ABSTRACT

It is becoming a trend to apply an emotional lens and to position emotions as central to educational interactions. Recently, affective computing has been one of the most actively research topics in education, attracting much attention from both academics and practitioners. However,despite the increasing number of papers published, there still are deficiencies and gaps in the comprehensive literature review in the specific area of affective computing in education.Therefore, this study presents a review of the literature on affective computing in education by selecting articles published from 2010 to 2017. A review protocol consisting of both automatic and manual searches is used to ensure the retrieval of all relevant studies. The final 94 selected papers are reviewed and relevant information extracted based on a set of research questions. This study classifies selected articles according to the research purposes, learning domains, channels and methods of affective recognition and expression, and emotion theories/models as well as the emotional states. The findings show the increased number and importance of affective computing studies in education domain in recent years. The research purposes of most affective computing studies are found to be designing emotion recognition and expression systems/methods/instruments as well as examining the relationships among emotion, motivation, learning style, and cognition. Affective measurement channels are classified into textual, visual, vocal, physiological, and multimodal channels, while the textual channel is recognized as the most widely-used affective measurement channel. Meanwhile, integration of textual and visual channels is the most widely-used multimodal channel in affective computing studies. Dimensional theories/models are the most preferred models for description of emotional states. Boredom, anger, anxiety,enjoyment, surprise, sadness, frustration, pride, hopefulness, hopelessness, shame, confusion,happiness, natural emotion, fear, joy, disgust, interest, relief, and excitement are reported as the top 20 emotional states in education domain. Finally, this study provides recommendations for future research directions to help researchers, policymakers and practitioners in the education sector to apply affective computing technology more effectively and to expand educational practices.

Keywords:

Teaching/learning strategies

Adult learning

Architectures for educational technology

system

# Introduction

Emotions are considered as individual experiences and responses which depend on the conditions in which they appear in human societies (Dupre et al., 2015). Emotions are an essential part of every human being that can influence behavior, thinking ability, decision making, resilience, well-being and the way humans communicate with each other (Morrish, Rickard, Chin, & Vella-Brodrick, 2018; Neophytou, 2013). The study of human-computer interaction is referred to as Affective Computing, which is based on programming and aims to diagnose and measure emotional expression (Lin, Wu, & Hsueh, 2014). R. Picard (1997) coined the term affective computing in 1994, and since then many systems have been developed using affective computing techniques. There has been an increasing number of studies over the past decade to promote the usefulness of human computer/robot interactions (Baker, D'Mello, Rodrigo, & Graesser, 2010; Fu, Leong, Ngai, Huang, & Chan, 2017; Wang, Huang, & Makedon, 2014). The design of affectaware interactions between humans and technology has become a significant research area among researchers into human-computer interaction (HCI) (Gil, Virgili-Gomá, García, & Mason, 2015). The main goal of affective computing is to recognize and measure people's explicit affective appearances and link them to their implicit emotions. Recognition of affective states is useful for analyzing users' reactions in order to elicit behavioral intentions and to create reasonable responses. This can develop an affective-aware system and improve its user interface in potential applications (Handayani et al., 2014). Human behaviors in individual and social communities are strongly influenced by emotions and need to be intensively examined in any human activity, such as e-Learning (Faria et al., 2017). The number of emotions affecting students learning styles and their academic achievements is increasing (Peterson, Brown, & Jun, 2015). Meanwhile, it is believed that there is significant relationships between emotional events and students' developed entrepreneurial competencies (Lackéus, 2014). Thus, it is becoming a trend to apply an emotional lens in emerging academic research and position emotions as central to learning (Baldassarri, Hupont, Abadía, & Cerezo, 2015; M.; Feidakis, 2016; Jiménez, Juárez-Ramírez, Castillo, & Ramírez-Noriega, 2017; Xu, 2018). This has further given rise to an important trend in the development of affective tutoring systems (ATSs), a type of intelligent tutoring systems (ITSs), which are systems with the ability to control learners' adverse emotions (Lin et al., 2014; Malekzadeh, Salim, & Mustafa, 2014). The incorporation of an emotion detection capability can significantly expand the use of educational technologies and offer additional opportunities to improve the overall distance learning outcomes as well as providing new chances for personalized instruction and low-cost delivery of teaching and learning programs (Arroyo, du Boulay, Eligio, Luckin, & Porayska-Pomsta, 2011; Cabada, Estrada, Hernández, Bustillos, & ReyesGarcía, 2018; Caballé, 2015; Shen, Xie, & Shen, 2014). Emotions could be the key factor that affects learning, as well as becoming the driving factor that promotes learning and engagement (Leony, Muñoz-Merino, Pardo, & Kloos, 2013; Lin et al., 2014; Muñoz, Mc Kevitt, Lunney, Noguez, & Neri, 2011). They also play a critical role in decision-making, timing, managing learning activities and thus, increasing the student's motivation in learning (Sandanayake & Madurapperuma, 2013). Therefore, to identify the relationships between the emotional, cognitive and motivational aspects of learning, it is crucial to have reliable methods of emotion recognition in an academic context (Burić, Sorić, & Penezić, 2016)

