certain facial Action Units (AU) are coded based on the muscular actions to produce them.

In a comprehensive review of research studies that detected engagement in the context of online learning, Dewan et al. (2019) found that the existing research methods fall into three main categories, namely automatic, semi-automatic and manual, and that detecting engagement facial expressions through the computer-vision methods is particularly promising in terms of accuracy and efficiency of dealing with engagement-related data. Several studies used facial expressions to examine three emotional states (i.e., boredom, confusion, and frustration), which are the focus of this research. Table 2 summarizes the list of facial AUs reported in the previous studies that are associated with boredom, confusion, and frustration. Certain AUs commonly appeared across studies, such as AUs 4 and 7 for confusion.

McDaniel et al. (2007) used the FACS to explore the relationship between learning activities and affective states when learners solved complex problems in AutoTutor, an intelligent tutoring system about Newtonian physics and computer literacy. The study used observation data and emote-aloud protocol to examine the affective states experienced by college students during learning with AutoTutor. They found that confusion was the most noticeable emotion from multiple facial features.

| Table 1 | Sample AUs in the facial action coding system by Ekman and Friesen (1978) |
|---------|---------------------------------------------------------------|
| **AUs** | **Description** | **Facial muscle** |
| AU1     | Inner brow raiser   | Frontalis, pars medialis |
| AU2     | Outer brow raiser  | Frontalis, pars lateralis |
| AU4     | Brow lowerer       | Depressor glabellae, depressor supercili, corrugator |
| AU7     | Lid tightener      | Orbicularis oculi, pars palpebralis |
| AU12    | Lip corner puller  | Zygomaticus major |
| AU14    | Dimpler            | Buccinator |
| AU41    | Lid droop          | Relaxation of levator palpebrae superioris |
| AU43    | Eyes closed table  | Relaxation of levator palpebrae superioris |
| AU45    | Blink              | Relaxation of levator palpebrae and contraction of orbicularis oculi, pars palpebralis |

*Source: IMOTIONS [https://imotions.com/blog/facial-action-coding-system/] (2020)*
Confusion was manifested by the presence of a lowered brow (AU4), coupled with the tightening eyelids (AU7), and a lack of a lip corner puller (AU12). Frustration had a significant correlation with AU12 (lip corner puller). They concluded that boredom was not easily distinguishable from facial features because AUs with statistical significance showed negative relationships. The difficulty of detecting boredom was speculated to be due to boredom’s similarity with an expressionless face, where people do not show noticeable movements on the face and stay neutral.

D'Mello et al. (2009) further examined learners’ affective states during learning with AutoTutor. This time, they used multiple methods in online and offline contexts: self-report, human judges, and automated detection. The study revealed high correlations between boredom and eyes closed (AU43). Confusion and frustration had more AUs with significant relationships. Confusion showed high associations with the expression of brow lowerer (AU4), lid tightener (AU7), and lip corner puller (AU12). Frustration was detectable from the presence of inner brow raiser (AU1), outer brow raiser (AU2), lip corner puller (AU12), and dimpler (AU14). Overall, they concluded that although confusion and frustration are easily detectable from facial expressions, boredom is more difficult to relate to specific facial features.

In complex learning situations, confusion is likely to lead to frustration if disequilibrium in cognitive understanding is not resolved (D’Mello & Graesser, 2012). Under the situation of learning a new concept, for instance, when a learner cannot reconcile their feeling of confusion (e.g., feeling stuck), they may feel frustration. Detecting indicators of frustration, therefore, has been critical to designing affect-sensitive learning environments (D’Mello et al., 2009). Grafsgaard et al. (2013a, 2013b) conducted a series of research to automatically detect learners’ frustration from facial movements. Grafsgaard et al. (2013a) used the program called Computer Expression Recognition Toolbox (CERT), which produces intensity values for different facial actions in FACS (Littlewort et al., 2011). They found that brow lowerer (AU4) and mouth dimpler (AU14) are positively associated with frustration. In another study, Grafsgaard et al. (2013b) used both the manual coding of FACS and CERT. The participants were college students in an introductory engineering course who used a web-based tutoring program called JavaTutor. They found that the intensity of brow lowerer (AU4) is associated with the increase in frustration. They described the combination of AU1, AU2, and AU4 as the prototypical expression of “fear brow” that often occurs during anxiety. Bosch et al. (2014) also used CERT to automatically detect facial AUs in the context of learning Python programming language. They found the occurrences of five affective states, namely, boredom, confusion, flow/engagement, frustration, and neutral. The analysis of facial expressions detected from novice programmers indicates that confusion is easily detectable by AU 1 (Inner brow raiser) or the combination of AU 1 and brow lowerer (AU4).

