EMOTIONS, TECHNOLOGY, DESIGN, AND LEARNING
Emotions and Technology

Communication of Feelings for, with, and through Digital Media

Series Editor
Sharon Y. Tettegah

Emotions, Technology, Design, and Learning

Volume Editors
Sharon Y. Tettegah and Martin Gartmeier
EMOTIONS, TECHNOLOGY, DESIGN, AND LEARNING

Edited by

SHARON Y. TETTEGAH
Professor, University of Nevada, Las Vegas
College of Education, Las Vegas, NV, USA
Beckman Institute for Advanced Science and Technology,
National Center for Supercomputing Applications, affiliate,
University of Illinois, Urbana, IL, USA

MARTIN GARTMEIER
Postdoctoral Research Fellow, Technische Universität München,
TUM School of Education, Susanne-Klatten Endowed Chair for Empirical Educational Research,
München, Germany
CONTRIBUTORS

Michael Atkinson
Boise State University, Boise, Idaho, USA

Whitney Baer
University of Memphis, Memphis, Tennessee, USA

Devshikha Bose
Boise State University, Boise, Idaho, USA

Ling Cen
College of Communication Engineering, Jilin University, Changchun, China

H. Chad Lane
Department of Educational Psychology & Illinois Informatics Institute, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

Danielle Clewley
University of Memphis, Memphis, Tennessee, USA

Joanna C. Dunlap
University of Colorado Denver, Denver, Colorado, USA

Shi Feng
University of Memphis, Memphis, Tennessee, USA

Arthur C. Graesser
University of Memphis, Memphis, Tennessee, USA

Daphne Greenberg
Georgia State University, Atlanta, Georgia, USA

Yungwei Hao
National Taiwan Normal University, Taipei, Taiwan

Jason Matthew Harley
Computer Science and Operations Research, University of Montréal, and Department of Educational and Counselling Psychology, McGill University, Montréal, Québec, Canada

David P. Hays
University of Memphis, Memphis, Tennessee, USA

Fengye Hu
College of Communication Engineering, Jilin University, Changchun, China

Robert J. Jensen
Emory University, Atlanta, Georgia, USA

Ulas Kaplan
CREATE Lab, New York University, New York, USA
Yanghee Kim  
Utah State University, Logan, Utah, USA

Patrick R. Lowenthal  
Boise State University, Boise, Idaho, USA

Jim Murtagh  
Boise State University, Boise, Idaho, USA

Sanghoon Park  
Instructional Technology Program, College of Education, University of South Florida, Tampa, Florida, USA

Bogdan Pătruț  
“Vasile Alecsandri” University of Bacau, Bacau, Romania

Jan L. Plass  
CREATE Lab, New York University, New York, USA

Enilda Romero-Hall  
University of Tampa, Tampa, Florida, USA

Candace L. Sidner  
Worcester Polytechnic Institute, Worcester, Massachusetts, USA

Diantha Smith  
Utah State University, Logan, Utah, USA

Roxana-Petronela Spatariu  
“Vasile Alecsandri” University of Bacau, Bacau, Romania

Jeffrey Thayne  
Utah State University, Logan, Utah, USA

Patricia Vela  
Emory University, Atlanta, Georgia, USA

Patricio A. Vela  
Georgia Institute of Technology, Atlanta, Georgia, USA

Breya Walker  
University of Memphis, Memphis, Tennessee, USA

Fei Wu  
College of Automation Science and Engineering, South China University of Technology, Guangzhou, China

Cindy S. York  
Northern Illinois University, DeKalb, Illinois, USA

Zhu Liang Yu  
College of Automation Science and Engineering, South China University of Technology, Guangzhou, China
FOREWORD

With respect to technology, it is important to place terms and tools within a historical context, given that in today’s society, when speaking to a person who is a Millennial (individuals who are born in the early 1980s to 2000), s(he) may tell you that technology is the Internet and smart phones. For the Millennial, then, technology may only mean digital or biotechnologies. If we were to speak broadly to some individuals from the Silent Generation, Boomers, Millennials, and Generation Y, technology may also mean automobiles, airlines, overhead projectors, flashlights, microwaves, ATMs, etc. Hence, technology in the twenty-first century can mean many things. For example, technology could mean software applications, hardware, social media platforms, functional magnetic resonance imaging, mobile technology, learning, and content management systems, to name but a few.

Humans and other animals have used tools for centuries; however, the most important aspect of any tool is how we use and interact with it and the emotional responses we experience while we interact with it, either physically or consciously. The focus of this book series is to provide a variety of conceptual, theoretical, and practical perspectives on the role of emotions and technology. Various psychological and social-emotional aspects of communicating through and with many types of technology are engaged in ways that extend our understanding of technology and its consequences on our lives.

A specific goal and purpose of this book series focuses on emotions and affective interactions with and through technology. In some cases, these interactions are user-to-user, supported by the technology. In other instances, these interactions are between the user and the technology itself. Let us take, for example, researchers who have used animated social simulation technology to measure emotions of educators (Tettegah, 2007) and others, who use biotechnology to measure decision-making and emotional responses of users of technology (Baron-Cohen, 2011; Decety & Ickes, 2009). In a recent article, Solomon (2008) points out, “One of the most critical questions about human nature is the extent to which we can transcend our own biology (p. 13).” I would argue that through our use of technology we, in fact, are attempting to extend and transcend our emotions by way of robots and other intelligent technological agents. As such, we should then ask ourselves: why are discussions of emotions and technology so important?
Inquiry regarding the nature of emotions is not new. In fact, examples of such forms of inquiry have been documented since the dialogues of Socrates and Plato. Researchers and practitioners in psychology, sociology, education, and philosophy understand the complicated nature of emotions, as well as [the importance of] defining emotions and social interactions. The study of emotions is so complicated that we still continue to debate within the fields of philosophy, education, and psychology, the nature of emotions and the roles of affective and cognitive processes involving human learning and behavior. The volumes in this series, therefore, seek to present important discussions, debates, and perspectives involving the interactions of emotions and various technologies. Specifically, through this book series on Emotions and Technology, we present chapters on emotional interactions with, from, and through technology.

The diversity of emotions, played out by humans with and through technology run the gamut of emotions, including joy, anger, love, lust, empathy, compassion, jealousy, motivation, frustration, and hatred. These emotional interactions can occur through interactions with very human-looking technologies (e.g., avatars, robots), or through everyday commonplace technologies (e.g., getting angry with an ATM machine when the user fails to follow instructions). Hence, understanding the ways in which technology affords the mediation of emotions is extremely important toward enhancing our critical understanding of the ways in which student minds, through technology, are profoundly involved in learning, teaching, communicating, and developing social relationships in the twenty-first century.

The majority of the chapters presented in the books in the series will no doubt draw on some of the recent, pervasive, and ubiquitous technologies. Readers can expect to encounter chapters that present discussions involving emotions and mobile phones, iPads, digital games, simulations, MOOCs, social media, virtual reality therapies, and Web 2.0/3.0 technologies. However, the primary focus of this series engages the readers in psychological, information communication, human computer interaction, and educational theories and concepts. In other words, technologies will showcase the interactions, however, the concepts discussed promise to be relevant and consistent constructs, whether engaging current technologies or contemplating future tools.

The whole book series began with a call for a single volume. However, there was such a huge response, that what was to be one volume turned into eight volumes. It was very exciting to see such an interest in literature that lies at the intersection of emotions and technology. What is very clear here
is that human beings are becoming more and more attached to digital technologies, in one form or another. In many ways, we could possibly posit the statement that many individuals in the world are inching their way toward becoming cyborgs. It is apparent that digital technologies are in fact more and more second nature to our everyday life. In fact, digital technologies are changing faster than we are aging.

The life of a new technology can be 6 months to 1 year, while the human lifespan ranges from 0 to 80+ years. With the aforementioned in mind, humans have to consider how their emotions will interact and interface with the many different technologies they will encounter over the course of such a lifetime. It seems as if it were only yesterday that the personal computer was invented and now we have supercomputing on a desktop, billions of data at our fingertips on our smartphone computers, and nanotechnology assisting us with physiological functions of living human animals. Regardless of the technology we use and encounter, emotions will play a major role in personal and social activities.

The major role that technology plays can be observed through the many observations of how humans become excited, frustrated, or relieved, when interacting with new technologies that assist us within our daily activities.

Our hope is that scholars and practitioners from diverse disciplines, such as: Informatics, Psychology, Education, Computer Science, Sociology, Engineering and other Social Science and Science, Technology, Media Studies and Humanities fields of study, will find this series significant and informative to their conceptual, research, and educational practices. Each volume provides unique contributions about how we interact emotionally with, through, and from various digital technologies. Chapters in this series range from how intelligent agents evoke emotions; how humans interact emotionally with virtual weapons; how we learn or do not learn with technology; how organizations are using technology to understand health-related events; to how social media helps to display or shape our emotions and desires.

Sharon Tettegah
University of Nevada, Las Vegas, USA

ACKNOWLEDGMENTS
I would like to give a special thank you to Martin Gartmeier, Dorothy Espelage, Richard Ferdig, WenHao David Huang, Grant Kien, Angela
Benson, Michael McCreery, Safiya Umoja Noble, Y. Evie Garcia, and Antonia Darder, and all of the authors for their reviews and contributions to this work.

REFERENCES


One key question investigated in educational technology research is how the design of instructional environments affects learners and learning processes. Thereby, the main focus in this volume is on the development and design of deep-learning processes, through which students can increase their knowledge and skills in complex and challenging subject areas, e.g., in academic disciplines (Calvo, D’Mello, Gratch & Kappas, 2015; Graesser, D’Mello & Strain, 2014). Up to now, this research has pursued a primarily cognitive focus while emotional and affective aspects of such learning have widely been neglected for the most part (Leutner, 2014; Park, Plass, & Brünken, 2014), with the exception of the field of affective computing (Calvo, D’Mello, Gratch & Kappas, 2015). In the area of emotions, design, and learning, this is evident when looking at the most influential theories in use in educational technology research, such as the cognitive theory of multimedia learning (Mayer, 2005), the multimedia principle (Fletcher & Tobias, 2005), or the integrated model of text and picture comprehension (Schnotz, 2005). These approaches have allowed for substantial improvements in our understanding of technology-based learning, yet without taking the role of learners’ emotions into account.

One of the most influential and widely cited publications in this field of educational research is the Cambridge Handbook of Multimedia Learning (Mayer, 2005). The subject index of that publication features no entries for “emotions,” “feelings,” or “affect.” However, interestingly, in the second edition of the same handbook (Mayer, 2014), these three keywords appear in the index at least once (“emotional design,” “affect in learning with micro-worlds,” “affect in learning with simulations,” “feeling of knowing”). This shows that research on the interplay between emotions, instructional design, and learning has still not yet been established as a major strand in instructional and educational technology research. However, the recent publication of pertinent articles (Calvo, D’Mello, Gratch & Kappas, 2015; Graesser, D’Mello & Strain, 2014) and special issues (e.g., Leutner, 2014) shows that the topic is gaining momentum. For several reasons, this can be seen as a necessary amendment of research on technology and design-based learning.
First, the importance of taking emotions into account when analyzing any learning process has long been recognized in educational research. It has been argued that an overly cognitive approach to analyze learning treats emotions as nothing more than a “commotion—an unruly inner disturbance” (Scheffler, 1977, as cited in Solomon, 1992, p. 45). Such a perspective not only neglects an important aspect of technology-based learning processes, but also has been described as being “utterly destructive of education” (Solomon, 1992, p. 46) in general. This is because it leads to educational processes artificially being separated in grotesque parts, such as “unfeeling knowledge and mindless arousal” (Scheffler, 1977, as cited in Solomon, 1992, p. 45).

Second, although substantial research on the role of emotions in education exists (see Pekrun, 2014), these aspects have so far “largely been neglected in research on technology-based learning and instruction” (Leutner, 2014, p. 174). This is disappointing, as scholars have long been arguing in favor of learners bringing their social assumptions to all “interactions with computers, television, and new media,” making these “fundamentally social and natural, just like interactions in real life” (Reeves and Nass, 1996, p. 5, original emphasis). The point made here is that many assumptions we have about learning, cognition and interaction apply equally well to real-life as well as to mediated contexts. Drawing upon this assumption, it seems simplistic to analyze technology-based learning without taking the feelings of learners into account. Given the large number of publications, which have drawn upon this point made by Reeves and Nass (1996, e.g., Clark & Mayer, 2011; Woolf, 2009), it seems it is about time to move the role of emotions into the focus of research on design and technology-based learning more strongly.

Third, even in studies, which do take the role of emotions for learning into account, emotions often, “are filtered through the notion of the expected utility function and a cognitive evaluation process” (Perlusz, 2004, p. 1). In fact, emotions are far more than just mental background images of our cognitive processes. One interesting piece of evidence relating to this point comes from research on decision-making of individuals who suffer from their emotions being impaired by brain lesions. Pertinent studies have shown that the inability of such individuals to make good decisions, to some degree, relates to the absence of input and guidance from an emotional aspect of their consciousness (Damasio, 2006). This underlines that emotions are crucial aspects of higher order processes, such as decision-making or learning, no matter if these occur off- or online. Hence, in order to improve
our understanding of the mechanisms behind the design of technology-based learning, focusing on the role that emotions play therein is promising.

**Envisioning the Empathic Learning Environment**

Research on the interplay between emotions, technology, learning, and design is still in an early stage. In an attempt to sketch the goals and promises behind pursuing this line of research, Graesser, D’Mello and Strain (2014, p. 473) have envisioned that, “in an ideal world, the computer system would put the student in a zone of optimal concentration that targets relevant knowledge about the subject matter, at a pace that delivers the right challenges to the particular student at the right time.” In most technology-based learning environments, the learning goals are cognitive. For successfully reaching these goals, however, the emotions a learner experiences during learning play an essential role. The higher-order goal pursued in research on emotions in e-learning is to gain knowledge about how to design learning environments in ways that allow for positive emotional experiences during learning. In this way, deeper and more persistent learning processes shall be facilitated. This idea has been expressed as the *affective mediation assumption* of learning behavior (Leutner, 2014). This means that the intensity and persistence with which learners cognitively engage with any learning program is mediated by emotional factors.

In order to achieve this goal, it is important to understand ways in which emotions matter in the design of technology-based learning and how research may contribute to make technology-based learning more emotionally sensitive. Mayer (2005) has differentiated between *basic* and *advanced principles* of multimedia learning: the level of *basic principles of multimedia learning* concerns basic design features of learning environments, e.g., regarding the integration of different media, the coding of information in words and pictures (modality), or in pictures and narrated text (redundancy). Moreover, the segmentation of information and the existence of cues regarding the importance of different materials matter. For many of these principles, substantial evidence exists today for the impact they have on cognitive learning outcomes. However, we do not yet know much about the influence of these aspects on the emotions of learners, which, in turn, are a mediator for cognitive learning outcomes. Pertinent studies are Magner et al.’s (2013) investigation of the influence that decorative illustrations in learning environments have on learning or Dunlap’s chapter on emoticons in the present volume. In contrast, *advanced principles* of multimedia learning concern more sophisticated aspects of instructional design, such as worked-out examples,
collaboration, animation, and interactivity, or sensitivity of learning environments to learners’ level of knowledge. Again, focusing the emotional besides the cognitive level adds a new dimension to all of these principles.