## Research significance and research qusitions

Affective computing continues to be one of the most actively research topics in education (Poria, Cambria, Bajpai, & Hussain,2017). Over the last decade, researchers have studied different aspects of affective computing in education/learning, such as the role of the teacher in managing students' emotions (Lavy & Eshet, 2018; Siu & Wong, 2016; Urhahne, 2015), the effect of gender, and the influence of shape and color on emotion and learning (Merz & Wolf, 2017; Plass, Heidig, Hayward, Homer, & Um, 2014), students' emotions when doing homework (Goetz et al., 2012), the impact of emotions related to achievement on students' decision-making and learning performance (Chen & Wu, 2015), the design of emotionally-aware models and platforms for educational environments (Baldassarri et al., 2015; Muñoz et al., 2011), and the relation between students' emotional states and the development of entrepreneurial competencies (Lackéus, 2014). However, more in-depth research is necessary to make these understandings applicable to real-world situations because research studies on affective computing, especially in the educational domain, are so far broader and less mature than in other domains (C. Wu, Huang, & Hwang, 2015). To date, few reviews have been conducted to synthesize the results of previous studies on affective computing in the educational domain. Malekzadeh, Mustafa, and Lahsasna (2015) examined studies related to the regulation of negative emotional states in the students' learning process. This review was limited only to empirical research published between 2008 and 2014 and examined different methods for dealing with the users' negative emotional states, such as “boredom”, “anxiety”, and “sadness”, during learning using computerized learning systems. Vogel-Walcutt, Fiorella, Carper, and Schatz (2012) conducted a cross-disciplinary study on the extant literature on the state of boredom, published between the 1980s and 2011, and critically reviewed the description, evaluation, and mitigation of the state of boredom within the educational settings. C. Wu et al. (2015) reviewed the research trends regarding affective computing in education between 1997 and 2013. They identified 90 relevant papers from selected databases and proposed five challenges and problems for affective computing implementation in education. Recently, Barcelos and Ruohotie-Lyhty (2018) focused on articles related to beliefs and emotions in second language teaching by covering theoretical frameworks.

However, the literature reviews conducted by Malekzadeh et al. (2015) and Vogel-Walcutt et al. (2012) merely focused on specific aspects of affective computing in education. Malekzadeh et al. only considered the negative emotional state of users, while VogelWalcutt et al. focused on the emotion state of boredom. In their review of the selected studies, C. Wu et al. (2015) did not consider the research purposes, emotion theories/models, and instruments used. Furthermore, they only covered the literature up to 2013 and stated that their review was limited to only seven journals. Meanwhile, review of Barcelos and Ruohotie-Lyhty (2018) is limited to second language teaching only.

Hence, our present study is a new attempt to fill the gaps and respond to Wu et al., 2015's call for more comprehensive review of affective computing in education/learning. Our study also conducts a more comprehensive examination and analysis of recent research findings in the field, in order to classify those studies based on their research trends, purposes, learning domains, affective recognition and expression channels and methods, emotion theories/models, instruments used, and the emotional states studied. This study attempts to classify the research purposes and to identify the major affective measurement channels and the methods used in each channel as well as to clarify the integration of different channels into the multimodal-based affective computing channel. Furthermore, this study discussed the different emotion theories/models and used the self-reported questionnaire instruments in affective computing studies. All these aspects are new and have not been addressed in the previous research.

To achieve the aim of this study, we systematically reviewed the recent relevant literature between 2010 and 2017 published in journals, conferences, and workshops, based on the guidelines developed by Kitchenham and Charters (2007). This was done to ensure a more comprehensive analysis of previous studies and to obtain a better understanding of current developments and trends in affective computing in education, which will provide recommendations for practitioners and academics for further research directions. Specifically, we have formulated the following four research questions to help us in managing the literature review and in achieving the research goals:

(RQ1) What are the trends in affective computing in education/learning as can be gauged from the selected papers?

(RQ2) What are the main research purposes and learning domains addressed the selected papers?

(RQ3) What are the main affective measurement channels and methods used in the selected papers?

(RQ4) What are the major theories/models of emotion adopted and the emotional states considered in the selected papers?

It is hoped that this review can help researchers in identifying areas which have already been investigated or require further investigation, as well as updating practitioners on the new developments of affective computing in the educational context.

The remainder of this study is structured as follows. Section 2 explains the research method used to conduct a systematic literature review in this study. Section 3 reports the results of the review of the selected papers, based on the four research questions. Section 4 discusses the outcomes and gives suggestions for future research based on the four research questions (RQs). Finally, Section 5 summarizes all the findings and draws conclusions.

# Review method

Kitchenham and Charters (2007) guideline is one of the most widely used and reliable step by step approach for performing systematic literature review (SLR) in all fields (Asadi & Dahlan, 2017; Elaish, Shuib, Ghani, & Yadegaridehkordi, 2017; Elaish, Shuib, Ghani, Yadegaridehkordi, & Alaa, 2017). Thus, in this study we followed Kitchenham and Charters (2007) guidelines in order to come out with accurate, clear, and transparent literature review. Kitchenham and Charters (2007) guidelines involve the following distinct activities: formulate the research questions; develop a review protocol; identify inclusion and exclusion criteria; develop the search strategy and study selection process; perform data extraction and synthesis; conclude and report the results. Details about this procedure and its associated steps are discussed in the following subsections.