Overall, although some AUs were commonly found across multiple studies, the mixed findings and interpretation indicate the need for additional research for clear classification and detection of emotions in academic situations. The gap in the existing literature can be found in two aspects. First, most research studies that used facial expressions to detect academic emotions were conducted with intelligent tutoring systems or in the learning situation that demands high level of cognitive activities such as problem-solving and programming. It is still unclear to what extent the findings from such cognitively demanding situations could be observed in the situation of asynchronous video-based learning where learners usually listen to video lectures and may not have specific achievement goals involving cognitively demanding tasks (e.g., taking a quiz) for motivation. Second, most existing studies have been conducted with learners from Western cultures. Traditionally, facial expressions were considered universal (Ekman, 1982). Recent studies, however, suggest that facial expressions can be influenced by cultural factors, especially socially constructed egos, cultural practices, and

Table 2  Facial AUs associated with boredom, confusion, and frustration

| Emotion      | Facial action units (AUs) | References                     |
|--------------|--------------------------|--------------------------------|
| Boredom      | AU4*, AU7*, AU12*        | McDaniel et al. (2007)         |
|              | AU43                     | D’Mello et al. (2009)          |
| Confusion    | AU4, AU7, AU12*          | McDaniel et al. (2007)         |
|              | AU4, AU7, AU12           | D’Mello et al. (2009)          |
|              | AU1, AU4                 | Bosch et al. (2014)            |
| Frustration  | AU12                     | McDaniel et al. (2007)         |
|              | AU1, AU2, AU12, AU14     | D’Mello et al. (2009)          |
|              | AU4, AU14                | Grafsgaard et al. (2013a, 2013b) |

*Negative correlation (−)
race (Baker et al., 2019; Mandal & Ambady, 2004). Hence, whether or not the existing classifications of detecting emotions can be applied to Asian learners must be examined, as these have been mainly examined with Western learners.

Method

Research Participant and Context

The participants of this study were 26 Korean adult learners between 20 and 35 years old who were recruited through convenience sampling. In recruiting the participants, we considered individual variables such as gender, academic majors, and prior knowledge. The participants include 12 men and 14 women, coming from various academic majors such as science and engineering, arts, sports, and humanities. They did not have any prior experience of taking the video lecture used in this study.

The experiment was conducted remotely in the participants’ homes to simulate authentic online learning situations during the COVID-19 pandemic, rather than in a controlled lab setting where researchers were present. The presence of observers and cameras has been documented to have a possible effect on participants’ behaviors when they are aware of being directly observed (Jewitt, 2012). Therefore, we chose an indirect observation method in a natural setting where the participants were less aware of the observer effect and felt comfortable about performing the given learning activity.

Video Material

For the experiment, the participants were asked to watch an online video lecture about “Human Subject Research,” as shown in Fig. 1. This video is 72 min long and covers the basic research issues and methods that researchers should adhere to before and during conducting research studies involving human subjects. Specific topics covered in the video lecture include the scientific and ethical validity of research proposals, procedures for obtaining legal consent from research subjects, measures for protecting personal information, bioethics, and safety procedures. Throughout the video, the female instructor appears on the left side of the screen and explains key concepts in a lecture mode with textual and visual aids.

The selection of this video lecture was done with two main considerations. First, we chose a rather long video clip for this study to detect and analyze negative emotions (i.e., boredom, confusion, and frustration). The selection of a long video clip was intentional because short video clips do not induce sufficient occurrences of emotional states. Longer video clips are likely to elicit a more diverse range of learners’ emotions regardless of whether they are positive or negative. Second, we chose video content that the participants had not taken before to control the effect of prior knowledge. The level of prior knowledge has been well documented to affect learning processes and outcomes (e.g., Shapiro, 2004; Taub et al., 2014). When recruiting the participants, we asked them whether they had taken any lectures on human subject research, and those who had not taken such courses were selected as the participants for the study.