However, what may be the most ambitious vision connected to researching emotions in technology-based learning is to create what could be called an “empathic” learning environment. Such an emotionally sensitive learning environment may have been enabled technologically to detect the emotions learners experience and to provide state-specific instructional support, e.g., in the form of comments from an intelligent animated agent. This scenario comprises two emotion-related technological functions, the detection of emotions a learner experiences as well as the algorithm-based reaction to these by the technology. A number of promising examples already exist today (Graesser, 2011).

In the area of advanced design principles, a further promising perspective is to design learning environments in ways that bring learners to empathize with technologically presented content. Imagine someone learning to become a physician, being confronted with an interactive, animated patient who describes his medical condition. Such a simulation could, if perceived as being realistic, induce vicarious emotional experiences in a learner. Moreover, coping with these vicarious emotions and reacting appropriately to the patient are important skills of physicians that could effectively be trained by means of highly developed, “empathic” educational technology. Of course, it has to be said that the empathy any machine can be programmed to exhibit is different from the empathy human beings are able to sense. However, we argue that the interactive, computer-based learning systems will be programmed to have some form of emotional intelligence in the future.

The present volume assembles research from the fields of psychology, computer science, instructional design, education, and learning sciences research, which advances our understanding of the interplay between emotions, technology, learning, and design. The chapters in this book can be grouped into three major strands of inquiry: the first section of this volume investigates the interplay between emotions and affect recognition systems in technology-based learning. The second section features reviews on emotions, affect, and design features of learning environments. The third section focuses on the role of interactions and design for learning.

EMOTIONS AND AFFECT RECOGNITION SYSTEMS

In Chapter 1, Art Graesser and his colleagues focus on intelligent tutoring systems (or “agents”) that adaptively communicate with learners in natural
language in order to optimize their learning. The learners in focus here are adults with low literacy skills seeking to improve in this area. The chapter describes how the emotions that these learners experience are tracked by an intelligent computer system. Different strategies of how the intelligent agents react to these emotions are also described.

In Chapter 2, Ling Cen and colleagues explore the ability of computer systems to recognize emotions from continuous speech. They describe an experiment involving an online learning environment equipped with a real-time speech emotion recognition system. The results show the system’s capacity to reliably recognize students’ emotional responses to the course. These results are relevant with regard to the customization of online courses to fit students with different learning abilities and to help students achieve optimal learning performance.

Tying in with the inquiry on pedagogical agents, in Chapter 3, Chad Lane discusses the question of how these impact on learners emotionally. To address this question, the empirical literature on pedagogical agents is reviewed in this chapter, where the focus is on emotions during and after learning interactions. The chapter concludes with a discussion of suggestions and ideas for promising directions of research that contribute to more fully understanding the apparent strengths of pedagogical agents.

Closing the first section of the present volume, Chapter 4, by Bogdan Pătruț and Roxana-Petronela Spatariu, focuses upon how pedagogical agents, which are designed to express emotions can bring benefits to educational processes. Drawing upon an initial literature review, the authors argue that pedagogical agents can encourage learners to care more about their progress and increase their experience of positive emotions. The authors present their own approach to the implementation of responsive agents at the end of the chapter.

REVIEWS ON EMOTIONS, AFFECT, AND DESIGN

The second section of the present volume features reviews dedicated to elucidate the interplay between emotions, affect, and design.

Chapter 5 by Jason Harley opens this section. This chapter provides an interdisciplinary overview of different methods used in research with computer-based learning environments to measure learners’ emotions. These methods include the coding of facial expressions and body posture, physiological variables, log-file data, and self-report measures. The review is guided by key questions, e.g., What are the factors that should inform the selection of different affective methodologies?
Chapter 6, by Yanghee Kim and her colleagues, is concerned with the use of advanced learning technologies in order to facilitate learners’ positive affect and, thereby, increase engagement and learning. This chapter introduces three tools that have been used to promote positive affect of learners: virtual peers (animated, on-screen characters), humanoid robots, and online videos. The central part of the chapter discusses how each of these technologies has been used to support positive affect in students who face various challenges in their learning.

Chapter 7, authored by Jan Plass and Ulas Kaplan, introduces an integrative approach to cognition and emotion aimed at understanding the role of emotional design in digital media for learning. They argue that this approach serves as an important step for making designs of multimedia learning more compatible with the real-life complexity of how the human mind functions and changes. Implications of this approach, which are discussed, affect the design and research on digital learning environments, such as simulations, games, and virtual worlds.

In Chapter 8, Joanna Dunlap and her colleagues investigate emoticons as visual tools, which instructional designers use to promote positive affect in learners. The chapter presents a review of the literature on emoticons used in support of online learning, with the goal of improving future practice in, and research of, online teaching and learning. Based on this review, instructional recommendations for online educators are provided.

The subsequent Chapter 9, by Patricia Vela and her colleagues, analyzes the role of emotions and learning in human–robot interaction. This chapter presents a theory-based introduction to the theory of emotions, to the value of emotions in education, and to the status quo of technology in education. Finally, it analyzes case studies in which emotions have emerged from robotics activities and from interactions with students.

INTERACTIONS, DESIGN, AND LEARNING

In the volume’s final section, the role of interactions and design for learning with digital media is discussed. In Chapter 10, Shangoon Park suggests affective scaffolding through virtual avatars (VAs; also known as pedagogical agents), as a way to improve learners’ emotional experience in online learning. Virtual avatars are relevant in the present context, as they can play the role of emotional facilitators by providing students with just-in-time emotional scaffolding. The authors review related literature that serves as the foundation of virtual avatar design for interest development.
Moreover, they discuss design factors for emotional scaffolding with VAs, by employing two communication modes: visual persona and verbal messages.

In Chapter 11, Enilda Romero-Hall focuses on the design of animated pedagogical agents in computer-based instruction and its effect in human-agent interactions. The chapter discusses recent research efforts on the effect of the emotional dimension of animated pedagogical agents on learning and on how research helps us understand which emotion elements are critical for the design of believable animated pedagogical agents.

Chapter 12, authored by Yungwei Hao, researches students’ sense of social presence when using different types of Web 2.0 technologies, namely blogs, wikis, social networking, social bookmarking, and virtual worlds. The chapter reports a study in which student teachers from Taiwan filled out questionnaires and took part in focus-group interviews. One result of the study is an analytic framework for Web 2.0 integration that considers students’ affective states while using the technologies.

Candace Sidner, in the final Chapter 13, focuses on collaborations and what is needed to make an intelligent computer agent a useful participant in collaborations. She discusses how non-verbal behavior plays a role in this respect and delves into emotional expression as a non-verbal behavior in collaborations, and what purpose it can and must serve to collaborators. As can be seen, emotional expression provides much more than a signal of the cognitive state of the expressor of the emotions.

Martin Gartmeier
TUM School of Education, TU München, Munich, Germany

Sharon Tettegah
University of Nevada, Las Vegas, USA

REFERENCES


CHAPTER 1

Emotions in Adaptive Computer Technologies for Adults Improving Reading

Arthur C. Graesser\textsuperscript{a}, Whitney Baer\textsuperscript{a}, Shi Feng\textsuperscript{a}, Breya Walker\textsuperscript{a}, Danielle Clewley\textsuperscript{a}, David P. Hays\textsuperscript{a}, Daphne Greenberg\textsuperscript{b}

\textsuperscript{a}University of Memphis, Memphis, Tennessee, USA
\textsuperscript{b}Georgia State University, Atlanta, Georgia, USA

One in six adults in the USA has low-level literacy skills (OECD, 2013). Some of these adult readers are from poor populations and others from countries with a different language. All face difficulties with daily literacy tasks (National Research Council, 2011). This chapter explores how intelligent tutoring systems (ITSs) with conversational agents (i.e., talking heads) have the potential to help them overcome these barriers in a way that is sensitive to both their cognitive and emotional states.

Imagine being in charge of helping a group of the struggling adult readers (hereafter called “adult readers,” as a technical term). One of the adult readers is 40 years old, with two full-time jobs, making less than the minimum wage. There is very little time to learn or practice reading. A second adult has many hours of free time because she is unemployed and trying to find a job, but has trouble comprehending and filling out the job application forms. A third adult comes from China, barely speaks English, and depends on relatives for support. A fourth adult is 18 years old, but has never received an adequate education in high school because he dropped out at the age of 15. He finds it difficult to pass the driving test to get a license to get to work. These four adults have very different profiles, but all of them experience stress as they attempt to reach their potentials in a print-rich environment.

Technology can rise to the occasion to help these adult readers. During the last 3 years, we have been developing an ITS with conversational agents to help these readers in our Center for the Study of Adult Literacy (CSAL, www.csal.gsu.edu), a major research effort that includes the University of Memphis, Georgia State University, University of Toronto, and Brock University.
The computer can recommend texts that fit with their interests and reading abilities; can help them communicate with peers for social interaction; and can deliver interventions that optimize learning. These computer applications are interactive and adaptive to the learner. Interestingly, the adaptivity is not confined to the knowledge, skills, and cognitive capacities of the learner. The computer application also adapts to emotions. All of the computer applications discussed in this chapter attempt to accommodate the emotions and moods of the learners.

This chapter describes our attempts to build affect-sensitive computer technologies that try to help these adults improve their reading comprehension. We have already documented, in previous publications, how we have developed, tested, and successfully applied affect-sensitive ITS to help college students learn STEM (science, technology, engineering, and mathematics) topics (D’Mello & Graesser, 2012; Graesser & D’Mello, 2012; Graesser, D’Mello, & Strain, 2014; Lehman et al., 2013). The first section summarizes this body of research. The second section describes how we have used what we have learned from college students on STEM topics, to develop computer applications for adult readers attempting to improve their reading comprehension.

In this chapter, we also focus on ITS applications with conversational agents (talking heads, hereafter called “agents”) that communicate with the learner in natural language. Agents have become increasingly popular in ITS and other contemporary learning environments. Examples of adaptive agents that simulate dialogs and that have successfully improved student learning are: AutoTutor (Graesser, 2011; Graesser et al., 2004, 2012; Nye, Graesser, & Hu, 2014); DeepTutor (Rus, D’Mello, Hu, & Graesser, 2013); Coach Mike (Lane, Noren, Auerback, Birth, & Swartout, 2011); Crystal Island (Rowe, Shores, Mott, & Lester, 2011); My Science Tutor (Ward et al., 2013); and Virtual Patient (Ferdig, Schottke, Rivera-Gutierrez, & Lok, 2012). These systems have covered topics in STEM (e.g., physics, biology, computer literacy), reading comprehension, scientific reasoning, and other difficult topics and skills. These systems engage the learner in a dialog, in which the human learner interacts with only one agent. The agent can be either a peer (approximately the same level of proficiency as the human); a student agent with lower proficiency (so that the learner can teach the agent); or an expert tutor agent. AutoTutor is a pedagogical agent that simulates the dialog moves of human tutors, as well as ideal pedagogical strategies. The approach of getting agents to simulate a tutor is a sensible first design for agents because human tutoring is known to be a very effective environment for improving student learning and motivation. Meta-analyses
that compare tutoring to classroom teaching and other suitable comparison conditions report effect sizes between $\sigma = 0.20$ and $\sigma = 1.00$ (Cohen, Kulik, & Kulik, 1982; Graesser, D’Mello, & Cade, 2011; VanLehn, 2011). Empirical evidence also supports the claim that AutoTutor and similar computer tutors with natural language dialog yield learning gains comparable with trained human tutors, with effect sizes averaging 0.8, range 0.3–2.0 (Graesser, 2011; Nye et al., 2014; Olney et al., 2012; Rus et al., 2013; VanLehn, 2011; VanLehn et al., 2007).

In addition to dialogs, there can be multiple agents interacting with the human. For example, three-party conversations, called trialogs, involve two agents and a human learner (Graesser, Li, & Forsyth, 2014). The two agents take on different roles, such as peers and tutors. For example, learners can observe vicariously two agents interacting, can converse with a tutor agent, while a peer agent periodically comments, or can teach a peer agent while a tutor rescues a problematic interaction. Agents can argue with each other over issues and ask what the human learner thinks about the argument. Examples of trialogs appear in the CSAL tutoring environments for adult readers, as well as several successful ITS with multiple agents, such as Betty’s Brain (Biswas, Jeong, Kinnebrew, Sulcer, & Roscoe, 2010); Tactical Language and Culture System (Johnson & Valente, 2008); iDRIVE (Gholson et al., 2009); iSTART (Jackson & McNamara, 2013; McNamara, O’Reilly, Rowe, Boonthum, & Levinstein, 2007); and Operation ARA (Forsyth et al., 2013; Halpern et al., 2012; Millis et al., 2011). There now follow some example trialogs that attempt to be sensitive to the affect states of adult learners.

**AFFECT-SENSITIVE ITSs FOR COLLEGE STUDENTS LEARNING STEM TOPICS**

An affect-sensitive computer technology requires a solution to three goals. First, it needs to identify what emotions, moods, or other affect states are experienced by the learners during learning. Second, it needs to track these affect states automatically during the course of learning. Third, it needs to respond to these affect states in a fashion that optimizes student learning and motivation. This section reviews the progress the field has made on these three solutions.

**What Affect States Do Learners Experience?**

During the last decade, researchers have vigorously investigated the moment-to-moment emotions that occur during learning with intelligent computer technologies (Baker, D’Mello, Rodrigo, & Graesser, 2010; Calvo & D’Mello,
The goal is to identify the learning-centered emotions, or what Pekrun (2006) calls achievement emotions, that routinely occur during learning tasks. We refer to these as emotions, even though they are amalgamations of cognitive and affective states that span from a couple of seconds to a minute or longer (D’Mello & Graesser, 2011). According to the meta-analysis of D’Mello (2013), the most common emotions are confusion engagement/flow, boredom, and frustration, with delight and surprise occasionally occurring but considerably less frequently. Anxiety also occurs when the learner is being assessed in examinations that are graded. It should be noted that these learner-centered emotions are very different than the six “basic” emotions investigated by Ekman (1992) that are readily manifested in facial expressions: sadness, happiness, anger, fear, disgust, and surprise. They are also different from the moods that occur over longer stretches of time.

Emotions nearly always accompany learning when students struggle with complex technical texts, challenging writing assignments, and difficult problems to solve. Many of the emotions are negative. For example, students experience confusion and frustration when they confront various obstacles in comprehension, production, reasoning, and problem-solving. Boredom occurs when the materials are either too easy or too difficult, which runs the risk of the learner’s disengagement from the learning environment and of mind wandering (Feng, D’Mello, & Graesser, 2013; Franklin, Smallwood, & Schooler, 2011). Fortunately, there are also positive emotions. Students experience curiosity and intense engagement when the topics interest them, and delight when there are deep insights and discoveries as challenges are conquered. Therefore, a mixture of positive and negative affective states is experienced during the moment-to-moment process of learning.