## 2.2 inclusion and exclusion criteria

In this study, we considered the articles from jounals, conferences and workshops published in the English language from January 2010 to December 2017. There were some reasons behind selecting this period of time in this study. First, this review is a complement of previous review studies (Malekzadeh et al., 2015; Vogel-Walcutt et al., 2012; C.; Wu et al., 2015) to provide more in-depth understanding of affective computing in education in recent years. Second, the term affective computing has been increasingly used in education related studies since 2010 (C. Wu et al., 2015). However, there is no effort that specifically focuses and reviews studies published from 2010 onwards. Therefore, this study systematically collected related studies from 2010 to 2017 to comprehensively classify and conclude affective recognition and expression channels, methods, models, and instruments in the education domain which have not been considered by previous researchers. The databases searched include ISI Web of Knowledge, ScienceDirect, IEEE Explore, and Springer Link. The search excluded editorials, prefaces, poster sessions, panels and tutorial summaries, articles that were not available in a complete version, not peer-reviewed or not written in English. It is noted that, to be included in this study, an article should meet all inclusion criteria and not should satisfy any of the exclusion criteria. Table 1 summarizes the inclusion and exclusion criteria

# Review results

## 3.2（RQ2）what are the main research purposes and learning domains addressed the selected papers?

## 3.1.1 Research Purposes

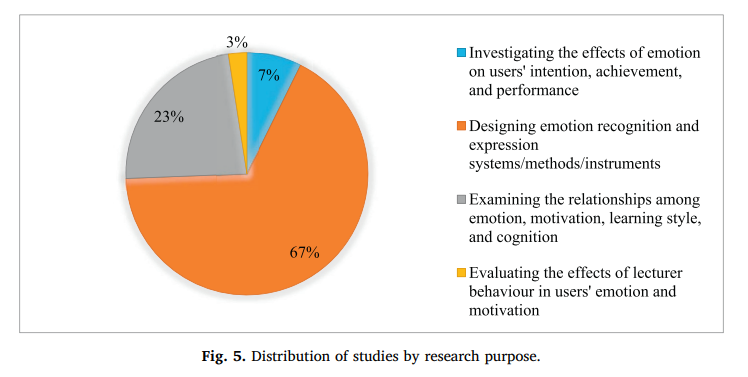
## In this research, each paper was classified into one of four categories based on its stated research purpose: (1) investigating the effects of emotion on users' intention, achievement, and performance; (2) designing emotion recognition and expression systems/methods/instruments; (3) examining the relationships among emotion, motivation, learning style, and cognition; and (4) evaluating the effects of a lecturer behavior in students' emotion and motivation. Fig. 5 shows that the most common research purpose was designing emotion recognition and expression systems/methods/instruments (67%), followed by examining the relationships among emotion, motivation, learning style, and cognition (23%), with only 7% of studies investigating the effects of emotion on users' intention, achievement, and performance, and just 3% evaluating the effects of lecturer behavior on users' emotion and motivation

### 3.2.2 learning domains

This study adopted the QS World University Rankings list, 2016, which identifies five popular learning domains, covering 42 subjects, as shown in Table 3.The main learning domains are “social sciences and management” (21%) and “engineering and technology” (19%), followed by “natural sciences” (15%), “arts and humanities” (6%) and “life sciences and medicine” (4%). 13% of studies conduct in mixed learning domains. However, a large group of studies (22%) do not specify any learning domain as they mainly focus on the design of affective recognition systems and investigation of emotional states.

## 3.3 （RQ3）what are the main affective measurement channels and methods used in the selected papers？

In this study, the affective measurement channels were classified as textual, visual, vocal, physiological, and multimodal, which is any integration of those channels. Fig. 6 shows a logic chart of affective measurement channels and the methods used in each channel. Most studies were reported in the textual channel, in which most studies used a self-reporting method, either through questionnaires (32 studies) or text (14 studies), and 3 studies used expert observation. In the visual channel, facial expression was the most widely used method (8 studies). In the physiological channel, the common affective measurement methods were electroencephalogram (EEG) (4 studies), eye tracking (3 studies), electrocardiography (ECG), heart rate variability (HRV), and skin conductance level (SCL) (2 studies each) (refer to Appendix A for definitions of physiological affective measurement methods). There were 16 studies in the multimodal channel. There were no studies that only applied the vocal channel (speech and/or prosody& intonation methods). However, the vocal channel is considered a distinctive channel as it is used as a part of the multimodal channels in some studies.



### 3.3.2 Multimodal channel Integration of different channels into the multimodal-based affective computing channel is shown in Table 5. According to this table, integration of textual and visual channels, used in seven (7) studies, is the most widely-used multimodal channel in affective computing studies. Integration of textual, visual, and physiological (three studies), and visual and vocal (two studies) were also carried out. Other integration strategies used included: textual + visual + vocal, textual + vocal + physiological, visual + vocal + physiological and textual + visual + vocal + physiological (each with one study).