Research Process

The overall research process includes four steps. First, the researchers explained the purpose of the experiment through a face-to-face or online meeting with each subject and obtained the consent form for the research participation. In addition, the researcher provided the participants with detailed instructions about the experimental procedure. Second, each subject took the video lecture for 72 min, and the whole learning process was recorded in

Fig. 1 Video lecture on “Human Subject Research.” (The slide explains the groups of human subjects who need special attention such as students–teachers, those in power relations, and socially disadvantaged groups)

Fig. 2 Video screen with the participant’s face at the corner in Zoom
Zoom. The subjects were asked to show their face in Zoom to record their facial expressions remotely. Figure 2 presents a screen showing both the video scene of the lecture that the learner was watching and the learner’s face. With this recording method, the learner’s learning process and facial expressions in specific scenes could be analyzed simultaneously. Third, the subject sent the recorded video back to the researcher via e-mail. Fourth, the researcher conducted an interview with each subject for 30 min on the following day of taking the video lecture.

Data Collection and Analysis

Data on Facial Expressions and the Affective States

The first research question was to examine the patterns between facial expressions and negative emotions (i.e., boredom, confusion, and frustration) that learners experienced during online video-based learning. To examine this research question, we collected data through the video recording of the whole online learning process, and the judges conducted a retrospective affect judgment procedure in the following steps. The manual analysis by human judges was chosen for the first research question because there were no established programs that automatically detect academic emotions that this study focuses on. Two human judges (also the authors of this paper) had experiences to code facial expressions and emotions through previous research projects prior to participating in this study.

First, we used a bottom-up approach to identify segments wherein notable facial expressions occurred from all the video recordings. Identified segments were divided into a unit of a scene based on the event in which facial expressions occurred. A scene number was assigned to each unit. A total of 498 scenes were extracted through this process.

Second, among the facial expressions extracted in the first step, we further identified the scenes that appeared to contain negative emotions (boredom, frustration, and confusion), considering the retention time. Typically, when certain facial expressions last for at least 10 to 20 s, the expressions carry a certain meaning (Balconi & Bortolotti, 2013). Accordingly, this study analyzed scenes where facial expressions were maintained for more than 10 s, which helped to exclude significant portions of transitory states of emotions. Moreover, it should be noted that the focus of the present study was on detecting three types of negative emotions; hence, other emotional states (e.g., positive emotions such as flow and engagement) were not considered in this process. The second step led to confirm a final set of 210 scenes for data analysis.

Third, we used a top-down approach to assign pre-determined codes to corresponding facial expressions and affective states. Two judges trained in Ekman’s FACS (Ekman & Friesen, 1978) manually analyzed the collected data. Table 3 presents the FACS codes used in this study. Two judges independently coded emotions and corresponding facial expressions in each unit of the scene. We also used the participant interview (explained in the next section) as a secondary source to gain situational information, as recommended in other research studies (e.g., D’Mello et al., 2009). When disagreement in coding occurred between the two judges, the final decision was made through their discussion and agreement. The inter-rater reliability measured in Cohen’s Kappa was 0.743. In general, Kappa scores greater than 0.75 indicate an acceptable agreement between coders.

Fourth, we conducted both correlation analysis and association rule mining techniques to examine significant patterns between facial AUs and affective states. As discussed earlier, developing an automated system detecting learners’ affective states has been a huge challenge due to the fuzziness of emotional data. Information on how certain facial expressions are significantly correlated to certain affective states is essential in building automated computing systems. Similar to D’Mello et al. (2009), we first conducted a correlation analysis using Spearman’s rank correlation coefficient to identify statistically significant associations between facial AUs and affective states in the collected data. Patterns with significant correlations were identified by using the threshold value of the representative metric of the association rule mining analysis (support is 5% or higher; confidence is 50% or higher). Next, the clustering technique was used to visualize and group the representative patterns of learners’ emotions. Facial AUs and affective states for each scene were mapped into two dimensions using principal component analysis. Clusters were classified by hierarchical clustering, which is an algorithm to group similar elements into groups based on the hierarchical structure of individual elements. This technique has the advantage of identifying the appropriate number of clusters through a more informative dendrogram than non-hierarchical methods, such as K-means.