The duration and sequencing of these emotions has also been investigated. D’Mello and Graesser (2011) documented that delight and surprise are much shorter in duration than boredom, engagement/flow, confusion, and frustration. Regarding the sequencing of emotions, the sequences depend on what events occur during the course of learning, but a number of general sequences have been identified (D’Mello & Graesser, 2012; Graesser & D’Mello, 2012; McQuiggan et al., 2010). At various points during a typical learning activity, the student is in a state of cognitive and emotional equilibrium that corresponds to engagement and possibly flow. Then
some event, stimulus, or thought occurs that creates an impasse (i.e., obstacle, goal blockage) and the student experiences confusion (and sometimes surprise when the phenomenal experience is abrupt). If the confusion is resolved, the student returns to engagement/flow, thereby completing a healthy cycle, which may include delight when a difficult impasse is conquered. If the impasse is not resolved and the goal is persistently blocked, then the student experiences frustration. As the student struggles with the frustration, there invariably are additional impasses and resulting confusion. Persistent failure leads to boredom and eventual disengagement from the task. Boredom can also lead to frustration if the student is forced to endure a lengthy learning session. These sequential patterns are compatible with a cognitive disequilibrium model that D’Mello and Graesser proposed, to explain the occurrence, during, and sequencing of, learning-centered affect states.

A variety of methodologies have used human judgments, as opposed to automated computer tracking, to explore the learning-centered emotions in the above research (Baker et al., 2010; Craig, D’Mello, Witherspoon, & Graesser, 2008; D’Mello et al., 2009; Graesser & D’Mello, 2012; Woolf et al., 2009). These include:

1. Trained judges observe college students during the course of learning and make judgments on what emotions occur periodically.
2. The learner and the learning environment are video- and audio-recorded during the process of learning. These records are observed retrospectively for the occurrence of emotions, by the learner, peers, teachers, or trained judges.
3. The learners are asked to rate their emotions or make judgments of the valence (positive versus negative) and intensity of their affective states during the process of learning.
4. The learner “emotes aloud” during the process of learning by expressing whatever emotions, moods, or other affect states are experienced. These human judgments can be categorical judgments or ratings on particular emotions either periodically (e.g., every 20 s) or on occasions, when they are salient to the person making the judgment.

Of course, human judgments of emotions have some methodological limitations. The criteria that the various individuals use in judging the emotions may differ, so the reliability of the judgments is modest. The reliability scores on the occurrence of an emotion in a particular category (measured as Cohen’s kappa, which adjusts for chance and varies between 0 and 1), are between 0.2 and 0.5, depending on the emotion; the highest our research
has documented is 0.7 between trained judges in point (2), above (Graesser & D’Mello, 2012). When disagreements such as these occur, it is difficult to defend a particular gold standard on what affective state is correct when comparing points (1), (2), and (3), above. Some learners have a minimal sensitivity to their own emotions, so their judgments are not perfect. Trained judges show the best agreement, but that may be attributed in part to the systematic training routine. The emote-aloud judgments in point (4), above are problematic because some learners may have trouble expressing their emotions and the emote-aloud task may somehow alter the learning experience. Similarly, the online judgment task in point (3) may also alter the learning experience. For these reasons, researchers exploring learning in computerized learning environments have turned to automated, non-intrusive technologies to track affective states, as discussed below.

**Automated Tracking of Affective States During Learning**

The automated classification of affect states varies in the extent to which they are invasive. Noninvasive sensing methods do not attach sensing devices to the learner and do not disrupt the normal stream of learning by probing the students with questions about their emotions or learning. The computers detect emotions of the learner by analyzing different communication channels and their interactions (Calvo & D’Mello, 2010; D’Mello & Graesser, 2010; Picard, 1997). The common communication channels include facial expression (D’Mello & Graesser, 2010; Grafsaard, Boyer, Phillips, & Lester, 2011; Kapoor et al., 2007); speech parameters (Litman & Forbes-Riley, 2006); body posture (D’Mello, Dale, & Graesser, 2012); and language and discourse interaction (D’Mello & Graesser, 2012). The accuracy of these automated detection methods is modest, but so is the agreement among human judges, as discussed above. Some biological channels of affect sensing are minimally invasive, such as a wrist band, but most are maximally invasive, such as an EEG or fMRI. These methods include the recording of heart rate, movements of the muscles, galvanic skin response, and brain activity (Calvo & D’Mello, 2010). However, most of these methods are invasive, in the sense that it is obvious to the students that they are being recorded by physical instruments that have contact with their bodies.

Once again, there is no gold standard for measuring what emotions the learners are actually experiencing during learning, because all measures are imperfect windows into emotional experience. The various measures correlate only modestly (kappas ranging from 0.2 to 0.5, see D’Mello & Graesser, 2010), with each measure having both virtues and liabilities. In light of these
indeterminacies in measurement validity, researchers often collect multiple measures and adjust the confidence in their conclusions according to the consistency of the results.

Most of our work on automated emotion detection has concentrated on three channels: the discourse interaction history; facial actions; and body movements—and combinations of these three channels (D’Mello & Graesser, 2010, 2012). The discourse interaction history includes events stored in the *AutoTutor* log file, the speech acts of student and tutor turns, and the knowledge states achieved by the student during the tutorial dialog. An analysis of the discourse interaction history provides a model of the context of an emotional expression. The facial actions and expressions and body pressure measurement systems are tracked by particular systems that are beyond the scope of this chapter. The point we wish to convey here is that the discourse history goes a long way (90% or higher, compared with all other sensing devices) in classifying students’ affect states (defined as a composite of many measures). As one example, Graesser and D’Mello (2012) reported that dialog interaction history showed accuracies of 63%, 77%, 64%, 70%, and 74% (50% is chance), in discriminating confusion, frustration, boredom, flow, and delight from neutral. The average across emotions was 70%. If we were to transform these scores to values comparable with kappa scores [i.e., $2 \times \frac{\text{score} - 0.5}{}$], the quantities would be 0.26, 0.54, 0.28, 0.40, and 0.48, respectively, or 0.39 overall. Such kappa scores are comparable with accuracy scores reported by other researchers in the literature who have attempted automated emotion detecting systems. They are also comparable with the reliability of human judgments.

It is beyond the scope of this chapter to specify the mechanisms of the computer automatically detecting affective states from the different channels of communication (see D’Mello & Graesser, 2010, 2012, for a description of these mechanisms); however, we will present some highlights of the channel that involves the language and discourse interaction history. Advances in computational linguistics (Jurafsky & Martin, 2008) have made it possible to interpret students’ natural language by segmenting the language within conversational turns into segments (such as speech acts), classifying the segments into categories (such as questions, assertions, expressive evaluations, and other speech act categories), and performing semantic evaluations of the quality of the student assertions. Quality can be assessed automatically by matching the verbal input of the student to representations expected by the computer. The semantic match algorithms go beyond keywords and into inferences derived from large text corpora and higher dimensional
semantic spaces (Landauer, McNamara, Dennis, & Kintsch, 2007). The history of these interpreted interactions in learning sessions is stored in log files, so that data mining and machine learning analyses can be performed. This allows researchers to discover what language and discourse patterns are diagnostic of particular learner emotions.

According to D’Mello and Graesser (2010), the dialog cues that trigger the emotions are quite different for the different emotions. The cues that accompany confusion tend to be short student responses, frozen student expressions (such as “I don’t know”; “Uh huh”), speech acts by the tutor that are indirect (such as hints), and early time phases during the student’s initial attempts to solve the problem or answer the questions posed by the tutor. In contrast, the cues that accompany frustration are negative tutor feedback and student responses that are locally good ideas but not globally good ideas. Flow/engagement tends to occur with lengthier answers, during early phases of the dialog, and after positive tutor feedback. Boredom tends to occur in later phases in the session or a particular problem and when the tutor tends to lecture with direct assertions. These dialog cues are important when we track emotions of adult learners in our conversational learning environments.

How Does the Computer System Respond to Learner’s Affect States?

Detecting an emotion is essential in an affect-sensitive learning environment, but it is also essential to respond in a manner that optimizes learning and motivation. One place to start is to consider what affect states are most predictive of learning. This question has indeed been explored in studies that consider both correlational and causal relationships between the affect states and learning. Graesser and D’Mello (2012) reported significant correlations between learning gains and some emotions. More specifically, there were significant positive correlations with confusion and flow/engagement, but a negative correlation with boredom, and zero correlations with delight and frustration. Confusion was the best predictor of learning gains among the various emotions investigated, which is consistent with the cognitive disequilibrium already discussed. The fact that learning correlates negatively with boredom and positively with flow/engagement are consistent with predictions from Csikszentmihalyi’s (1990) analysis of flow experiences. Moreover, a series of studies have shown that confusion has a significant mediating role in causing learning gains (D’Mello, Lehman, Pekrun, & Graesser, 2014; Lehman et al., 2013). It takes a minimal amount of knowledge to be confused and those who deliberate over the confusion tend to learn more than those who do not
deliberate or who have insufficient knowledge of what the issues are. In summary, an important step in predicting learning is a combination of sufficient knowledge, experiencing cognitive disequilibrium, and having a deep engagement in the material to the point of being in a “flow” experience.

Our contention is that ITS can be engineered to optimize learning in an affect-sensitive manner. Some of the ITS would include agents to guide the process of the learner interacting with ITS systematically. For example, Graesser, Hu, Nye, and Sottilare (2015) identify the following techniques to promote motivation:

1. Offer texts, multimedia, simulations, or games that deliver information, tasks, and scenarios at the student’s zone of optimal challenge; not too easy or too difficult but just right.
2. Provide an engaging story narrative to sustain interest and coherence. The narrative would ideally be integrated with the academic subject matter and promote its value.
3. Reward the student with points/resources that are extrinsically reinforcing or with experiences that are intrinsically reinforcing.
4. Give the active student control over the interaction in order to allow autonomy and self-regulation.
5. Give the insecure student materials he/she can successfully master in order to build confidence and self-efficacy.
6. Interact with the student in an adaptive, turn-by-turn conversation or collaboration to promote interactivity and social presence.
7. Give the student timely feedback on his/her actions, so it is clear where the student stands in mastering the complex material.
8. Give feedback and guidance on the student’s emotions, so he/she can monitor the coordination between emotions and learning.

Other researchers also have lists of these techniques, but it is beyond the scope of this chapter to provide a comprehensive coverage of approaches to optimize motivation and emotional experiences. The next section describes our approaches to help adult readers through intelligent learning environments with dialogs and trialogs.

**BUILDING ITS WITH DIALOGS AND TRIALOGS FOR STRUGGLING ADULT LEARNERS**

This section identifies methods we have implemented to help adults read at deeper levels of comprehension. These methods include principles of intelligent adaptivity, conversational dialogs and trialogs, and games, in a manner that accommodates the affective states of the learners. These systems have
not been fully tested on adult readers, but they are based on a body of ITS research summarized in the previous section. We first focus on dialogs and then trialogs. For each, we describe how affect-sensitivity, intelligent adaptivity, and game design can potentially enhance learning and motivation.

**Dialogs**

*AutoTutor* implements the mechanisms of human tutoring that have been documented in previous research (Cade, Copeland, Person, & D’Mello, 2008; Graesser et al., 2011; Graesser, Person, & Magliano, 1995). A typical tutor presents example texts, problems to solve, or difficult questions for the two to collaboratively work on. By “collaborative,” we mean that the student and tutor take turns talking and performing actions until a correct set of responses eventually emerges. The primary mechanism is *expectation and misconception tailored* dialog in both human tutors and *AutoTutor* (Graesser et al., 2012; Graesser, Li, et al., 2014). There is a set of expectations (i.e., sentences, phrases, quantities, or other good answers) that the tutor wants covered and a set of anticipated misconceptions (i.e., bugs, errors, false claims) that are corrected along the way. A poor tutor simply lectures by always articulating or otherwise expressing the expectations. A better tutor tries to get the students to be more active by expressing information or showing their work. That requires dialog moves, which encourage the students to take some initiative. Eventually, most or all of the expectations are covered, so the example text, problem, or question is considered completed.

There are different categories of dialog moves that occur in both human and computer tutoring. The primary categories are listed below, they are: short feedback, pumps, hints, prompts, prompt completions, assertions, summaries, mini-lectures, corrections, answers, and off-topic comments.

*Short feedback.* The feedback is either positive (“yes,” “correct,” head-nod); negative (“no,” “almost,” head-shake, long pause, frown); or neutral (“uh huh,” “okay”).

*Pumps.* The tutor gives non-directive pumps (“Anything else?” “Tell me more”) to get the learner to do the talking or to take some action.

*Hints.* The tutor gives hints to get the learners to do the talking or take action, but directs the learners along some conceptual path. The hints vary from being generic statements or questions (“What about X?” “Why?”), to speech acts that nudge the learner toward a particular answer. Hints promote active student learning within the boundaries of relevant material.
Prompts. The tutor asks a very leading question in order to get the learner to articulate a particular word or phrase. Sometimes learners say very little, so these prompts are needed to get the learner to say something specific.

Prompt completions. The tutor expresses the correct completion of a prompt.

Assertions. The tutor expresses a fact or state of affairs.

Summaries. The tutor gives a recap of the answer to the main question or the solution to the problem.

Mini-lectures. The tutor expresses didactic content on a particular topic.

Corrections. The tutor corrects an error or misconception of the learner.

Answers. The tutor answers a question asked by a learner.

Off-topic comments. The tutor expresses information unrelated or tangentially related to the subject matter.

These different categories have different functions. Short feedback and corrections function to ensure that the information covered is accurate. Pumps, hints, and prompts are attempts to get the student to do the talking and doing. In contrast, prompt completions, assertions, summaries, and mini-lectures deliver correct information to the student. There is a continuum of dialog moves from attempts to get the student to express information at one end, to the tutor expressing information at the other end: pumps < hints < prompts < prompt completions < assertions < answers < summaries < mini-lectures. Active student learning is normally encouraged so the distribution of these dialog moves reflects the ability of the student and/or the tutor’s attempts to encourage active student learning. The off-topic comments typically serve as a motivational function. To recap, the main pedagogical functions of these discourse moves are to promote information accuracy, active student learning, and student motivation.

There are additional, less frequent, dialog moves that serve other functions. Floor management is an important function that the tutor uses to pass the turn to the student. For example, a tutor’s request for summary (“Please summarize the answer to the question”) encourages the student to give the summary rather than the tutor, which is a move that also promotes active student learning. Instead of the tutor answering a student question, the tutor can request the student to answer the question in a counter-question move (“How would you answer your question?”). Similar floor management and promotion of active student learning functions occur in the tutor’s request for student to read moves (“Please read the text aloud”) and request for correction moves (“Explain how this
answer is incorrect”). Unfortunately, human tutors rarely perform these functions that explicitly pass the floor to the student instead of the tutor providing this information. These dialog moves also serve the function of promoting self-regulated learning in the student, a function that human tutors rarely exhibit in tutoring protocols (Graesser et al., 1995). The main method of the tutor passing the floor to the student is through pumps, hints, and prompts.

Below is an example dialog that a tutor might have with an adult reader. This is a fictitious dialog that illustrates the discourse patterns of the students:

Move 1. **Tutor**: Please read the following text. [Request for student to read].