## 3.4 (RQ4) what are the major theories/models of emotion adopted and the emotional states considered in the selected papers?

### 3.4.1 Emotion theories/models

Two main steps are defined for an affective recognition system to measure a user's emotions. First, the system collects the data from users' emotions then it predicts the user's emotional states based on a predefined model and previously recognized emotions, and also from the stored data (Jaques, Vicari, Pesty, & Martin, 2011, pp. 599–608). Generally, a categorical model aims at explaining the cognitive process that elicits an emotion from predefined categories of emotions, whereas a dimensional model refers to a twodimensional model of affects, where emotions are considered to be a combination of arousal/learning and valence/affect. Table 6 shows the frequency analysis of different emotion theories/models and their description in terms of dimensional or categorical emotional states. The table shows that the majority of studies (20 papers) used the dimensional theory/model for description of emotional states, in which the control-value theory of achievement emotions proposed by Pekrun (2006) is the most widely-used theory. Kort, Reilly, and Picard (2001) learning spiral model and Russell (1980) circumplex model of emotions are the second and third most commonly used frameworks, respectively. However, other dimensional theories/models such as the Limited Capacity Model of Motivated Mediated Message Processing (Lang, 2006), Whissell's evaluation-activation space (Whissel, 1989), Posner's Circumplex model of affect (Posner, Russell, & Peterson, 2005), and Feidakis's emotion model (Michalis Feidakis, Daradoumis, Caballé, Conesa, & Gañán, 2013), were each used once. Nine studies used categorical-based theories, and these include five which used the Facial Action Coding System (FACS) (Ekman, 1992; Ekman & Friesen, 1971) and four using Ortony, Clore and Collins' Structure of Emotions (OCC) (Ortony, Clore, & Collins, 1988)

### 3.4.2 emotional states

Table 7 shows the frequency analysis of the top 20 emotional states used in the studies in the educational domain. Owing to the diverse emotional states involved in education, the majority of the less frequent emotional states are not included in the table. The table shows that the top 20 emotional states used in affective computing studies are boredom, anger, anxiety, enjoyment, surprise, sadness, frustration, pride, hopefulness, hopelessness, shame, confusion, happiness, natural, fear, joy, disgust, interest, relief, and excitement. The top emotions reported in the table indicate that researches in the education domain are more concerned about negative emotions. This also is supported by (Malekzadeh et al., 2015) who reported that many empirical researches have focused on applying different techniques for managing the negative emotional state of the learners to improve learning productivity during learning episode. Meanwhile (Vogel-Walcutt et al., 2012), emphasized on controlling negative state of boredom within educational settings for enhancing educational practices. Generally, many researchers strongly agree that ITSs would significantly improve learners' performance if they can manage negative emotional states of the learners (D'Mello & Calvo, 2013; Tian et al., 2014).

### 3.4.3 Distribution of emotional states by affective measurement channels

The distribution of the research papers by the top 20 emotional states and affective measurement channels is shown in Fig. 7. Among the measurement channels, the textual channel is mainly used to measure boredom, anger, anxiety, and enjoyment; the visual channel is mainly used for measuring surprise, anger, sadness, and disgust; the physiological channel is mainly used for measuring boredom, frustration and disgust; and multimodal channel is mainly used for measuring anger, surprise, frustration, confusion, and natural emotions

# 4. discussion

## 4.1  (RQ1) what are the trends in affective computing in education/learning as can be gauged from the selected papers?

From Fig. 4 it has been observed that there is a trend for more research and interest in affective computing, indicating the increasing importance of this subject in the educational domain. This finding was supported by (C. Wu et al., 2015) who reported a growing trend of affective computing research in education domain. It is expected that almost all learning applications and platforms will have an embedded capability to detect and monitor learners' emotions in the near future. More specifically, considering the rapid development and importance of mobile learning in the current educational systems (Elaish, Shuib, Ghani, & Yadegaridehkordi, 2017; Elaish, Shuib, Ghani, Yadegaridehkordi et al., 2017; Yadegaridehkordi & Iahad, 2012), developing emotional models that can be integrated into mobile devices such as tablet and mobile phone is more substantial (Subramainan, Yusoff, & Mahmoud, 2015). Therefore, the educational environments of the future might be different and be more sophisticated than the current situation. Meanwhile, given the significant role of educational institutions in the development of societies, knowledge and technologies, developing strategies to fulfill their historic mission of teaching and research becomes crucial for various stakeholders (Olcay & Bulu, 2016). In this sense, policymakers and practitioners in education are urged to make appropriate plans to allocate the necessary resources for supporting future research and development in affective recognition and expression.