Data on Learners’ Perceptions about Emotional Experiences

The second research question was to examine the learners’ overall perceptions about their emotions experienced during online video-based learning. This was done qualitatively through a video-stimulated recall interview that is “a research technique in which subjects view a video sequence of their behavior and are then invited to reflect on their decision-making processes during the videoed event”
Table 3  Action units in FACS and codes for facial expressions used in this study

| Facial action unit (au) | Description         | Example         |
|-------------------------|----------------------|-----------------|
| AU1                     | EY inner brow raiser | ![Image](image1.png) |
| AU4                     | EY brow lowerer      | ![Image](image2.png) |
| AU7                     | EY lid tighten       | ![Image](image3.png) |
| AU12                    | MO lip corner puller | ![Image](image4.png) |
| AU14                    | MO dimpler           | ![Image](image5.png) |
| AU41                    | EY lid droop         | ![Image](image6.png) |
| AU43                    | EY eyes closed       | ![Image](image7.png) |
| AU45                    | EY blink             | ![Image](image8.png) |

*EY eyes, MO Mouth*
We viewed the video-stimulated recall interview as a relevant method because it allows collecting information on learners’ emotions about specific moments or events.

The two judges who coded the facial expressions watched the recorded video submitted by the subject and identified the areas for interview questions (e.g., moments of showing negative emotions) by scene numbers. They then prepared the video clips of the selected areas as the materials for simulating the subject’s recall. The interviews were conducted via Zoom for about 30 min on the next day after the subject watched the video lecture. For an effective stimulated recall interview, the researcher asked the participant to view the video clips and asked the interview questions. The interviews were semi-structured to allow the participants to freely express their emotional experiences and to stimulate their introspective thinking process. The following interview questions were crafted to explore whether the learners had any difficulty during the online video-based learning: “Do you usually have any problems or difficulties when watching the video lecture? If so, why?” and “How much did you feel concentrated in the overall learning process?”.

All interview sessions were video-recorded and transcribed for analysis. After organizing the transcripts for each interview question, common themes were identified by a constant comparison method. The researcher read the transcripts repeatedly to identify themes indicating the learner’s emotions, behavioral conditions, and consequences. A particular focus was given to examine potential factors that affected the learners’ emotions related to boredom, confusion, and frustration during the learning process.

**Results**

**Patterns of Facial Expressions and the Affective States**

In a total of 210 scenes, we identified 101 incidences of boredom, 82 confusions, and 62 frustrations. Since the dual codes were allowed, the total number of affective states exceeds the total number of scenes. During the analysis process, two AUs were removed due to the low observation count (AU7) and the low inter-rater reliability (AU43). As summarized in Table 4, we found that certain AUs had significantly positive correlations with three affective states examined in this study: (a) Boredom: AU14, AU41, and AU45; (b) Confusion: AU4 and AU12; and (c) Frustration:

| Facial action unit (AUs) | N   | Kappa score | Boredom (N = 101) | Confusion (N = 82) | Frustration (N = 62) |
|-------------------------|-----|-------------|--------------------|--------------------|---------------------|
| AU1 EY inner brow raiser| 28  | 0.576       | -0.32*             |                    | 0.45*               |
| AU4 EY brow lower       | 22  | 0.805       | 0.20*              | -0.22*             |                     |
| AU12 MO lip corner puller| 13  | 0.581       | 0.24*              | -0.17*             |                     |
| AU14 MO dimpler         | 90  | 0.694       | 0.15*              |                    |                     |
| AU41 EY lid droop       | 12  | 1.000       | 0.21*              | -0.15*             | -0.16*              |
| AU45 EY blink           | 29  | 0.670       | 0.25*              |                    | -0.23*              |