Move 2. **Student**: [Student reads the text about treating a burn wound and the tutor occasionally corrects the student’s mispronounced words].

Move 3. **Tutor**: Very good. [Positive short feedback] What is the topic of this text? [Prompt].

Move 4. **Student**: About burns.


Move 6. **Student**: I don’t know.

Move 7. **Tutor**: What about this sentence? [Hint].

Move 8. **Student**: It says to put on medicine.

Move 9. **Tutor**: Not really. [Short negative feedback] What do you do with the ice? [Prompt].

Move 10. **Student**: Put it on the burn.

Move 11. **Tutor**: Right. [Short positive feedback] You put the ice on the burn. [Prompt completion].

... (later).

Move 12. **Tutor**: Now let’s recap what the text is saying … [Summary].

This excerpt illustrates a typical dialog that a human tutor might have with an adult reader. There are a number of ways that this dialog could be modified, however, in order to be more affect-sensitive and to enhance motivation. We now turn to some techniques designed to improve this dialog to improve the affective dimension as well as learning.

**Give Students Choices**

The tutor could have given the student the opportunity to select the text in Move 1, instead of assigning it. Affect and motivation are enhanced when students are given choices (Deci & Ryan, 2002; Lepper & Woolverton, 2002; Pekrun, 2006)—one of the well-known principles of games...
(Tobias & Fletcher, 2011)—and thereby they can follow their interests (Tobias, 1994). A choice among 2-3 unattractive alternatives can even improve their motivation.

**Assign Texts Within the Student’s Ability**

Adult readers have limited self-efficacy on reading tasks and sometimes low self-esteem, so it is wise to assign texts well within their zone of proximate development. Computers can automatically scale texts on difficulty level (Graesser, McNamara, et al., 2014; Nelson, Perfetti, Liben, & Liben, 2011), whereas adult readers are routinely classified on their grade level of mastery (Greenberg, 2008). Therefore, the grade level of the text should be lower than the grade level of the reader, or, at most, slightly above their level. Lepper and Woolverton (2002) have also proposed a clever technique in which the tutor expresses that the task is difficult (even though the task is actually well within the student’s zone of mastery), but still encourages the student to give it a try. When the student readily handles the task, the student gets a boost in self-efficacy and self-esteem. It should be noted that students with higher ability and confidence would not benefit from this approach, but instead would be assigned tasks that present challenges and that extend the boundaries of their zone of proximal development (Brown, Ellery, & Campione, 1998; Vygotsky, 1978). An intelligent, adaptive assignment of texts at the right difficulty level should help affect, motivation, and learning.

**Give Supportive Short Feedback**

The wording of the positive, neutral, and negative feedback (Moves 3, 5, 9, 11) should be tailored to the ability levels and personality of the student. Human tutors tend to be reluctant to give negative feedback, even when the students’ contributions are incorrect (Graesser et al., 1995). They tend to be polite rather than giving face-threatening feedback that runs the risk of lowering the student’s confidence. D’Mello and Graesser (2012) have reported that students with low knowledge learn STEM topics better with supportive empathic feedback (e.g., “I know the material is difficult,” “Most students miss this,” “You can get this if you keep trying.”) more than short succinct unemotional feedback. (“That’s incorrect.”) In contrast, high knowledge learners actually showed lower learning from these empathic short responses, perhaps because they appear artificial or patronizing. Once again, intelligent adaptation of the short feedback to the learner is important.

We once created a tutor that gave playfully rude feedback by simply changing the wording of the short feedback. For example, the following...
short feedback was given for negative and positive short feedback, respectively: “You’re obviously not as smart as I thought you are” and “Aren’t you the little genius.” The rude tutor was much more engaging to knowledgeable adults than was the unemotional prude tutor, but most low knowledge learners would presumably not appreciate this approach. We suspect that the personality profile of the computer tutor needs to mesh with the attitudes of the learner. Some learners will be insulted by the rudeness, crudeness, and attempts at humor that teases the learner. The impact on learning, however, remains an open question.

**Gauge the Prospects of Active Learning**

Dialog moves 1, 3, 5, 7, and 9 are all geared to encourage the student to do the talking. This effort is noteworthy under the banner of active learning, but what if the learner is incapable? It is important to carefully gauge how verbose and accurate the student is in their answers so that they will not be dispirited from inadequate performance. Nothing worthwhile will be gained by repeatedly insisting that a low knowledge, low verbose student generates academic material. They need to start slow, starting with yes/no, then a word, then a phrase, then a sentence, and then a paragraph. They need to have models of good performance, as addressed in the next section. On the other hand, students with higher knowledge and verbosity should be encouraged to express themselves. We also know that coherent, verbose expressions are highly diagnostic of deep engagement and flow experiences (D’Mello & Graesser, 2012). Once again, the automated tutor needs to be adaptive.

**Gauge Who Should Summarize**

We know that summaries and reflections are excellent vehicles for learning. But who should provide these verbal expressions? The student or tutor (see Move 12)? This is where precise, dynamic, adaptive adjustments need to be made. The tutor needs to gauge when the student is stumbling and the tutor needs to step in. When does confusion lead to frustration, then to boredom, then to disengagement? It is a complex balance between the student and tutor providing the information.

**Trialogs**

A second agent, in trialog conversations, can enhance pedagogical agent design (Graesser, Li, et al., 2014). This involves two computer agents interacting with a human. Multiple agents have been incorporated in many learning environments with agents, as discussed earlier in this chapter. The major
advance in our recent research is how trialogs can be analyzed systematically to optimize a number of principles of learning, affect, and motivation.

Below are some trialog designs that can be productively implemented for adult learners. These designs can be contrasted with the more sophisticated designs for more advanced learners.

Design 1. *Vicarious learning with human observer.* Two agents interact and model ideal behavior. The two agents can take on different roles. Low skilled learners particularly benefit from this.

Design 2. *Vicarious learning with limited human participation.* The same as point (1) except that the agents occasionally turn to the human and ask *prompt* questions in order to promote engagement and assess the human’s comprehension.

Design 3. *Tutor agent interacting with human and student agent.* There is a tutorial dialog with the human, with the student agent contributing and receiving feedback. The agent takes the heat for bad answers that mirror the human answers.

Design 4. *Expert agent staging a competition between the human and a peer agent.* A competitive game (with score points) occurs between the human and peer agent, as the tutor agent guides the competition.

Design 5. *Human teaches/helps a student agent with facilitation from the tutor agent.* The human teaches the student agent as the tutor agent rescues problematic interactions. Higher skilled learners might particularly benefit from this.

Design 6. *Human interacts with two peer agents that vary in proficiency.* The peer agents vary in knowledge and skills that are adjusted according to the learner’s skill level.

Design 7. *Human interacts with two agents expressing contradictions, arguments, or different views.* The discrepancies between agents stimulate cognitive disagreement, confusion, and potentially deeper learning, but this may be beyond the zone of adult readers.

The vicarious learning designs (Designs 1 and 2, above) are appropriate for learners with limited knowledge, skills, and actions, which is the population of interest in this chapter. Design 4 is motivating for learners by virtue of the game competition, whereas Design 3 minimizes negative feedback to the human. These are the designs relevant to adult readers. The others are relevant to more high skilled learners.

An example trialog should clarify the nature of the discourse. This is a conversation with an adult learner interacting with two agents in a trialog. It should be noted that these readers have difficulty writing, so it is beyond the abilities of most of them to type much verbal information. The best
many of them can do is click on multiple choice alternatives, drag and drop information, or toggle on alternatives. In the example in Figure 1.1, Christina is the tutor agent, Jordan is the student agent, and the human is Jack. Christina and Jordan are talking heads at the edge of the main interface. The design team decided to put these agents as “guides on the side” rather than “sages on the stage” because they wanted the focus of attention to be on the interface with texts and multimedia rather than having an intense direct focus on the conversation with the agents.

The trialogs incorporate features designed to optimize affect-sensitivity, motivation, and learning. As in the case of the dialogs described earlier, the designer builds in student choice, assignment of text at the right difficulty level, and a judicious implementation of short feedback, active student learning, and summarization. For example, the first text to read in a lesson starts out at a medium level of difficulty, whereas a subsequent text is easier or more difficult, depending on student performance in the first text. Regarding short feedback, the tutor agent never blames the student by saying “You are incorrect” but rather attributes the problem to the selection (“That’s incorrect”), highlights the option in red, or asks the student agent for their views (see answer (8)). Regarding active student learning, students are not expected to write a lengthy summary of several sentences, which is well beyond their zone of proximal development. Instead, there is a gradient of active learning as one processes along the following input options: (1) click on a word in a text to answer a question; (2) click on a sentence in a text when asked a question; (3) click an option in a multiple choice question with sentences; (4) drag and drop a sentence from text into a particular category (e.g., designating whether a sentence is a fact, an opinion, or an emotion trigger?); (5) rate the quality of a text summary; (6) type in a word to answer a question; and (7) type in a sentence that reflects an inference about the text. At this point, we are varying these seven student input options among the various questions in the lessons. We can thereby investigate performance on these questions among students with different ability levels. With these data to hand, in a later version of the lessons, we can intelligently assign input categories at or near each student’s zone of proximal development.

The trialogs in CSAL’s AutoTutor incorporate design features that go beyond dialog. Already mentioned is the game competition, in which the human student and student agent compete on a task and accumulate points, as in, e.g., in Jeopardy. In this game mode, we dynamically select the student agent’s answers throughout the lesson so that the human always wins or ties;
(1) Christina (Tutor Agent): Sam is preparing for an interview and he is not sure about how he should look. Based on this section, Jack, what should Sam do?

(2) Jack (Human Student): [Clicks on first answer option and is correct]

(3) Christina: Great! The correct answer is that Sam should look his best for the interview. We know this because it is supported by statements in the text.

(4) Jordan (Student Agent): You mean like when Sam's friend Earl says, you better make sure you clean up?


(6) Christina: Which of these supports the statement, Sam should look his best for the interview?

(7) Jack: [clicks on the second sentence that starts with ‘They all joked…’ and is incorrect]

(8) Christina: Okay. What do you think Jordan? Which sentence gives Sam advice on how he should look?

Figure 1.1 Example trialog with an adult learner.
the goal is to boost self-efficacy and self-esteem in these adult readers. In a helping mode, the student agent is having trouble with a task (such as comprehending a job application form) and turns to the human by asking questions for help and the tutor agent steps in for additional help as needed; the fact that the human can help another person once again serves the goal of boosting the human student’s self-efficacy and self-esteem. The helping mode is presumably more motivating than a grill-test mode in which the tutor assesses the human student with frequent questions and feedback. It is also an improvement over the lecture mode in which the tutor and student agents take turns lecturing to the human student. Yet another mode is collaborative problem-solving, where the student agent, tutor agent, and human collaboratively solve a practical problem, such as filling out a medical form. We are currently exploring a variety of modes that go well beyond the “schoolish” lecture and test modes that are not particularly well suited to affect and motivation.

At this point in the CSAL project, we have nearly completed development of the first version of 30 lessons with the AutoTutor triologs. The topics of the lessons target comprehension skills that range from figuring out the meaning of a word from context to making plausible inferences, asking good questions, and detecting a good summary. We are empirically evaluating each lesson on dimensions of affect, motivation, and learning. Affect is assessed by having the adult learners rate their experience on scales of affect, perceptions, and attitude (e.g., how much they like the learning environment, how useful they find it), as well as recording and coding their facial expressions at different points in the lesson. For example, we can see when they are confused, frustrated, bored, disengaged, surprised, or delighted (D’Mello & Graesser, 2012; Graesser & D’Mello, 2012). Motivation is assessed by rating scales, time spent on different phases of the lesson, and dropout rates. Learning is assessed by performance on the questions and other tasks during the lessons, as well as objective pre-tests and post-tests. Our hope is that our attempts to optimize emotions during learning will result in higher motivation and learning. Of course, it remains to be seen whether our designs are successful in this population of adult readers.

**CHALLENGES AND LIMITATIONS**

The agent environments with dialogs and triologs hold considerable promise in helping students learn in ways that are sensitive to their cognitive, emotional, and motivational states. However, there are a number of challenges and
limitations in these environments that are worthwhile to pursue in future research. We want to end by addressing two of these challenges and limitations.

One challenge is that, sometimes, the human learner is not well “understood” by the computer. Such comprehension gaps can be irritating to the learner—so irritating that the learner disengages and gives up. Students discover this when they receive incorrect short feedback from the computer tutor or when the tutor’s response does not coherently address what the learner just expressed. Although the computer’s assessment of student quality is surprisingly accurate, there are errors and there is a risk of the learner’s blaming the computer. Interestingly, human tutors also misunderstand the learner (about as often as the computer tutor!), but there is a blind spot to such errors because the human tutor is assumed to be competent and human tutors are more adept at detecting and managing such communication gaps.

A second challenge is that a computer tutor cannot handle any student question that is asked and any topic that the student wants to pursue. Instead, the computer can only address the subject matter being tutored, i.e., there is very limited mixed-initiative communication. This can be frustrating to a student who wants to pursue a different topic, ask a genuine question, or pursue their interests at a rapid pace of inquiry. Perhaps this problem will disappear after thousands of ITSs are developed and are readily accessible. Until then, these agent-based ITSs tightly drive the agenda and learning experience, which limits self-regulated learning by the human.

This is a small selection of the limitations of the agent-based learning environments that help students learn by holding a conversation in natural language. We recognize many of the obstacles, but that should not undermine the enormous progress that has been made during the last two decades. We have adequately documented that these systems help learning in many populations of students, subject matters, and skills. The major challenge is to design them to optimize motivation and the emotional experience so that the students do not become bored, and disengage.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (SBR 9720314, REC 0106965, REC 0126265, ITR 0325428, REESE 0633918, ALT-0834847, DRK-12-0918409, 1108845); the Institute of Education Sciences (R305H050169, R305B070349, R305A080589, R305A080594, R305G020018, R305C120001); Army Research Lab (W911NF-12-2-0030); and the Office of Naval Research (N00014-00-1-0600, N00014-12-C-0643). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, IES, or DoD. The Tutoring Research Group (TRG) is an interdisciplinary research
team comprised of researchers from psychology, computer science, physics, and education, at University of Memphis (visit http://www.autotutor.org, http://emotion.autotutor.org, http://fedex.memphis.edu/iis/).

REFERENCES


CHAPTER 2

A Real-Time Speech Emotion Recognition System and its Application in Online Learning

Ling Cen\textsuperscript{a}, Fei Wu\textsuperscript{b}, Zhu Liang Yu\textsuperscript{b}, Fengye Hu\textsuperscript{a}

\textsuperscript{a}College of Communication Engineering, Jilin University, Changchun, China
\textsuperscript{b}College of Automation Science and Engineering, South China University of Technology, Guangzhou, China

INTRODUCTION

In human speech, both linguistic information and paralinguistic information associated with implicit messages, such as emotional states of the speaker, are conveyed. Human emotions are the mental and physiological states associated with the feelings, thoughts, and behaviors of humans. The emotional state expressed by a human subject reflects not only the mood but also the personality of the human subject. In human–human verbal communication, it plays an important role by reflecting the speakers’ responses to the outside world. The same words expressed in different emotions, for example, can delivery quite different meanings. Identification of the emotional states conveyed in speech is therefore quite critical for achieving effective communications between humans.