## 4.2  (RQ2) what are the main research purposes and learning domains addressed the selected papers?

Crompton, Burke, Gregory, and Gräbe (2016) and Wingkvist and Ericsson (2013) emphasize the importance of determining research methods and purposes in order to interpret and share results and facilitate knowledge transfer. However, previous reviews on affective computing have not attempted to review and classify the research purposes. This study presents a new finding to fill the gap. According to Fig. 5, of the 94 studies reviewed in this research, 67% considered “designing emotion recognition and expression systems/methods/instruments” as the primary research purpose. As C. Wu et al. (2015) and Malekzadeh et al. (2015) pointed out, the increased application of affective computing in education has only started in recent years. Thus, it is not surprising that there still has been a lack of emotion recognition and expression systems/methods/instruments. Therefore, most of the researchers have focused their efforts on designing and proposing emotion recognition and expression systems/methods/instruments for educational environments (Caballé, 2015; Khalfallah & Slama, 2015; Salmeron-Majadas et al., 2014; Sandanayake & Madurapperuma, 2013; Yang, Alsadoon, Prasad, Singh, & Elchouemi, 2018). In view of this development, researchers and practitioners in education are urged to incorporate new devices and equipment, such as intelligent sensors, cameras, speech prosody and intonation recognition, in the design and development of affective recognition systems. Further, exploiting the latest information technology such as cloud computing, green information technology and internet of things, in the design process could affect a dramatic change in affective computing. Meanwhile, by considering the significant impacts of color features like chroma, hue or lightness on emotions, concentrating on emotional differences and color preferences in different learning situations can be another research direction in education domain (Sokolova & Fernández-Caballero, 2015). Other research purposes include examining the relationships among emotion, motivation, learning style and cognition (Baldassarri et al., 2015; Harley, Carter, et al., 2015; Heidig, Müller, & Reichelt, 2015; Tj et al., 2015), and investigating the effects of emotion on users' intention, achievement, and performance (Chen & Wu, 2015; Chung, Cheon, & Lee, 2015; Lackéus, 2014).  
Although the design of an affective recognition system or method is the first step in integrating such systems into the educational environment, it is important to examine the performance, effectiveness, and usability of these systems, which will help in designing more efficient and applicable systems and methods. Therefore, researchers and practitioners should examine and appraise these factors as soon as a new affective recognition system or method is introduced into the educational sector. Studies on affective computing in education have been widely conducted in different learning domains. The results show that affective computing is generally applied in courses related to social science and management field (Finch, Peacock, Lazdowski, & Hwang, 2015; Heidig et al., 2015; Peterson et al., 2015; Urhahne, 2015), and to a considerably lesser extent in Life sciences & medicine field. Therefore, it is hoped that researchers in the different disciplines can collaborate in designing and developing suitable applications for under-represented courses. So far, the potential use of affective learning in latest open and online learning platforms such as Massive Open Online Courses (MOOCs), M-learning (Mobile learning), and CSCL (Computer Supported Collaborative Learning) have not critically explored. Hence, future study should be conducted on emotion-sensitive computerized MOOCs, Mlearning, and CSCL to provide more facilitated and personalized learning environments (M. Feidakis, 2016).

## 4.3  (RQ3) what are the main affective measurement channels and methods used in the selected papers?

Afzal and Robinson (2011) classified channels of nonverbal behavior into visual, vocal, and physiological. C. Wu et al. (2015) reported the use of conventional questionnaires, skin conductance response, facial expression, heartbeat, EEG and EMG, as some of the instruments used in affective computing in education/learning studies. However, based on the papers reviewed, no study has comprehensively classified the affective measurement channels and methods used in educational environments. This study identified five major categories of affective measurement channels: textual, visual, vocal, physiological, and multimodal (see Fig. 6). Each channel employs different methods: the textual channel uses self-report by questionnaire or by text, and expert observation; the visual channel uses facial expression, head pose, body gesture, eye gaze, and keystroke dynamics; the vocal channel uses speech and prosody and intonation; the physiological channel uses EMG, ECG, HR, HRV, EOG, EEG, EDA, BVP, SCL, breathing rate, temperature, and eye tracking, while the multimodal channel employs any integration of the above-mentioned channels. This study presents new findings not reported previously. The results show that speech and prosody and intonation have not been used alone in any of the studies reviewed. Therefore, it can be concluded that a single vocal channel is the least useful channel, as it is only applicable as a part of multimodal channel for detecting emotional states in education. It is widely believed that in emotional speech processing, the emotional speech changes according to the different acoustic features (Scherer, 1986). The individuals show their feelings not only by the acoustic features, but also with the content they intend to deliver. Different phrases, words, and syntactic structures can make different kinds of expression styles and results. Consequently, the lack on the capture and the analysis of more detailed/reliable physiological features limits emotional speech synthesis and recognition systems (Tao & Tan, 2005). Therefore, affective interaction multiagent systems need to be designed to precisely capture emotions of a human via voice, facial expression and physiological signals (C. Wu et al., 2015). Meanwhile, in speech emotion recognition, learning representations for natural speech segments which can be utilized effectively under unconstrained and noisy settings is a substantial challenge. Thus, it is more effective to collect and transfer information and knowledge through other channels in addition to the individual voice channel (Albanie, Nagrani, Vedaldi, & Zisserman, 2018).  
Table 8 presents a brief comparative overview of different channels. Although the use of textual methods is easy and cheap, they are struggling with some challenges such as cultural and language differences, not being real time, and not accurate enough (Broekens & Brinkman, 2013; Lin et al., 2014). Visual methods provide extra information and are practically more deployable. However, they associate with noisy sensors that are largely unscalable, image processing problems, and privacy issues. Vocal sensing provides accurate information through integration into interactive user interfaces. However, it largely is limited to dialogue-based learning systems. Physiological analysis requires tightly controlled environmental conditions as well as specialized and fragile equipment, thus may not always be suitable for learning environments (Chen & Sun, 2012; Chen & Wu, 2015; Gil et al., 2015). The literature on affective computing recommends that integrating different input sources can enhance the outcome of affect recognition. However, technical challenges associated with integrating channels and difficulties associated with managing and interpreting huge amount of data generated from various channels are the basic issues that need to be considered when using multimodal channels (D'Mello & Kory, 2012)