*All listed correlations are statistically significant at p < 0.05

| Facial action unit (AU) | Affective state |
|-------------------------|-----------------|
| AU1 EY inner brow raiser| + *             |
| AU4 EY brow lower       | + *             |
| AU12 MO lip corner puller| + *         |
| AU14 MO dimpler         | + *             |
| AU41 EY lid droop       | + *             |
| AU45 EY blink           | + *             |

*All listed rules are selected by the criteria (Support > 5%, Confidence > 50%)
AU1 \((p < 0.05)\). The state of boredom appears to be detectable by a dimpler in the mouth area (AU 14, \(r = 0.15\), drooping eyelids (AU 41, \(r = 0.21\)), and blinking eyes (AU45, \(r = 0.25\)). A lowered brow (AU 4, \(r = 0.20\)) coupled with pulling a lip corner (AU 12, \(r = 0.24\)) appears to be significantly associated with the expression of confusion. The analysis indicates that frustration has only one facial action with a significantly positive correlation, which is raising an inner brow (AU 1, \(r = 0.45\)).

Table 5 shows the results of associating rule mining, excluding AU7 and AU43, due to low incidence and Kappa scores. The results derived from the correlation analysis and the association rule mining technique are confirmed to be matched, indicating the strong patterns observed for these facial expressions and the affective states.

From the clustering analysis, 210 scenes were classified into four clusters (Fig. 3). Overall, we found that a dimpler (AU14) is a good indicator to group the states of boredom, confusion, and frustration. Looking at the characteristics of each cluster, in Cluster 1, AU14 was mainly connected with the emotions of boredom. Cluster 2 showed the boredom state, but AU14 was not observed. In Cluster 3, AU14 was found to be connected by feelings of confusion and frustration. Cluster 4 showed emotions of confusion and frustration, but AU14 was not observed.

Emotional States Experienced by Participants

Through the video-stimulated recall interview, we intended to understand the overall emotional experiences that the participants felt while watching the online video lecture (RQ2). Two common themes were drawn from the analysis of the interview data: (a) emotions about the learning content and (b) emotions related to the way the learning content was delivered.

First, the participants shared that sustaining their motivation was difficult because the video content was boring and uninteresting. 23 participants mentioned some negative remarks such as “The video was not interesting” and “It was boring.” Specifically, they mentioned that the way the instructor presented the information on the learning content was unilateral, mainly relying on the textual information. Simply reading the textual content without elaborated explanations was the greatest factor causing the boredom state. Important information was presented visually on the screen, and thus, the instructors’ explanation of simply reading it at the same tone and speed was not perceived to be interesting, thereby causing learners’ state of boredom. Before the experiment, we asked the participants about their interest in the topic of the video content, and most of them responded positively such as “I’m curious about the topic.” Hence, learners who are interested in the topic at the beginning may change their state from curiosity to boredom if the content presentation has no variations. During the interview, several participants mentioned “It was difficult to watch the video one-way” as the cause of their boredom. Video content is continuously presented to learners regardless of their level of understanding and interest at temporal moments. During the analysis of video clips, we observed that learners often showed the behaviors of taking notes, changing body postures, and biting their lips, which are interpreted as conscious or unconscious actions taken to overcome the boredom state.

Next, we examined how much the participants concentrated on the overall learning process. Ten participants responded that although they were bored and had difficulty concentrating, some momentary concentrations occurred when the interesting content appeared. Specifically, they felt interested and concentrated when the instructor explained the content with concrete examples such as the actual cases that happened to violate the research ethics. The learners felt temporary interest by recalling their personal experiences when the examples were presented. However, the interest generated through that moment did not last long. 16 participants indicated that they could hardly concentrate throughout the learning process. In summary, the potential reasons for such affective states include lack of concentration, uninteresting content, and one-way content delivery method. In particular, the situational factor appears critical in causing negative emotions because online video-based learning demands learners’ self-regulation.
Discussion and Conclusion

Discussion of Key Findings

Recognizing learner’s motivational problems in remote online learning during the COVID-19 pandemic, this study attempted to examine the patterns of learner’s academic emotions and facial expressions detected in asynchronous video-based learning. This study focused on, among various academic emotions, boredom, confusion, and frustration, which are important academic emotions that may inhibit the learning process and need additional support. In this section, we revisit the key findings related to each research question.