As computer-based applications receive increasing attention from both academia and industry, the human-computer interaction (HCI) technology has also advanced rapidly over recent decades. Similar to human–human communication, one essential enabler of natural interaction between human and computers is the computer’s ability to understand the emotional states expressed by the human subjects and deliver a personalized response accordingly. Motivated by the demand for human-like machines, automatic recognition of emotional states has been investigated for over the past two decades (Amir, 2001; Cen, Dong, Li, Yu, & Chang, 2010; Cen, Ser, & Yu, 2008; Cen, Ser, Yu, & Cen, 2010; Cowie, 2001; Dellaert, Polzin, & Waibel, 1996; Lee & Narayanan, 2005; Morrison, Wang, & De Silva, 2007; Nguyen & Bass, 2005; Nicholson, Takahashi, & Nakatsu, 1999; Petrushin, 1999, 2000; Scherer, 2000; Ser, Cen, & Yu, 2008; Ververidis & Kotropoulos, 2006; Yu, Chang, Xu, & Shum, 2001; Zhou, Wang,
Yang, & Chen, 2006). Modeling and analysis of emotions from human speech span across several fields, including psychology, linguistics, and engineering. In engineering, speech emotion recognition has been formulated as a pattern recognition problem that mainly involves feature extraction and emotion classification. Speech emotion recognition has found increasing applications in practice, e.g., in security, medicine, entertainment, education. However, the research work on speech emotion recognition is mainly conducted on pre-processed databases that, in general, consist of isolated utterances or phrases. Emotional states are recognized based on these isolated sentences. This limits its applications in practice. Compared with emotion recognition on a pre-processed database, more challenges, such as identification of emotion onset, lasting, and change, recognition efficiency for real-time processing, etc., are faced in real-time speech emotion recognition that aims at detecting emotional states from continuous and spontaneous speech.

With the development of communication technology and human-computer interface, online learning has attracted increasing interest. It has many advantages in comparison with traditional classroom face-to-face learning. First, online learning provides lots of flexibility and convenience to students. With only a computer and Internet connection, students can complete their learning courses anywhere and anytime without struggling to squeeze in a fixed class schedule. Second, the online learning environment is able to provide a wider choice of courses for students, which can both effectively expand their education scope and inspire their interest in different fields. Third, taking a course online is much cheaper than traditional educational programs with expensive tuition fees. Moreover, cost and time spent on the trips to and from the classes can be saved. Fourth, students are allowed to complete online learning courses in an environment in which they feel comfortable. In addition, if they can study from home, they do not need to worry about issues, such as transportation to/from campus, meals in campus, and finding study rooms. However, online learning lacks the interaction between teachers and students. Unlike the traditional face-to-face learning environment, where teachers are aware of students’ responses to the delivered material and can adjust course contents and delivery speed accordingly, teachers in the online learning environment cannot see how the students feel about the ongoing course, from their facial and verbal expressions. This makes it impossible for online learning to adapt course delivery modes to fit students’ learning ability. Solutions to solve or relieve this problem have become important in the development of online learning programs.
In recent years, the impact of students’ emotions on effective learning has been investigated in the communities of education and data mining (Kort, Reilly, & Picard, 2001). Positive emotions can yield good feeling, improve thinking skills, increase the tendency toward greater creativity, help problem-solving, and enhance efficiency and thoroughness in decision-making (Isen, 2000). Indeed, both affective and cognitive functions are integrated in the human brain, and affective functions play very important roles in brain learning. Detecting the affective states of students and understanding their responses to the delivered material can help customize the course in online learning systems to fit each specific student. For example, for high-ability students, the course can be delivered at a higher speed, and for those who are very interested, knowledge extension and extra examples may be introduced; while for those who feel confused, question and answer sessions or more detailed explanations could be added; and for those students who feel bored, funny and interesting activities can be introduced to attract their attention. Course providers can thus adjust the teaching content and delivery speed to satisfy the variety of students. A negative response can alert students to focus on learning and adapt themselves to study. The aforementioned process can enhance the performance and efficiency of learning, and also bring more fun into the learning process.

In this chapter, we first present our work on the development of a real-time speech emotion recognition system. This system is able to accept both pre-recorded speech data and continuous speech recorded in real-time, and also detect the emotional states expressed in speech as it is played back in the former, or as it is recorded in the latter. A friendly graphic user interface (GUI) is provided to display recognition results, including target category and timing information of individual segments, as well as an analysis of emotion frequency statistics based on the whole input data. Experiments with both pre-recorded datasets and real-time recording expressed in four emotional states have been carried out, and the average accuracies of 90% and 78.78% have been achieved, respectively. Second, the application of the developed real-time speech emotion system in online learning is explored with an experiment in a simulated online learning environment. The results have shown that our emotion recognition system can efficiently understand the student’s response to the course, which makes it possible to customize online courses for each student taking the same course, but with different learning abilities, in order to achieve optimal learning outcome.

The remaining part of this chapter is organized as follows. The proposed real-time speech emotion recognition system is presented in the next
section. The experiment results are illustrated and then there follows a description of its application in online learning and the numerical results in the simulation study. Finally, there are concluding remarks.

REAL-TIME SPEECH EMOTION RECOGNITION SYSTEM

As mentioned above, for better applications of speech emotion recognition in practice, automatic recognition of emotion from continuous speech is required. Here, we present a real-time speech emotion recognition system that we have developed.

Graphic User Interface

A user-friendly GUI is provided to allow convenient use of the system without the need to understand the technical details of speech emotion recognition. Figure 2.1 shows a screenshot of the system interface. It mainly includes six parts highlighted using rectangles and labeled with red numbers, which are elaborated on below.

1. The system accepts three types of speech data source, i.e., real-time recording from a microphone, a pre-recorded audio file, and a dataset consisting of multiple audio files. There are two working modes in the system, i.e., online and offline modes. In the online mode, the speech is recorded in real-time, and as recording the emotional states are recognized. In the offline mode, emotional states are detected either individually from an audio file or in a batch from a dataset.

2. If the classification model for emotion recognition has been trained, the model files can be imported to the system for unseen data testing.

3. The three buttons here are used to control the system running for speech processing, i.e., Start, Pause, or Stop.

4. The waveform of the speech signal under processing is illustrated.

5. The timing information of each speech segment and resultant emotion category are displayed.

6. The analysis of emotion frequency statistics is given based on the whole input data upon completion, i.e., the recording is terminated in online mode or all files in the given data set have been processed.

Major Functions

The major functions of the system are shown below:

1. Real-time recording via microphone
2. Creating audio file or database
3. Loading audio file or database
4. Displaying the waveform of the speech signal that is processed
5. Detecting the start- and end-points of speech
6. Automatic segmentation of continuous speech
7. Speech signal pre-processing by means of pre-emphasis, framing, and windowing, followed by feature extraction
8. Training an emotion recognition model
9. Recognizing the emotional states from an audio file or from a speech database
10. Recognizing the emotional states from continuous speech in real-time
11. Displaying recognition results, including the categories and timing information with a resolution of 0.01 s
12. Statistical analysis of emotion frequency for a speech dataset or real-time recording

Figure 2.1 The interface of the real-time speech recognition system.
Methods and Procedure

In our system, emotion recognition is formulated as a supervised learning process, in which the emotional states involved in a speech signal is recognized based on its acoustic features by using a classification model. Figure 2.2 illustrates the framework in supervised classification. A feature set is extracted from the training data to capture the important and discriminative attributes of speech signals. A training model is then constructed by feeding pairs of feature sets and the target values of emotion categories into the learning algorithm of support vector machines (SVMs). In the prediction process, the same features are extracted from unseen speech data, which are fed into the obtained training model to yield predicted labels for the target emotions.

Besides the parts of the feature extraction and model learning for classification, the entire system is mainly composed of six parts: speech detection, speech segmentation, signal pre-processing, feature extraction, emotion recognition, and statistical analysis of emotion frequency.

Figure 2.3 illustrates the system flowchart. First, voice activity detection (VAD) is carried out to detect the presence of human speech. If the speech signal detected is quite long, it will be divided into a few segments. The divided signal is then pre-processed by means of pre-emphasis, framing, and windowing. Next, features are extracted from the processed signal. In our system, the features we use include three short-term cepstral features, i.e., linear prediction-based cepstral coefficients (LPCC); perceptual linear prediction (PLP) cepstral coefficients; and mel-frequency cepstral coefficients (MFCC). Based on the extracted features, the emotional states are
calculated by building a classification model, where in our implementation, the SVM is used as the learning algorithm. These will be elaborated on further below.

**Voice Activity Detection**

VAD, also known as speech detection, aims to detect the presence or absence of speech and differentiates speech from non-speech sections. It is important in a variety of speech-based applications, especially in speech coding and speech recognition. Various VAD algorithms have been developed in the literature, based on different principles, e.g., detecting sudden

![Figure 2.3](image-url)  
**Figure 2.3** The flowchart of the proposed speech emotion recognition system.
changes in energy, spectral, or cepstral distances, in order to satisfy different
requirements from various features and compromises among latency, sensi-
tivity, accuracy, and computational cost.

In real-time speech emotion recognition, correct detection of speech is
an enabling requirement, which directly reflects the resultant recognition
accuracy. In our system, the VAD is carried out based on short-time energy.
The signal is divided into overlapped frames. It is based on the assumption
that the signal within a frame is stationary or quasi-stationary. The frame shift
is the time difference between the start points of successive frames, and the
frame length is the time duration of each frame. We extract the signal frames
of length 25 ms from the signal at every interval of 10 ms.

Let
\[ X_i = [x_1, x_2, \ldots, x_N] \]  
be the signal in the \( i \)th frame with audio samples of \( x_n \), and \( \text{Std}_i \) be the stan-
dard deviation of \( X_i \) measured in logarithm, and then we have:

\[ \text{Std}_i = 20 \times \lg \sqrt{\frac{\sum_{n=1}^{N} (x_i - \bar{X})^2}{N - 1}} \]  

(2.2)

If the condition of:

\[ \text{Std}_i > (M_{\text{Std}} - L_1), \]
\[ \text{Std}_i > -L_2, \]  

(2.3)
is satisfied, the frame is considered as a speech frame, where \( M_{\text{Std}} \) represents
the maximum level of the standard deviation among all frames. In Equation
(2.3), \( L_1 \) and \( L_2 \) are two thresholds that are set according to the environment
situation and background noise, e.g., in the experiments presented later in
this chapter, they are set as:

\[ L_1 = 30, \]
\[ L_2 = 60. \]  

(2.4)

In this way, the presence and absence of speech can be detected. A pause
is defined as a continuous non-speech section with a length longer than 0.3 s.
In order to bear efficient information in emotion recognition, the speech
signal captured from real-time recording should be no less than 6 s. If the
speech section between two successive pauses is shorter than 6 s, it will
be combined with next speech section until the length of speech is equal
to or longer than 6 s, which will be assigned an emotion category in its
entirety.
**Speech Segmentation**

After the speech signal is captured via real-time recording, segmentation is carried out for the purpose of silence removal and speech partition by using the method given in Giannakopoulos (2010). Continuous speech segments and in-between silence that will be removed are detected based on short-time energy and spectral centroid of speech signals with pre-defined thresholds.

Let $E_i$ be the short-time energy of the $i$th frame as

$$E_i = \frac{\sum_{n=1}^{N} |x_n|^2}{N}, \quad (2.5)$$

and $C_i$ be the “gravity” center of the $i$th frame’s spectrum, as

$$C_i = \frac{\sum_{n=1}^{N} (n + 1) X_i^{\text{fft}}(n)}{\sum_{n=1}^{N} X_i^{\text{fft}}(n)}, \quad (2.6)$$

where $X_i^{\text{fft}}$ is the discrete Fourier transform (DFT) coefficients of the $i$th frame. The thresholds for both feature sequences are calculated as:

$$T = \frac{w \cdot m_1 + m_2}{w + 1}, \quad (2.7)$$

where $m_1$ and $m_2$ are the first and second local maxima in the histogram of the corresponding feature sequences, and $w$ is a user-defined weighting coefficient.

Once the two thresholds are estimated, the speech segments are formed by successive frames with feature values larger than the corresponding thresholds. A post-processing for merging overlapping segments is then conducted to obtain the final segments. Finally, all segments shorter than 3 s will be combined with adjacent segments in order to avoid segments, which are too short with inefficient information in emotion recognition.

Figure 2.4 illustrates an example for speech detection and segmentation. The top of Figure 2.4 shows the original speech signal; the resultant signal by removing the endpoint silence is shown below that; the divided signal is then depicted, with different segments highlighted in the figure; and the merging of short segments is displayed at the foot of the figure.

**Signal Pre-processing**

In order to emphasize important frequency components in the signal, a pre-emphasis process is carried out on the speech signal by using a finite impulse response (FIR) filter, called a “pre-emphasis filter,” given by:
In Equation (2.8), the coefficient $a_{\text{pre}}$ can be chosen typically from $[-1.0, 0.4]$ (Picone, 1993). In our implementation, it is set to be:

$$a_{\text{pre}} = - \left( 1 - \frac{1}{16} \right) = -0.9375,$$

so that it can be efficiently implemented in fixed point hardware. The filtered speech signal is then divided into frames with a length of 25 ms at every interval of 10 ms. A Hamming window is then applied to each signal frame to reduce signal discontinuity in order to avoid spectral leakage.

**Acoustic Features**

After the speech signal captured from real-time recording is pre-processed, acoustic features are extracted, which will be used in emotion recognition. The features we used are three short-term cepstral features, i.e., PLP Cepstral Coefficients, MFCC, and LPCC, as shown below:
1. There are 54 features of PLP:
   a. 18 PLP cepstral coefficients
   b. 18 Delta PLP cepstral coefficients
   c. 18 Delta Delta PLP cepstral coefficients
2. There are 39 features of MFCC:
   a. 12 MFCC features
   b. 12 Delta MFCC features
   c. 12 Delta Delta MFCC features
   d. 1 (log) frame energy
   e. 1 Delta (log) frame energy
   f. 1 Delta Delta (log) frame energy
3. There are 39 features of LPCC:
   a. 13 LPCC features
   b. 13 Delta LPCC features
   c. 13 Delta Delta LPCC features

By fusing the PLP, MFCC, and LPCC features, a vector with a dimension of $R^M$ is achieved, where $M = 132$ is the total number of the features extracted for each frame.

**Support Vector Machines (SVMs)-Based Learning Model**

These features, as the representation of original data, are used in the classification model, where the SVMs are employed as the learning algorithm. The SVMs, developed by Vapnik and his colleagues at AT&T Bell Labs in the mid-1990s (Vapnik, 1995), have become of increasing interest in classification. They are has shown to have better generalization performance than traditional techniques in solving classification problems. In contrast to traditional techniques for pattern recognition, based on the minimization of empirical risk learned from training datasets, the SVM aims to minimize the structural risk to achieve optimum performance.