The textual channel was the most commonly used affective measurement channel, as the vast majority of studies utilized a selfreported questionnaire method (see Fig. 6). This could be due to the advantages of the questionnaire, such as its reliability, validity, ease of use, meaningful feedback, cheapness, and the fact that it is not dependent on the use of any special equipment (Broekens & Brinkman, 2013; Lin et al., 2014). In addition, as shown in Table 4, AEQ is the most popular questionnaire in affective computing studies. AEQ was proposed by Pekrun, Goetz, Titz, and Perry (2002) and Pekrun, Goetz, Frenzel, Barchfeld, and Perry (2011) to recognize emotional states experienced by students. The instrument measures enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, and boredom in different situations using 24 scales. AEQ has been used successfully in affective computing studies (Moga, Sandu, Danciu, Boboc, & Constantinescu, 2013; Muis et al., 2015; Pekrun, Cusack, Murayama, Elliot, & Thomas, 2014; Urhahne, 2015). However, since the educational environment provides a good source of different emotional experiences, researchers are encouraged to develop a multidimensional questionnaire that is comprehensive enough for measuring different emotions for assessing complex relationship between cognitive and motivational aspects of learning in different teaching and learning activities and situations (Burić et al., 2016). It is often difficult for individuals to recognize and report their own emotions under different situations. Meanwhile, controlling and monitoring the dynamic student's characteristics and emotions during learning activities is problematic through textual method (Chrysafiadi & Virvou, 2013). According to Poria et al. (2017), the way humans naturally express their emotions and feelings is typically multimodal combination of the textual, audio, and visual modalities. Thus, according to Vogel-Walcutt et al. (2012), practitioners and researchers should explore optimal combination self-reported instruments and other physiological affective recognition methods for more accurate assessment. In this situation, the large amount of collected information could help in recognizing hidden emotions, as well as achieving more comprehensive understanding of learners' behavior in the real environment (D'Mello & Kory, 2012).   
The integration of textual and visual channels is the most widely used multimodal channel in affective computing studies. Facial expression is the best direct method for accurately detecting emotional states specially for virtual learning environments (Baldassarri et al., 2015; D'Mello & Kory, 2012; Yang et al., 2018). Meanwhile, according to Tao and Tan (2005), the quality of facial simulations has enhanced significantly due to the recent advances of hardware and matching software. Therefore, multimodal-based studies have mainly tried to solve the challenges of the textual channel (e.g., the Hawthorne effect, and lack of capability to recognize changes in affective states) by taking advantage of the visual channel, especially the facial expression method (Lin et al., 2014; Salmeron-Majadas et al., 2014; Santos et al., 2013; Tj et al., 2015). Visual + vocal was reported as another common multimodal channel (Alepis & Virvou, 2011; Hussain et al., 2011). According to Tao and Tan (2005), audio-visual mapping is a powerful method for capturing and managing users' emotions and feelings in a desired system. This finding supports the findings of D'Mello and Kory (2012) who conducted meta-analysis and found that the vast majority of multimodal-related studies are bimodal and generally focused on audio-visual affect recognition.

Some researchers believe that multimodal approaches help to overcome the constraints of individual channels and improve the accuracy over the best unimodal channels (Banda & Robinson, 2011, pp. 200–207; D'Mello & Kory, 2012; Michalis; Feidakis, Caballé, Daradoumis, Jiménez, & Conesa, 2014; Harley, Bouchet, et al., 2015). However, some theoretical, methodological, and measurement challenges still need to be addressed before these types of channels can be used for measuring learners' emotions empirically. Furthermore, arising from the findings, some questions still remain on how to integrate different channels to achieve better results, how to deal with the different data types or formats that different methods provide, how to manage disagreements on emotional states among various methods at a particular time, how to get recognition results; these need to be answered in future in-depth studies (Poria et al., 2017). Thus, researchers are urged to investigate the integration of different methods in order to propose multimodal-based affective recognition systems that can comprehensively pattern human emotional states under different teaching/learning conditions. Moreover, more efforts should be done in the parameters integration (Tao & Tan, 2005).   
Generally, textual and visual can be appropriate affective recognition channels in financially restricted situations and/or in situations that the available equipment is limited. Visual methods provide extra information and are practically more deployable. Non-invasive physiological processes are appropriate in collecting and understanding the emotional state of the individuals without having any physical contact. Multimodal approaches help to overcome the constraints of individual channels and improve the accuracy over the best unimodal channels. However, challenges associated with managing actual aspects of affective states and behavior, technical possibility, and practical matters make the selection of appropriate channel for emotional detection and interpretation more difficult. Meanwhile, limitations of privacy, ethics, and comfort limit the deployment, design, and implementation of the most fit sensing technologies. Thus, researchers, policymakers and practitioners in education sector need to adopt more appropriate affective recognition channels and methods based on specific situations and on the availability of enough equipment and financial resources by evaluating and comparing benefits and challenges associated to each channel and its methods.