To answer the first research, this study confirmed that statistically significant relationships exist between certain facial expressions and learners’ emotions. Table 6 is an extended version of Table 2 with the findings from this study for easy comparison. Confusion is the affective state that was easily detectable with the presence of brow lower (AU4) and pulling a lip corner (AU12). Given that these associations were also reported in other studies (e.g., Bosch et al., 2014; D’Mello et al., 2009; McDaniel et al., 2007), AU4 and AU12 are interpreted to be a rather universal expression on the face, indicating the state of learners’ confusion. Frustration was detectable from only one facial expression: inner brow raiser (AU1). Other facial AUs such as AU2, AU4, AU12, and AU14, which were reported to be associated with frustration, were not significant indicators in this study. As seen in Table 4, AU12, the action of pulling a lip corner, appears in both confusion and frustration emotions in the previous studies, making its interpretation of related emotions ambiguous. In the context of this study with Korean adult learners, AU12 appeared to be a better indicator of confusion than frustration.

Interestingly, boredom appeared to have correlations with multiple facial AUs. This finding is rather different from the existing literature that found the difficulty of detecting the affective state of boredom with particular facial AUs (e.g., Craig et al., 2008) and even described boredom as mirroring an emotionless face (McDaniel et al., 2007). D’Mello et al. (2009) suggested that although AU43 (eye closure) was a significant indicator of boredom, it should be used as a secondary source because of the low statistical power of prediction. In the present study, boredom was detectable from certain facial activities such as dimpler (AU14), lid droop (AU41), and blink (AU45). Fatigue/tiredness and the nature of learning materials may explain the different findings between this study and other studies. This study used a long video clip (72 min) mainly in a lecture mode to detect the occurrence of negative emotions. Other studies such as McDaniel et al. (2007) and D’Mello et al. (2009) used an intelligent tutoring system that embeds intelligent functions and problem-solving activities demanding active cognitive processing. The long lecture video may have led the participants to feel tired and show various signs of fatigue/tiredness on their face.

One interesting finding related to fatigue and boredom is that although the participants were given the flexibility to control their learning pace, they did not show the behavior of controlling the video speed or pausing the video. Overcoming fatigue may have done more unconsciously such as partially closing the eyes and making dimples with the mouth muscles rather than through conscious behaviors like pressing a pause button to stop the learning process. On the whole, boredom and dimpler (AU14) were the most frequently occurred emotion and facial expressions. Dimpler is an expression of learners’ action to overcome drowsiness from boredom to continue the learning process. During the interviews, most participants shared that even though they felt bored, they tried to finish watching the video lecture and generally remembered the content of the video well. Therefore, dimpler (AU14), while indicating the boredom state, can also be seen as an expression of learners’ willingness to complete a task under a boring situation.

The second research question aimed to examine learners’ perceptions of their emotional experiences during online video-based learning. We wanted to know the learners’ opinion about their overall experiences during the learning process. The analysis of the video recall interviews indicates that the participants attributed their negative emotions to the nature of learning content and delivery method. They used the expression “boring” and “not

| Emotion     | Other studies* | This study          |
|-------------|----------------|---------------------|
| Boredom     | AU43           | AU14, AU41, AU45    |
| Confusion   | AU1, AU4, AU7, AU12 | AU4, AU12          |
| Frustration | AU1, AU2, AU4, AU12, AU14 | AU1 |

*Bosch et al. (2014), D’Mello et al. (2009), Grafsgaard et al. (2013a, 2013b), McDaniel et al. (2007)
interesting” frequently during the interview to indicate that the source of their boredom came from their interaction with the video content itself. As mentioned earlier, the video content was designed mainly in a lecture mode where the instructor explains the textual information presented in each scene. Such a unilateral, static mode of presentation seems to put the learners in a low arousal state (e.g., relaxation) where they easily turned to the state of boredom. Noteworthy, even the participants who appeared to be curious about the lecture topic also failed to sustain their motivation. On the whole, the interview results imply that a control system or feedback is needed to help learners return to the learning situation at the moment when they become bored and are unable to concentrate.