A single SVM itself is a classification method for two-category data. One-versus-all and one-versus-one are two commonly used methods for multiple-class classification problems (Fradkin & Muchnik, 2006). In the one-versus-all method, one SVM is built for each of emotions and used to distinguish this emotion from the rest. The emotion category of an utterance is determined by the classifier with the highest output, based on the winner-takes-all strategy. In the one-versus-one method, one SVM is built to distinguish between every pair of categories, which assigns an utterance to one of the two emotion categories. The final classification decision is made according to the results from all the SVMs based on a max-wins voting strategy.
Statistics Analysis of Emotion Frequency

An emotion category is given to each segment of continuous speech series captured in real-time recording. At the end of recording, statistics of emotion frequency is analyzed, based on classification results, given as:

\[ f_i = \frac{ne_i}{NE}, \]

(2.10)

where \( f_i \) denotes the frequency that the \( i \)th emotion is detected, \( ne_i \) is the number of segments that are classified into this emotion category, and \( NE \) is the total number of speech segments with valid prediction of emotional states.

EXPERIMENTS

To illustrate the effectiveness of our system, experiments are conducted with both the pre-recorded data for offline testing and the continuous speech recorded simultaneously as recognition for real-time testing. The data used in the experiments were recorded from one male speaker in a normal office environment at a sampling frequency of 16 kHz.

Offline Experiment

In this experiment, acted speech with four emotional states: neutral, happy, angry, and sad, were recorded. There were 60 sentences, each spanning around 4-20 s in the corpus for every emotion. Among them, 75\% were used for training the classification model and the remainder for testing. The confusion matrix for emotion recognition is listed in Table 2.1, where the first line gives the actual classes and the first column lists the predicted classes. The diagonal elements represent emotion recognition accuracy at which the predicted label is equal to the true label, while off-diagonal elements are those that are mis-labeled by the classifier. Higher diagonal values of the confusion matrix indicate better classification performance. It can be seen from Table 2.1 that, although the data in both training and testing are

<table>
<thead>
<tr>
<th></th>
<th>Neutral</th>
<th>Happy</th>
<th>Angry</th>
<th>Sad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>86.67</td>
<td>13.33</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Happy</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Angry</td>
<td>0</td>
<td>20</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>Sad</td>
<td>0</td>
<td>6.67</td>
<td>0</td>
<td>93.33</td>
</tr>
</tbody>
</table>
uniformly distributed among emotional states, the recognition accuracy is
diversely distributed with a mean value of 90%. The accuracy achieved based
on a small training dataset with only 45 short spoken utterances for each cat-
egory indicates that our system is quite competitive in offline testing.

**Real-time Recording and Recognition**

In this experiment, continuous speech spanning 8.2 min was recorded with
four emotions that were alternately expressed during the recording, each of
which spanned for around 2 min, with an interval of 30 s. The dataset
recorded in the previous offline experiment was used to train the learning
model. The 8.2-minute recording was divided into 68 segments, among
them 19 with neutral state, 18 with happiness, 17 with anger, and 14 with
sadness. The results are shown in Table 2.2. It can be seen from Table 2.2
that the accuracy of the anger class is slightly higher than that in the offline
testing. The accuracies achieved in the other categories are not less than
80%, except in the sad state. The average accuracy among all emotion cat-
egories is 78.78%. With the same training model, the average performance
in the real-time experiment is lower than that based on the pre-recorded
audio data, which was partially caused by the inaccuracy in speech detec-
tion and segmentation, sound variety in different recordings, and different
effects in various background noise. However, the average accuracy of
78.78% achieved indicates that our system can work properly in real-time
applications.

Figure 2.5 depicts the screenshot of the system output in the real-time
recording and recognition experiment. The statistics of the detected emo-
tions are shown. In the actual recording, the frequency of each of the emo-
tions is 25%, while they vary around 25% in the classification results that are
deviated a little from real labels. To cater for the requirements of some appli-
cations, we also counted the statistics for the positive class, which contains
happiness and the negative class, which includes anger and sadness, as shown

<table>
<thead>
<tr>
<th>Table 2.2 Confusion matrix in the real-time experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
</tr>
<tr>
<td>Neutral</td>
</tr>
<tr>
<td>Happy</td>
</tr>
<tr>
<td>Angry</td>
</tr>
<tr>
<td>Sad</td>
</tr>
</tbody>
</table>
APPLICATION IN ONLINE LEARNING

Use of Emotion Detection in Online Learning

In the previous sections, we have shown a speech recognition system that can detect emotional states from continuous speech signal in real-time. As discussed in the Introduction section, this system, if combined together with traditional online learning systems, can be useful for making learning more effective and interesting. Specifically, it may be used as follows:

in Figure 2.5. Together with the emotion frequency statistics, the time duration of recording is displayed.

Figure 2.5 Screenshot of emotion recognition results with real-time recording.
1. In the online learning environment, there lacks interaction between teachers and students. Detecting the emotional states from the verbal cues of the students in the online learning environment can help teachers understand their responses to corresponding learning content.

2. Sometimes, the emotional states of humans reflect their true response to the outside world more in facial expression than speech linguistic information, since the former is more difficult to keep hidden. Not all students wish to share their real thoughts in class. For example, when a teacher asked “have you got what I am teaching?,” some students may answer “yes” although they may not understand it well. The real response could be hiding and may not be consistent with the speech contents. Emotional states, however, can disclose their hidden thoughts. As a teacher, it is quite important to understand the true thoughts of students about the course.

3. According to the true response of students, the system can help teachers to assess student learning qualities and performance. For example, if a student is detected as being confused, he/she is more likely incapable of understanding what he/she is learning; if a student is detected with unchanged natural emotion, most of the time during class, it may be considered that he/she may be unable to concentrate on the lesson; positive emotions can be considered as active and confident responses to learning.

4. Understanding students’ true responses helps to customize online learning systems by responding to different users’ emotions. This can then be used for course providers to adjust tutoring contents and the delivery speed of teaching to fit each student’s learning ability. The courses can be delivered faster for the students with a higher learning ability, knowledge extension, and extra examples can be introduced for those who feel interested. Furthermore, question and answer sessions could be more specific or more detailed explanations could be properly added for the students who feel confused. Additionally, funny and interesting activities can be introduced for those students who feel bored.

5. In some cases, students themselves are not very sure how to describe their thoughts about the course. With the help of emotion detection, the learners can consciously be aware of their inner response, and accordingly adapt themselves to a better learning status during various learning stages.
Challenges

Although applying emotion detection in online learning can be useful, it is a challenging task.

1. During lessons, students speak at irregular intervals, which require computational efficient algorithms for real-time speech detection, recording, processing, and emotion recognition.

2. In online learning, there may be various students who are learning simultaneously in various environments, with different background noise. The words of the students have to be transmitted via communication systems, which reduces the speech quality at the receiving end of the transmission systems. This can largely decrease the accuracy of speech emotion recognition.

3. It is well-known that speech from various speakers has different features. In general, the accuracy of speaker-dependent emotion recognition is much higher than that of speaker-independent recognition. In online learning, students are not as similar as would be found in a traditional classroom, and their ages and backgrounds will be various. Pre-training for every specific student before using a system is not convenient and would bring an additional burden to practical applications.

4. Human emotional states are not discrete, but continuous. There is a lack of a definite model for each emotion. Any subtle change in intonation, loudness, or pitch of speech may lead to another emotion class. Emotion detection is, in general, based on classification techniques, which require a pre-defined set of emotion classes, e.g., boring, interesting, to reflect students’ responses. More classes can better represent student responses, but may reduce classification accuracy.

5. Since the emotional states are detected from verbal cues, the courses should be properly designed to encourage all students to speak frequently, for example, asking students to answer questions verbally; encouraging them to raise questions verbally during class; frequently asking about their understanding levels; introducing some short discussion session in between teaching hours, etc. In general, active classes are more likely to collect diverse verbal cues and consequently capture students’ responses more accurately.

Experiment Scenario and Results

We now go on to illustrate how the proposed real-time system can be applied in a simulated learning environment. Our focus is to explore the possibility of applying emotion recognition in online learning environments.
**Experiment Scenario**

Scenario: One male university student gave his spoken response by reading from a list of pre-defined comments according to his feelings, while he was watching an online course regarding arts with somewhat difficult content on the Internet. Although the experiment was carried out with a simplified and simulated scenario, the process followed the practical online learning environment.

**Speech Data**

The recording spanned around 10.3 min. The response was spoken with one of three emotional states, i.e., negative state (e.g., sadness); positive state (e.g., happiness); and neutral state. The emotional states expressed for the spoken responses of the student were not specified in advance; the emotional states were decided by the feelings of the student during the course learning.

**Results**

While recording, the emotional states were detected from the spoken comments. The 10.3-minute speech signals were divided into 72 segments. The statistics of the results are shown in Figure 2.6. It can be seen from Figure 2.6 that the negative state accounts for 33.13% and the positive accounted for 43.64%.

**Discussion**

The results indicate that the learning process is enjoyable to this student, as there are more segments with positive emotions than those with negative states. The negative emotions detected are mostly associated with anger and sadness, since the course was, in parts, difficult to understand and unclear. This agrees with the feedback received from the student on learning completion.

By detecting the emotional states from students’ verbal clues, courses can be customized according to the responses of the student. When negative emotion is detected, the original course video can be paused, with an alert shown to the student. Additionally, extra explanations or examples can be shown onscreen instead. If the negative emotion is found to be continuous, a couple of short questions can be given to check the learning situation of the student. When positive emotion is detected, the original course video is paused as well, and the extension of the currently delivered knowledge is given to the student.
Future Work

In subsequent work, we intend to extend our research by exploiting an adaptive speaker-independent model that is trained with sufficient amounts of speech data from various speakers, which will be more competitive in the applications of a real-time systems. We will also extend this research in a practical online learning environment, where university students enrolled in different grades and majors can be fully connected via the Internet and equipped with microphones. During these courses, students will be encouraged to raise and answer questions and all spoken utterances will be collected.
and transmitted to a server machine. The sentiments expressed will then be analyzed in real-time, and accordingly, the course will be adapted to fit each student. Further research on speech enhancement and real-time processing will be conducted to increase emotion detection accuracy when introducing more emotion categories.

**CONCLUSION**

This chapter reports our development and research of a real-time emotion recognition system. This system aims at detecting emotional states from continuous and spontaneous speech. It consists of six parts: voice activity detection, speech segmentation, signal pre-processing, feature extraction, emotion classification, and statistics analysis of emotion frequency. The results of the experiments with pre-recorded data for offline testing and with continuous speech in real-time testing have shown that the proposed system can perform well in multi-working modes. The application of the real-time emotion recognition system in online learning has been explored. A preliminary experiment has been carried out in a simulated online learning environment and the results achieved demonstrate that the real-time emotion recognition system is helpful in understanding students’ responses to delivered material during the learning process, which can then be used as clues to customize the online learning journey to fit students with different learning abilities, by adjusting the teaching content and delivery speed, accordingly.

**REFERENCES**


CHAPTER 3

Pedagogical Agents and Affect: Molding Positive Learning Interactions

H. Chad Lane
Department of Educational Psychology & Illinois Informatics Institute, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

INTRODUCTION

Emotions play an important role in conversations. From basic communicative acts, such as agreement or disagreement, one’s tone of voice, to the positive and negative appraisals of conversational events, the intricate dance of a conversation involves emotional signals at every turn and can produce powerful emotions in its participants (Parkinson, Fischer, & Manstead, 2005). The more poignant social emotions, such as pride, embarrassment, guilt, and confusion, can have long-lasting consequences with the potential to induce relatively stable changes in beliefs, behavior, and memory (Barsade, 2002; Kensinger & Corkin, 2003). It is this potential to promote positive, long-term changes (in learning, specifically) that frames the questions addressed in this chapter. In particular, it addresses: (1) how emotions arise and are influenced during learning interactions; (2) the stability and robustness of changes in learner emotions after their learning experiences.

Learning-related conversations take place in context, and that context influences the power of those conversations. For example, research on the quality of classroom relationships (e.g., between students, or between student and teacher) is known to improve learning and achievement (Martin & Dowson, 2009), and thus it is critical to consider the roles and backgrounds of those involved. Interestingly, when one of those conversational participants is a machine, research has shown that people are naturally inclined to assign social agency to those machines (Reeves & Nass, 1996). As a result of this natural tendency, the question of how to enhance computer-based learning outcomes has led researchers to investigate the presence of pedagogical agents, a type of embodied conversational agent (Cassell, Sullivan, Prevost, & Churchill, 2000). Pedagogical agents engage in a variety of activities, such as helping learners...
solve problems (as tutors often do), asking and answering questions, providing encouragement, or even role-playing as a fellow learner. Most engage in some form of natural interaction and use nonverbal communication, including speech, facial expressions, and gestures. The primary purpose of using pedagogical agents is to simulate social interactions associated with learning, and broaden the communicative bandwidth of educational technologies.

Although early pedagogical agent researchers had high expectations, in the roughly 20 years of empirical research on pedagogical agents, evidence for their effectiveness is generally considered to be mixed (Heidig & Clarebout, 2011; Schroeder, Adesope, & Gilbert, 2013; Veletsianos & Russell, 2014). Although some studies have demonstrated learning gains due to the presence of an agent (Schroeder et al., 2013), many have also suffered from methodological flaws, suggesting that observed learning gains are often due to effective instructional design, which of course does not require the use of an embodied pedagogical agent (Clark & Choi, 2005).

This chapter explores the question of how pedagogical agents impact learners emotionally, both during and after learning interactions. The empirical status of pedagogical agents is presented, as well as issues related to affect and emotions that help shape learning experiences. The chapter concludes with proposals for new directions of research that align more directly with the apparent strengths of pedagogical agents.

PEDAGOGICAL AGENTS: A BRIEF HISTORY

Since their inception, researchers have sought to build more human-like computer tutors. For example, some of the earliest intelligent tutoring systems (ITSs) used natural language to engage students in dialogues that simulated human-human tutoring (Wenger, 1987). These natural language interactions were usually driven by the tutor and consisted of conceptual questions, directed lines of reasoning, and Socratic dialogue (i.e., a series of questions and answers intended to reveal a student misconception or gap in knowledge). Interest in affect also arose early in the history of ITS research, with analyses of how human tutors manage learner motivation (Lepper, Woolverton, Mumme, & Gurtner, 1993) and how game-like activities can be used to increase engagement in learning (Burton & Brown, 1979).

The next logical step in pursuit of more human-like computer tutors, embodiment, was not far behind. Building on seminal work to create conversational agents, Steve was created to help learners acquire procedures for
operating complex equipment in a virtual world (Rickel & Johnson, 1997). Steve inhabited the learning environment with the learner; he moved around the space, looked at the appropriate buttons and switches, demonstrated procedures, and even could undo learner mistakes. It was argued that increasing the communicative bandwidth of such systems via gestures would benefit learning (Johnson, Rickel, & Lester, 2000).