|  |  |  |  |
| --- | --- | --- | --- |
| Channel | Method | Pros | Cons |
| Textual | Self-reported by questionnaire  Self-reported by text  Expert observation | -Easy to implement and use  -Cheap and not dependent on the use of any special equipment | -Cultural and language differences  -Not accurate enough |
| Visual | Facial expression  Head Pose  Body gesture | -Natural and observable  -Cheap equipment exception for gesture  -Extra information  -Practically deployable | -Time and resource consuming  -Noise  -Image processing problems  -Privacy issues |
| Vocal | Speech  Prosody & Intonation | -Natural, discernible  -Accurate  -Integrated into interactive user interfaces  -Practically deployable | -Limited to dialogue-based systems  -Time and resource consuming  -Cultural and language differences |
| Physiological | EEG  ECG  HRV  BVP  SCL  HR  Eye tracking | -Can be extended for real-time processing  -Easy to access the bio-signals | -Unobservable  -Issues with comfort and privacy  -Requires tightly controlled environmental conditions  -Specialized and fragile equipment  -Low recognition accuracy on dominance emotion  -Difficult to interpret collected data |
| Multimodal | Textual, Visual, Physiological  Textual, Visual  Textual, Visual, Vocal, Physiological  Textual, Vocal, Physiological  Visual, Vocal  Textual, Visual, Vocal  Visual, Vocal, Physiological | -Overcome the constraints of individual channels  -Improve the accuracy over the individual channels | -Technical challenges associated with collecting adequate and realistic data  -Difficult to manage and interpret huge amount of data generated from various channels |

## 4.4 (RQ4) what are the major theories/models of emotion adopted and the emotional states considered in the selected papers?

In previous studies, emotional states were discussed under categorical and dimensional emotion theories/models. Categorical models reflect discrete emotions such as fear and anger, to model a person's emotional states, whereas dimensional models represent a person's affective states in a multi-dimensional space, such as valence-arousal or learning-affect spaces (D'Mello & Kory, 2012). Table 6 shows that the selected studies on affective computing in education mostly focused on applying a dimensional theory/model for description of emotional states (Finch et al., 2015; Harley, Bouchet, et al., 2015; Moga et al., 2013; Muis et al., 2015), of which the most popular is the control-value theory of achievement emotions proposed by Pekrun (2006). This finding is supported by Peterson et al. (2015) who stated that achievement emotions-related research has been dominated by Pekrun's theory. According to the control-value theory of achievement emotions, students' beliefs regarding their cognitive quality are significantly linked to their control and value appraisals of the academic environment, which in turn, influence their emotional learning results (Pekrun, 2006).  
The FACS proposed by Ekman and Friesen (1971) and OCC proposed by Ortony et al. (1988) are the two main categorical models in affective-related studies. Ekman and Friesen introduced anger, fear, disgust, surprise, joy, and sadness as six “basic emotions”. Gil et al. (2015) found Ekman's model to be the most commonly-used categorical affective computing cognitive model. However, according to Banda and Robinson (2011, pp. 200–207) the FACS developed by Ekman and Friesen is the most widely used system for the coding and measurement of facial movements. On the other hand, Ortony et al.'s OCC model is one of the well-known cognitivebased psychological models for recognizing and interpreting emotional states (Jaques et al., 2011, pp. 599–608) because it is based on the cognitive theory of emotions and is easy to implement computationally. This model confidently explains the appraisal and the cognitive process that causes an emotion, from twenty-two fixed emotional categories. However, neither FACS nor OCC include boredom and interest, which are important and relevant affect states in the learning environment (R. W. Picard et al., 2004).  
Since emotional diversity is closely related to the selected theoretical foundation (Gross, 1998), limiting the range of emotions for theoretical reasons might lead to missing important parts of students' affective states. In comparison to the categorical approach, dimensional theories/methods are more preferable because they consider and describe a wide range of emotional states (Michalis Feidakis et al., 2013). This can be one of the reasons why affective computing-related studies in education generally employed dimensional theories/models for description of emotional states. However, according to (D'Mello & Kory, 2012), a mixed classification can also be employed to broaden the emotional states studied. Since performing mixed classification is relatively rare (D'Mello & Kory, 2012), future researches can focus more on application of this classification in education domain. Generally, only a few studies have explicitly described and explained the emotion theory/model they used for affective recognition in education. Thus, a gap exists between theory and practice. Theories of affect in learning need to go through practical testing through different affective measurement channels and methods and evolve in the real learning processes in order to provide more effective, personalized, and  
adaptive emotional feedback in learning environments.