Implications

Given the massive transition to online video-based learning during the COVID-19 pandemic, the results of this study can provide both pedagogical and technical implications concerning how to detect and intervene in potential issues associated with learner’s motivation and affective states. Content delivery-centered lectures are essential because of the high time efficiency. However, problems arise when a learner feels bored, but scaffolding to mediate the boredom is not provided in learning environments (Lajoie et al., 2020). This study suggests potential interventions to mediate learners’ affective states in three aspects: (a) automatic emotion detection, (b) video content design, and (c) instructor-level strategies.

First, this study suggests that for online video lectures to draw learners’ positive emotions like concentration and flow, the academic emotions must be analyzed through learners’ facial expressions, and learners must be provided with emotional feedback to reduce negative emotions. One potential solution can be in a form of adaptive online video lectures in which the content and feedback experienced by learners change accordingly with the automatic detection of changes in their emotions. Specifically, an adaptive system can provide feedback on specific emotions that appear more frequently since the high occurrence of certain facial expressions is a strong indicator of associated emotions. For example, the high frequency of dimpler (AU14) denotes an expression to overcome the current state of boredom, signaling the moment to provide affect-sensitive feedback to help learners control their boredom state (D’Mello et al., 2009). The development of such adaptive content, however, poses several technically challenging tasks because systems need intelligent computing functions to automatically detect learners’ facial features and to map with emotional states. When sufficient data are accumulated and applied to the existing emotional recognition technology, it is possible to design more interactive online video systems that consider learners’ emotional states.

Second, we suggest the need for the strategic design of video lecture content to induce learners’ positive emotions. The sudden transition to fully remote online learning during the COVID-19 outbreak was unpredicted, and most instructors found themselves unprepared for teaching in online learning environments (Bonk, 2020). Due to the time constraint coupled with technical difficulties, many video lectures were produced and delivered in a unilateral lecture mode. Compared with the face-to-face situation, the isolated nature of online learning makes it more difficult for learners to manage their attention and engagement. Breaking down video lectures into shorter clips is a good design strategy that can be easily incorporated to help engage learners and increase their concentration. Further, following the notion of interest and anxiety in motivation theory (Ainley, 2006), we suggest that more cognitively demanding activities (e.g., asking complex problems and self-explanation) and interesting examples can be inserted in video lectures to minimize a low state of arousal (e.g., relaxation, boredom, and tiredness) and to elicit a certain level of anxiety and interest involving positive activation of emotion (e.g., curiosity and concentration). Learners’ feeling of interest can direct their attention to information-seeking and problem-solving behaviors that are more cognitively demanding and help them escape from the passive state of low arousal.

Lastly, this study suggests that instructor-level interventions are essential to intervene in learners’ motivational and affective experiences (Allen et al., 2006). Simply delivering a lecture video should not be the main task of online teaching. Instructors should conduct pre-lecture and/or post-lecture activities to check the level of learners’ understanding of video lectures. Although we suggested the need for developing affect-sensitive and adaptive systems, detecting all the various types of learner emotions, regardless of whether they are positive or negative, is technically challenging. For instance, mixed findings related to the facial expressions of boredom imply the ambiguity and difficulty of automatic emotion detection in academic settings. Boredom, if it cannot be automatically detected, needs human interventions.

D’Mello and Graesser (2012) proposed the dynamic model of affective states that explains when and how confusion, frustration, and boredom occur based on the equilibrium/disequilibrium theory. When a learner faces cognitive disequilibrium between their current understanding and new information, it leads to a state of confusion. After feeling confused, the learners’ failure to revert to the flow state elicits frustration. If disequilibrium is not resolved for a substantial time, it eventually leads to boredom. Given this sequence of confusion frustration
boredom, instructors need additional scaffolding strategies to prevent learners from falling into a state of boredom due to disequilibrium for a long period. Scaffolding strategies include providing additional materials to help learners understand the content better, providing feedback to clear any confusion or misconceptions, and assigning tasks more relevant to learners’ cognitive level of understanding to help them gain more confidence. Instructors can also consider increasing social interaction to elicit positive emotions among students, given that emotions can be contagious and collective (Xing et al., 2019).

Limitations and Areas for Future Research

Some limitations of this study should be noted. First, generalizing the results of this study is limited to similar research contexts and participant profiles because the number of subjects and the amount of data were rather small. More research studies are needed to explore how different subject profiles, such as age and prior knowledge, may affect learners’ emotional experiences.