Compelling arguments were made for why an embodied agent should achieve greater learning gains than disembodied counterparts. For example, theories promoted by Lev Vygotsky in the 1950s, who argued that learning is a fundamentally social act, provided a strong inspiration for more human-like computer tutors. Vygotsky proposed that learning is socially constructed during interactions between people (Vygotsky, 1978), and so an embodied agent who can speak, listen, sense, and exhibit nonverbal behaviors would necessarily increase the social nature of a computer-based learning experience. A second, related idea often suggested for why pedagogical agents should promote learning, is that they are believed to enhance the motivation to learn. In other words, effective learners are those that remain engaged in the cognitive activities necessary for learning, even when they are confused and even when they do not understand. A patient, friendly, supportive agent, it was argued, should help learners stay engaged. As will be shown, many of these early compelling arguments are still a work-in-progress, due in large part to the complexity of realistically simulating human communicative behaviors, but also because of the emerging research on the relationships between social behavior and learning.

**DESIGN AND IMPLEMENTATION OF PEDAGOGICAL AGENTS**

It is a nontrivial design and engineering task to build an effective pedagogical agent. To do so, it is often useful to think of them as having both internal and external properties, both of which are relevant to the learning experience (Dehn & Mulken, 2000). Internal properties consist of the pedagogical methods used to promote understanding and performance, such as the provision of feedback and question-asking. External properties refer to the visible and audible features of the agent, such as appearance, animations, and voice. Krämer and Bente (2010) suggest that external properties can be further broken down into static as well as dynamic features. Static properties include the (relatively) fixed traits of an agent, such as gender, race, hairstyle, and dress. Dynamic properties are those that vary during an interaction, such as speech, gaze, gestures, and other nonverbal behaviors.
The decision to create and use a pedagogical agent is therefore a complicated one, and one that involves a variety of choices. Table 3.1 shows a small subset of the design space that addresses many of the physical aspects of creating a new agent, such as how it should look, sound, and behave. Contextual factors involve what role the agent should play (e.g., a domain expert, such as a tutor) and how a learner will interact with it. More complicated decisions need to be made regarding a pedagogical agent’s history, personality, and unique traits. A substantial amount of empirical research has gone into identifying the consequences and impacts of these choices, and the status of this work is summarized in the next section.

**EMOTIONS DURING TUTORIAL INTERACTIONS**

The emotional ups and downs of learning are widely acknowledged as important considerations in the design of any instructional intervention. How a human tutor balances the learner’s affective state with the cognitive activities of learning, such as solving problems or developing conceptual understanding, has been a topic of considerable debate. In their seminal analysis of expert human tutors, Lepper et al. (1993) were among the first to address issues related to motivation and affect in tutoring. They describe a body of techniques employed by human tutors for managing learner confidence, evoking curiosity, ceding control, promoting positive attributions,
and more. Tutors in their studies attended to student emotions at almost all times, and delicately balanced concrete learning goals with learner’s emotional states. In fact, they conclude that expert human tutors “devote at least as much time and attention to the achievement of affective and motivational goals in tutoring as they do to the achievement of the sorts of cognitive and informational goals that dominate and characterize traditional computer-based tutors” (p. 99). Since the publication of this article, intelligent tutoring researchers have radically embraced the challenge to improve the ability of computer tutors to address motivation and affect. In this section, two significant developments in the design of pedagogical agents that build on Lepper and colleagues’ important early work are highlighted: politeness and empathy.

**Politeness**

The idea that a machine should be less like a machine may seem fraught with irony, but it is however appropriate, given that humans naturally bring their social expectations and assumptions to human–computer interactions and treat machines as social actors (Reeves & Nass, 1996). In a learning setting, politeness is most apparent when a tutor or coach is faced with the challenge of helping, correcting, or otherwise supporting a learner who may be struggling. Minimizing negative emotions in learning can be achieved, in part, via politeness and by using simple “face-saving” measures (Brown & Levinson, 1987). For example, a face-saving method for delivering a hint to a learner at an impasse might adopt a collaborative tone, such as: “How about we solve for x?” rather than a directive (e.g., “You need to solve for x.”). Other strategies mirror those highlighted by Lepper et al., such as highlighting task difficulty (“This next problem is really tough.”) and suggesting that it is common for students to find a particular problem difficult. Politeness can have an impact on outcomes, both cognitive and affective, as discussed below.

**Emotion and Empathy**

Building again on the idea that people treat computers as social actors, research has also shown that these effects are amplified when a virtual character is used in the interaction (Gratch et al., 2007). Anecdotal evidence of this effect can be found in stories of people asking personal questions of Siri and other voice-based interfaces, but most importantly the possibility that

---

an agent-based interface can connect with users on an emotional level has implications for how pedagogical agents might productively interact with learners. Whether it be to convey enthusiasm (Bettencourt, Gillett, Gall, & Hull, 1983); strengthen relationships with students (Martin & Dowson, 2009); or to support learners as they struggle with negative academic emotions (Lepper et al., 1993), human tutors and expert teachers often express emotions that are connected to learning events. Not surprisingly, the role of empathy and use of emotional expressions has received attention from pedagogical agent researchers.

As part of a larger goal to cultivate trust and build longer-term relationships with learners, empathetic agents seek to directly address learner emotions by employing appropriate verbal and nonverbal behaviors to express emotions (Kim & Baylor, 2006). For example, emotional communication can be used to stage a learning experience (e.g., “This looks fun!”) and to express solidarity and concern in the midst of problem-solving (e.g., concentration and concern for success). Studies looking at the impact of empathetic responses from pedagogical agents suggest that they directly impact the learner’s experience. In particular, affective responses based directly on learner emotions, such as stated interest and frustration produced gains in both overall interest as well as self-efficacy (Kim, Baylor, & Shen, 2007). Researchers have also investigated the construction of data-driven models of empathic behaviors in a learning context. Here, models of affective responses and empathic behaviors were found to be believable and appropriate (based on human judges), when characters in an educational game mirror and support the emotional states of the learners (e.g., frustrated players find similar frustration in the characters around them) (McQuiggan, Rowe, & Lester, 2008).

Krämer and Bente (2010) identify additional, specific emotional signals that are made possible because of embodiment, including immediacy cues (e.g., smiling and proximity); quality of movement (e.g., posture); and eyebrow raising (which can be used for a wide range of reasons, including surprise, delight, to question, make an offer, etc.). Of course, verbal signals also can play an important role in conveying emotions, such as excitement, disappointment, anxiety, concern, and more. The complexities introduced by embodiment are substantial, but the promise of leveraging them in the service of achieving learning goals is similarly great. The next section summarizes the known impacts of engaging learners emotionally, and the conclusion returns to the promise of fully elaborated nonverbal communication to enhance learning outcomes.
ASSESSING THE IMPACTS OF PEDAGOGICAL AGENTS

Empirical research on pedagogical agents generally falls into one of three groups: (1) studies that focus on the simple presence of an agent; (2) those that focus on appearance or visible features; (3) studies that consider the behaviors or pedagogical policies. Given the wide variability in the design, role, interactivity, and use of pedagogical agents, research on their effectiveness can be difficult to generalize. Studies on pedagogical behaviors and policies (e.g., feedback) are less common, as they do not technically require an embodied agent—in fact, a large body of research exists investigating such questions with intelligent tutors (VanLehn, 2011).

This section briefly addresses the known impacts of pedagogical agents on cognitive outcomes and then shifts to affective outcomes. Although closely related, they are discussed separately as a convenience. Further, for the sake of focus, this section concentrates on pedagogical agents in the role of expert—that is, agents that were built to provide help, ask questions, and so on. Although substantial attention has gone into evaluating other roles, such as teachable agents (Biswas, Leelawong, Schwartz, & Vye, 2005), using pedagogical agents as coaches or tutors remains the most common.

Cognitive Outcomes

A fundamental challenge in evaluating pedagogical agents in terms of how well they can promote learning lies in the fact that ITSs (that are disembodied) have historically proven highly successful (Anderson, Corbett, Koedinger, & Pelletier, 1995; VanLehn et al., 2005). In terms of effect sizes, the best intelligent tutors regularly achieve gains of one standard deviation or more (VanLehn, 2011), which suggests there may be very little room left for improvement in terms of cognitive gains. In fact, incorporating any additional interactive components in a virtual learning environment inherently increases cognitive load because it competes for the learner’s attention. Thus, one way to view the incorporation of a pedagogical agent is in terms of return on investment: adding an animated agent has associated costs—so, what is the payoff? Do they enhance learning?

A decade-old review suggested there was no clear evidence for the use of pedagogical agents (Gulz, 2004), and at least two more recent reviews convey a murky view of the literature (Heidig & Clarebout, 2011; Veletsianos & Russell, 2014). These articles highlight numerous mismatches between the hypothesized impacts of agents and results of experimental studies. A recent meta-analysis from Schroeder et al. (2013) paints a slightly more positive
picture: in a review of 43 controlled studies (3088 participants) comparing learning environments with and without an embodied agent, they conclude that agents had a “small, but positive effect on learning ($g = 0.19, p < 0.05$)” (p. 20). The authors also report an even higher effect size with younger participants (4–7th grade: $g = 0.86, p < 0.05$; and K–12: $g = 0.86, p < 0.05$). Schroeder et al. conclude that their results provide support for social agency theory—that the use of pedagogical agents causes increases in learning due to social interaction between the learner and agent. Further, rapid advances in animation, graphics, sound, and motion capture could radically change the nature and outcomes of the studies included in these reviews.

A theme of the many reviews conducted in the last decade is that the broad question of whether to include an agent or not is growing increasingly less useful. As with many educational technologies, it is not the simple presence of the technology that matters, but rather how, when, and under what conditions they are used. For example, earlier studies on modality suggested that it was narration (audio) that was most responsible for gains in learning (Moreno, 2005); however, the recent meta-analysis by Schroeder et al. (2013) does not support this conclusion, as larger effect sizes were found for agents who use onscreen text rather than narration. If any conclusion can be drawn about using pedagogical agents to enhance learning outcomes (beyond what is possible without them), it is that they seem to provide modest improvements for some learners and in certain situations; however, it is very much a work in progress as to which learners and what kinds of situations (Gulz, 2004; Schroeder & Adesope, 2014).

Affective Outcomes

In addition to a cognitive focus, research on pedagogical agents has also investigated their impacts on learner perceptions, beliefs, and emotions. The belief that embodiment (of a tutor, co-learner, etc.) will lead to a more meaningful learning experience is somewhat consistent throughout the literature. In other words, enabling a machine to emulate emotions, such as empathy and enthusiasm, it is believed, will create improved conditions under which learning is more likely to occur.

One of the key reasons researchers believe pedagogical agents should have an impact on affective outcomes can be derived from social agency theory (Atkinson, Mayer, & Merrill, 2005; Krämer & Bente, 2010). In particular, it suggests that social cues from an agent induce a kind of obligation to respond and interact. If pedagogical agents can engage learners socially and encourage
learners to pursue meaningful goals, then motivation should increase. Seminal work on pedagogical agents seemed to provide support for this hypothesis. In particular, the persona effect was proposed and investigated by Lester et al. (1997), which stated that “the presence of a lifelike character in an interactive environment” can have a positive effect on student’s perceptions of their own learning experiences. Follow-up work by Moreno and Mayer (2000) provided additional support for the hypothesis; however, these early studies were criticized for lacking sufficient controls (Clarebout, Elen, Johnson, & Shaw, 2002).

To date, and much like the status of cognitive impacts, there is some evidence that pedagogical agents impact learners on an emotional level. In this section, that evidence is discussed in order to set the stage for a discussion of future work.

**Self-perceptions and Feelings Toward Learning**

How learners perceive learning content (e.g., positively or negatively) as well as themselves is important when considering longer-term desirable outcomes, such as career choices and lifelong learning. Although there are occasions when the absence of an agent produces more positive perception (Schroeder & Adesope, 2014), learners generally prefer having the help of an agent over not having one (Baylor, 2009; Moreno & Flowerday, 2006). Earlier studies suggest that pedagogical agents led to higher feelings that learning was effective (Atkinson, 2002), with more generally positive feelings toward the learning environment (Moundridou & Virvou, 2002).

In terms of agent behaviors, the politeness effect, defined as the impact of face-saving strategies on learning (as described earlier), has yielded improved learning outcomes and learner self-perceptions in multiple settings and audiences (McLaren, DeLeeuw, & Mayer, 2011; Wang et al., 2008); however, some recent evidence suggests that the effect may fade with time and over extended learning interactions (Brom, Bromová, Děchtěrenko, Buchtová, & Pergel, 2014). There is also evidence that pedagogical agents can be used to enhance self-regulated learning (Azevedo & Hadwin, 2005; Graesser & McNamara, 2010). In general, however, it seems that pedagogical agents are able to enhance learner feelings and perceptions positively.

**Motivation, Interest, and Self-efficacy**

How motivated students are, to learn and stay engaged with learning material, is a critical question for developers of educational technologies. Because
social engagement (with teachers, students, parents, etc.) is central to fostering overall engagement in school (Christenson, Reschly, & Wylie, 2012), pedagogical agent researchers often promote them as a potentially transformative technology in terms of building engagement. Unfortunately, the number of published empirical studies that investigate specific engagement hypotheses is limited (Domagk, 2010). The more general question of how the presence of an agent impacts learner motivation is a more common target. In general and as in our section on cognitive outcomes, affective outcome findings are mixed, at best. Some of the positive findings in terms of motivation, interest, and self-efficacy do support the basic claim that pedagogical agents can improve computer-based learning experiences for learners, or at least some of them.

Self-efficacy, defined as the strength in one’s belief in one’s own ability to complete tasks or goals (Bandura, 1997), is intimately related to motivation. As learners pursue and conquer learning tasks, their confidence builds and the inherent joys of expertise are flirted with. As with human coaches and teachers, the presence of an agent has also been found to produce increases in self-reported belief in utility of the topic, as well as in self-efficacy in success at the topic (Rosenberg-Kima, Baylor, Plant, & Doerr, 2007). Furthermore, some research suggests that social affordances of agents (speech and nonverbal behaviors) promote a positive affect and in some cases, improved transfer of learning (Krämer & Bente, 2010). More recently, agents have been linked to affective states known to be conducive to learning. For example, the personality of an agent (specifically, enthusiasm and “energy”) has led to gains in learners’ self-efficacy in informal learning settings (Lane et al., 2013).

A PATH FORWARD FOR PEDAGOGICAL AGENTS

Pedagogical agents have come a long way in 20 years. During this time, they have evolved from largely expressionless, robotic characters to empathetic and caring supporters of learning. The trend to make machines more human-like seems to be having a positive influence on learning outcomes, at least in particular ways, and in certain situations. They have progressed from simple demonstrations, to the focus of large-scale studies that address cognitive and noncognitive outcomes. This chapter ends with three related recommendations for future research that build on these themes and seek to advance the links between emotions, learning, and outcomes, using pedagogical agents as a vehicle for these advances.
Increase Nonverbal Fidelity

Despite advances in animation and video-game technologies, pedagogical agents still remain far behind in terms of their use of nonverbal behaviors when compared with expert human teachers. The use of gestures during teaching can reinforce concepts and support comprehension, and when used appropriately, can have a direct impact on learning outcomes (Alibali et al., 2013). Further, Krämer and Bente (2010) argue that virtual agents (in general) lack the sophistication to adequately convey naturalistic and appropriate nonverbal behaviors and consider the matter closed. Thus, it is likely that the full potential of pedagogical agents has not yet been explored, specifically to explore how they might promote learning via nonverbal signals and appropriate use of gestures.