Table 7 shows the 20 top emotional states that have been used in studies in the educational domain. The table indicates that most of the studies focused on the management of negative emotions such as boredom, anger, and anxiety, and these are perceived as hindrance in educational environment. This result is supported by Malekzadeh et al. (2015) and Vogel-Walcutt et al. (2012) who reported that many empirical researches have focused on applying different techniques for managing the negative emotional state of the learners to improve learning productivity during learning episode. According to Vogel-Walcutt et al. (2012), more in-depth research is still required to recognize technology-based assessment and qualification approaches that focus on individual negative academic emotions such as boredom. Although negative emotions are very important, researchers and other concerned stakeholders need to explore additional strategies aimed at managing the positive emotions, in order to enhance efforts in achieving the desired academic goals. Further, observing how different emotional states influence the learning processes and outcomes in different learning approaches (such as online learning, face-to-face learning, mobile-learning, game-based learning) could be another direction for future studies. The previous studies have paid little attention to specific academic-related emotions with respect to specific context, age groups, gender, and subject domains. Therefore, researchers and practitioners could work towards providing a unique categorization of academic-related emotions, based on different age groups, genders, and subject domains, in order to accelerate and improve the process of affective computing with minimum resource requirements in educational environments. Another interesting direction for future study is to find ways of promoting and triggering positive academic-related emotions or to prevent negative ones in educational environments. For example, as suggested by (Vogel-Walcutt et al., 2012), teachers could permit learners to be flexible in choosing their topic for a particular class project rather than assigning a certain topic to them. This method can trigger positive emotions and alleviate negative feelings of students in performing the assigned activity.  
Fig. 7 shows the distribution of the top 20 emotional states by affective measurement channels. This figure can provide some insights on selecting appropriate affective recognition channel and method based on different emotions under consideration or vice versa. There is little reports that some of the physiological signals are associated with the “basic emotions” such as sadness, anger, and disgust (Calvo & D'Mello, 2010). According to Li, Cheng, and Qian (2008), facial expressions are used to classify expressions into basic emotions such as surprise, fear, anger, and disgust. However, this study goes a step further and explores the relationship between the top emotional states and affective measurement channels in education domain. This is a new finding that has not been reported in previous studies. This finding provides a guideline to identifying the channel that is most appropriate for recognition of special emotional states. For example, textual methods are more suggested for exploring boredom, anxiety, anger, and enjoyment emotional states, while multimodal methods are more appropriate for detecting anger, surprised, frustration, disgust, and confusion emotional states. The channels can also be ranked, and in the event that some of them are unavailable, for instance, due to limitation of resources and equipment, a comparable or the next best channel set can be selected. For example, the textual channel is the first option for measuring boredom states. However, multimodal-based channels can be considered as a substitute in special situations, such as for students with certain disabilities (such as blindness) and for non-literate older who cannot involve in textual methods.

# Conclusion

This study has presented an overview of affective computing in education. A systematic literature review method was adopted from Kitchenham and Charters (2007) to answer four research questions on different aspects of affective computing in education. The review covers the studies published between 2010 and 2017 and indexed in ISI Web of Knowledge, ScienceDirect, IEEE Explore, and Springer Link databases. A stepwise systematic process was performed and inclusion/exclusion criteria were applied to filter and select the relevant studies. The final selected papers were reviewed and relevant information extracted based on a set of research questions. Eight new findings are presented that have contributed to answering the four research questions: (1) the continues number of published materials shows that the importance of affective computing in education has significantly increased in recent years; (2) the research purpose of most affective computing studies is based on designing emotion recognition and expression systems/methods/instruments, as well as on examining the relationships among emotion, motivation, learning style, and cognition; (3) although affective computing has been actively investigated in the social sciences and management domain, a large number of studies do not specify any learning domain as they mainly focus on the design of affective recognition systems and investigation of emotional states; (4) affective measurement can be classified into textual, visual, vocal, physiological, and multimodal channels, with the textual channel being the most widely-used affective measurement channel, in recent years; (5) AEQ is the most popular questionnaire used in affective computing research in education studies; (6) integration of textual and visual channels is the most widely-used multimodal channel in affective computing studies; (7) dimensional theories/models are preferred for description of emotional states, among which the control-value theory of achievement emotions proposed by Pekrun (2006) is the most widely-adopted theory; and (8) boredom, anger, anxiety, enjoyment, surprise, sadness, frustration, pride, hopefulness, hopelessness, shame, confusion, happiness, natural emotion, fear, joy, disgust, interest, relief, and excitement, are the top 20 emotional states. Finally, this study provides recommendations and insights for future research directions. By reviewing the recent studies on affective computing in education and providing new findings, it is hoped that this study can provide policymakers and practitioners in the education sector new insights into the more effective application of affective computing technology. The findings can also serve as a basis for educational researchers who look for new research opportunities and directions in this field.