Second, detecting facial expressions was conducted manually by human judges, and the level of accuracy may be different from that of automated affective computing programs. Future research may include both automatic analysis and human judgment to supplement the limitation of each method.

Third, the study was conducted with long video content to increase the possibility and occurrences of detecting negative emotions that this study focused on. Our findings may not apply to a situation where learners use video clips that are short and/or invoke more positive emotions. Additionally, we found that the level of fatigue and on-screen behaviors can be useful information to better understand the flow of learners’ affective states. Future research can collect non-intrusive physiological data such as eye tracking, bodily postures, and electroencephalogram (EEG), examine on-screen behavior patterns, and use different forms of video (e.g., shorter video clips and interactive presentations) to gain more comprehensive understandings of learners’ emotional expressions.

Fourth, video recording was conducted in the participants’ homes to minimize the observer effect. However, we do not rule out the possibility that the participants could display or suppress certain affect states because they were aware of being recorded. The literature on video research suggests that removing the observer effect completely is not possible methodologically and technically, and a more productive approach is to analyze data showing the influence of observers, rather than treating them as bad data (Jewitt, 2012). Also, the observer effect does not last long because subjects tend to quickly forget the presence of a camera (Callanan et al., 2007). Hence, we suggest that future research can corroborate video data with other types of measures such as surveys and participant interviews to gain validity of interpretations. Conducting a training session to make participants feel more comfortable and less aware of the camera effect is also a relevant strategy that future research can employ.

Despite these limitations, the contribution of this study is that it analyzes the relationship between facial expressions and affects through multifaceted data centering on learners’ emotional experiences in the context of online video-based learning. Moreover, this study presents implications that can inform the design of a more motivational online learning content. Based on the relationship between affects and facial expressions confirmed in this study, future research can be conducted to develop adaptive online video-based learning systems that are sensitive to learners’ affective states.

Acknowledgements This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A8030048).

References

Ainley, M. (2006). Connecting with learning: Motivation, affect and cognition in interest processes. Educational Psychology Review, 18(4), 391–405.

Ainley, M., Corrigan, M., & Richardson, N. (2005). Students, tasks and emotions: Identifying the contribution of emotions to students’ reading of popular culture and popular science texts. Learning and Instruction, 15(5), 433–447.

Allen, M., Witt, P. L., & Wheelless, L. R. (2006). The role of teacher immediacy as a motivational factor in student learning: Using meta-analysis to test a causal model. Communication Education, 55(1), 21–31.

Anders, A. (2015). Theories and applications of Massive Open Online Courses (MOOCs): The case for hybrid design. International Review of Research in Open and Distance Learning, 16(6), 1–13.

Ashwin, T. S., & Guddeti, R. M. R. (2020). Impact of inquiry interventions on students in e-learning and classroom environments using affective computing framework. User Modeling and User-Adapted Interaction, 30, 759–801.

Baker, R. S., D’Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners’ cognitive-affective states during interactions with three different computer-based learning environments. International Journal of Human-Computer Studies, 68(4), 223–241.

Baker, R. S., Ogan, A. E., Madaio, M., & Walker, E. (2019). Culture in computer-based learning systems: Challenges and opportunities. Computer-Based Learning in Context, 1(1), 1–13.

Balconi, M., & Bortolotti, A. (2013). The “simulation” of the facial expression of emotions in case of short and long stimulus duration. The effect of pre-motor cortex inhibition by rTMS. Brain and Cognition, 83(1), 114–120.

Bonk, C. J. (2020). Pandemic ponderings, 30 years to today: Synchronous signals, saviors, or survivors? Distance Education, 41(4), 589–599.
Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners’ levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior, 39*, 356–367.

UNESCO Institute for Statistics (UIS). (2020). *Survey on national education responses to COVID-19 school closures*. Retrieved from https://infogram.com/final-unesco-education-covid-19-data-1hke60d1x7m525r?live

Xing, W., Tang, H., & Pei, B. (2019). Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of MOOCs. *The Internet and Higher Education, 43*, 100690.

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