Strengthen Links Between Agent Behaviors and Learner Emotions

One of the arguments of this chapter is that pedagogical agents provide unique opportunities (over non-agent enabled learning environments) to increase engagement and connect with learners emotionally. The argument rests on the emotional potential of simulating meaningful social interactions and has roots in social agency theory. Although evidence is mounting for this claim, the findings described in this chapter regarding agent impacts on motivation, interest, self-efficacy, and others, are promising. Alongside research on agents, work on affect-aware educational technologies has led to computer-based learning environments that detect and respond to learner emotions (Arroyo et al., 2009; Calvo & D’Mello, 2011; D’Mello, 2013).

A large portion of this work focuses on emotions that naturally arise during learning, for example, as a result of struggling with problem-solving or learning material. Although very important, an emerging category of educational technologies are those that not only respond to emotions, but also induce specific emotions that are beneficial for learning. For example, Lehman et al. (2011) intentionally induce confusion in learners with two pedagogical agents who disagree, which is known to be a positive precursor for conceptual learning. Confusion can become negative, however, if it devolves into frustration. This suggests that pedagogical agents should strive to carefully manage the emotional states of learners, helping those on the brink of frustration and disengagement and challenging those who may be approaching boredom. Using conversational and nonverbal strategies, pedagogical agents have a wide range of communicative strategies available to achieve such a balance.
Build Real Relationships

The final suggestion for future research on pedagogical agents again builds on the nature of human-to-human learning interactions. Perhaps unsurprisingly, a substantial body of literature suggests that the quality of interpersonal relationships teachers have with students has a profound impact on learner emotions, motivation, and achievement (Martin & Dowson, 2009). How successful teachers connect with students, whether it be through effective nonverbal communication or proper interpretation of student emotions, can act as a blueprint for future pedagogical agent research. The idea that an agent can stay with a learner and be persistently available has already taken hold in health education. Relational agents have been shown to promote better understanding of health issues and adherence to medication (Bickmore & Picard, 2005). As pedagogical agent research moves forward, it will be important to track if techniques intended to build relationships are even possible and if so, how they impact learners over time. The novelty effect can be a problem when using pedagogical agents over extended periods (Schroeder & Adesope, 2014), and so there may be hidden challenges in determining proper behavior and eliminating predictable patterns (to a degree, at least).

CONCLUSION

The bottom line is that the decision to add a pedagogical agent implies a cost, not just in terms of engineering, but also in the working memories and attention of learners. It is perhaps no surprise that pedagogical agents do not produce blanket improvements in learning outcomes. Human teachers vary widely in their effectiveness, and so do pedagogical agents. The bright side is that we have full control over how agents look, behave, instruct, and interact with learners. Animation and artificial intelligence do bring compelling capabilities to computer-based learning environments that demand further attention from researchers. No one has seriously argued for a “teacherless” classroom or for minimalist learning environments void of social interaction. In fact, the opposite is often true when learning science researchers study such factors (Bransford, Brown, & Cocking, 2000), and so the motivation to pursue more socially enabled computer-based learning environments is strong.

Research conducted thus far has produced guidance on how to design agents, when to use them, how they should behave and emote, and what roles they might take in learning settings. As educational institutions and
learning sciences research continue to evolve beyond an exclusive focus on cognitive gains, the impacts of more human-like educational technologies will certainly require further investigation. If pedagogical agents increase the chances of educational technologies to influence learner emotions, then we may see an increase in blurred lines between education, emotional outcomes, and behavior change goals. For example, Heckman’s widely-cited research on the power of early intervention programs for disadvantaged children firmly places noncognitive skills, such as coping and resilience, as a key ingredient for completion of school and reductions in crime, divorce, and drug dependence (Heckman, 2006). If agent-based learning environments can play a supportive role in promoting such profoundly important outcomes, researchers need to continue unpacking and understanding their role in virtual learning environments.

REFERENCES


B. J. Fogg (Eds.), *Persuasive technology: Vol. 4744* (pp. 214–222). Berlin Heidelberg: Springer.


CHAPTER 4

Implementation of Artificial Emotions and Moods in a Pedagogical Agent

Bogdan Pătruț, Roxana-Petronela Spatariu
“Vasile Alecsandri” University of Bacau, Bacau, Romania

INTRODUCTION

In the educational process, the communication between the student and the teacher can be verbal or non-verbal. Both interlocutors and participants can express their emotions. A student can express emotions regarding the relation with the content of the lesson, and toward the kind of knowledge acquired. A teacher can express emotions related to the evolution of the student during the educational process.

As far as a student is concerned, it has been shown that emotions have significance on the ability to learn new information or to solve problems. There are various studies that interpret particular affective states of a student, many pointing out that there is a strong connection between a student’s emotion and a student’s learning.

Both neuroscience and psychology literature show the link between affect and performance on cognitive tasks. Lisetti (1999a, 1999b, 1999c) considers that emotions have a significant influence on a wide range of cognitive tasks, such as decision-making, planning, adapting to new environments, and learning. Picard (1997) agrees, claiming that “emotions play an essential role in rational decision-making, perception, learning, and a variety of other cognitive functions.”

The consequences of a positive affect can have a great impact on cognitive tasks, such as creative problem-solving and cognitive organization, facilitating them and increasing the intrinsic motivation for performing tasks.

A teacher is appreciated according to the ability to assess students correctly; to show feedback and empathy; to anticipate gaps in the student’s knowledge; and to guide the training process in an adequate manner. In addition, a good teacher manages to involve the student when the latter
has become lazy or bored. A teacher’s style, gestures and the emotions may be significant elements, which contribute to pedagogical success.

Intelligent tutoring systems are computer-based educational systems that provide immediate and personalized instruction or feedback to learners, usually without intervention from a human teacher. Despite the evolution of intelligent tutoring systems, the secret ingredient of “emotions” makes the teacher an irreplaceable element in learning.

What if we integrate emotions into intelligent tutoring systems? Starting from this question, this chapter has three objectives: (1) to make a short state-of-the-art review in the domain of theories of emotions and affecting computing; (2) to describe a method for dealing with artificial emotions in an intelligent tutoring system; (3) to present the educational benefits of embedding emotions in intelligent tutoring systems.

**THEORETICAL APPROACHES TO EMOTIONS**

Later, we present a literature review showing different theoretical approaches and models for recognizing emotions (natural or artificial) in any kind of environment, be it educational or social.

**The Discrete Approach**

The discrete approach used to model emotions is based on a distinction between “basic” emotions and all the other emotions. Basic or “primary” emotions are emotions that correspond to a unique and universal facial expression, being fast, hard-wired stimulus-response patterns. They form spontaneously and are connected to a unique state of feelings. Primary emotions are not based on memory, expectation, and higher-order cognition. Other emotions are either mixes of several basic emotions or special cases of basic emotions and they are based on memories and expectations, leading to cognitively elaborated behaviors (Walter, 2007).

**The Two-dimensional Approach: Circumplex Model**

The dimensional approach differentiates between emotions according to two or more characteristics that are given some value for all emotions. The Circumplex model has two such dimensions: valence and activity (Russell, 1980).

Russell’s Circumplex Model of Affect (Figure 4.1) focuses on subjective experiences, and consequently, emotions within these dimensions might not be placed at exactly the same levels for all the people.
Emotions are distributed in a system of coordinates where the x-axis measures the valence of emotions from negative to positive, and the y-axis specifies how actively or passively the emotion is experienced. Therefore, the model, illustrated in a two-dimensional graph, results in a two-valued vector for each emotion, and makes it possible to compare emotions with each other (Figure 4.1). While Russell provides a comprehensive set of emotions, these do not perfectly match in the application of learning, and they are too numerous for self-assessment tests (Shen, Leon, Callaghan, & Shen, 2007).

For example, both happiness and pleasure have a positive valence, but happiness is an active emotion, while pleasure is a passive emotion. Fear and contentment are almost completely opposite emotions, fear being both active and negative, while contentment, on the other hand, being passive and positive.

**The Three-dimensional Approach**

The third dimension is a new dimension, the “emotional intensity,” and sorts the related emotions according to their intensity.

**Plutchik’s Multidimensional Approach**

Plutchik combines the discrete approach with the dimensional idea of ordering emotions according to selected characteristics, resulting in relations between emotions that can be represented in a cone similar to the RGB-color-model (Figure 4.2).
The OCC Model

Ortony, Clore, and Collins (1988), the authors of The Cognitive Structure of Emotions, introduced the OCC model.

This model is also a discrete approach to modeling emotions, with the only difference being that the authors interpret emotions as reactions to either consequences of events, actions of agents, or aspects of objects. The goal of the model is to predict and explain human emotions, or more precisely, to determine which emotions are most likely to occur under certain circumstances. By dividing emotions into reactions to events, actions, and objects, and also by differentiating between positive and negative reactions, we can identify different types of emotions (Walter, 2007).

The OCC model divides emotions into 22 categories, depending on the reactions to situations that appear as goal-relevant events, acts of an agent, or attractive/unattractive objects (Figure 4.3).

The intensity of the emotion types is determined by different variables, such as desirability, praiseworthiness, appeal, sense of reality, proximity, unexpectedness, arousal, likelihood of an event, or familiarity of an object.
Figure 4.3 Structure of emotion types in the OCC Model of Emotions. (Source: Ortony et al. (1988)).
CAPTURING THE STUDENT’S EMOTIONS IN THE LEARNING PROCESS

The role of “affect” or emotion in learning is undeniable. Certainly, teachers know that emotions play a crucial role in motivation, interest, and attention. It has been proven that, for example, a slight positive mood, apart from making the student feel better, makes him think differently. It can bring flexibility in matters of problem-solving, greater creativity, as well as better choices in decision-making.

These findings suggest that emotion may be an important factor in learning and show that the human brain is a system in which both affective and cognitive functions are integrated with one another. Therefore, there are various cognitive theories that study emotions (Shen et al., 2007).

Research carried out by the AutoTutor group has provided evidence for a connection between learning and the affective states of confusion, flow, and boredom (D’Mello, Taylor, Tapp, King, & Graesser, 2007).

The OCC model has established itself as the standard appraisal model (Ortony et al., 1988).

There are various studies that use the OCC theory explicitly for recognizing student’s emotions in educational games, for example, Conati and Zhou’s work in the educational game Prime Climb (Conati & Zhou, 2002). Moreover, educational applications (Hanjalic & Xu, 2005) in video content are tagged with emotion to support personalization for applications, such as generation of “video highlights” or personalized recommendations for video films.

By means of their cognitive theory, Kort and Reilly (2002a, 2002b) explain the role of affect in learning. They proposed a four-quadrant learning spiral model that has not yet been empirically validated, in which emotions change while the learner moves through the quadrants and up the spiral.

Figure 4.4 suggests six possible emotion axes that may arise during learning. Figures 4.5 and 4.6 interweave the emotion axes shown in Figure 4.4 with the cognitive dynamics of the learning process.

In Figure 4.5, the positive valence (more pleasurable) emotions are represented on the right; the negative valence (more unpleasant) emotions are on the left. The vertical axis is called the Learning Axis, and it symbolizes the construction of knowledge upward, and the discarding of misconceptions downward (Kort & Reilly, 2002a, 2002b).
### Figure 4.4 Emotion sets relevant to learning.

<table>
<thead>
<tr>
<th>Emotion Set</th>
<th>Anxiety</th>
<th>Worried</th>
<th>Discomforted</th>
<th>Comforted</th>
<th>Hopefulness</th>
<th>Confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety-confidence</td>
<td>Anxiety</td>
<td>Worried</td>
<td>Discomforted</td>
<td>Comforted</td>
<td>Hopefulness</td>
<td>Confident</td>
</tr>
<tr>
<td>Listlessness-fascination</td>
<td>Ennui</td>
<td>Bored</td>
<td>Indifferent</td>
<td>Interested</td>
<td>Curious</td>
<td>Fascinated</td>
</tr>
<tr>
<td>Frustration-euphoria</td>
<td>Frustrated</td>
<td>Puzzled</td>
<td>Confused</td>
<td>Insight</td>
<td>Enlightenment</td>
<td>Euphoric</td>
</tr>
<tr>
<td>Dispiritedness-enthusiasm</td>
<td>Dispirited</td>
<td>Disappointed</td>
<td>Dissatisfied</td>
<td>Satisfied</td>
<td>Thrilled</td>
<td>Enthusiast</td>
</tr>
<tr>
<td>Terror-excitement</td>
<td>Terrorized</td>
<td>Dread</td>
<td>Apprehension</td>
<td>Calm</td>
<td>Anticipatory</td>
<td>Excited</td>
</tr>
<tr>
<td>Humiliation-proud</td>
<td>Humiliated</td>
<td>Embarrassed</td>
<td>Self-conscious</td>
<td>Pleased</td>
<td>Satisfied</td>
<td>Proud</td>
</tr>
</tbody>
</table>

### Figure 4.5 Four-quadrant model relating phases of learning to emotions.

### Figure 4.6 Circular and helical flow of emotion in the four-quadrant model.
In the first quadrant (I), the learner has positive emotions and accumu-
lates knowledge in a constructive way, working through the material with
ease. When discrepancies start to appear between the information and the
learner’s knowledge structure, they move to quadrant II, which consists
of constructive learning and negative affect, where they experience affective
states, such as confusion (Shen et al., 2007).

For example, when solving a puzzle, a student gets a bright idea, inves-
tigates how to implement a solution and then builds a simulation. If the
student runs the simulation and it fails, the student sees that the idea has
some part that do not work—it thus needs to be diagnosed and reconstructed
(Kort & Reilly, 2002c).

If the learner fails to figure out the puzzle, the student moves into quad-
rant III, where negative affects and un-learning are experienced.

After the misconceptions are discarded, the learner moves into quadrant
IV where the learner is still not sure exactly how to go forward. The quad-
rant is marked by un-learning and positive affect. However, the students can
acquire new insights and search for new ideas that can propel them back into
quadrant I. As learners move up the spiral, they become more competent
and acquire more domain knowledge (Shen et al., 2007).

If one visualizes a version of Figure 4.5 (and Figure 4.6) for each axis in
Figure 4.4, then at any given instant, the student might be in multiple quad-
rants according to different axes.

The education purpose is not to keep the students in the quadrant I, but
to help them see that the cyclic nature is natural in learning, and that if lands
on the negative side, it is an inevitable part of the cycle. The aim is that the
students orbit the loop and that they learn to propel themselves.

A third axis can be envisioned, the “cumulative knowledge” axis, as
coming out of the plane of the page.

The above movement from Quadrants I-II to III-IV can be visualized as
an orbit, and when this third dimension is added, an excelsior spiral can be
obtained.

The learner may experience multiple cycles until completion of the
learning exercise in building a complete and correct mental model associated
with a learning opportunity. It is noted that the orbit gradually spirals around
the cumulative knowledge axis and does not close on itself (Kort &
Reilly, 2002c).

The authors of the model used it on a fully automated computer program
that recognizes a learner’s affect by monitoring facial features, posture pat-
terns, and onscreen keyboard and mouse behaviors.
An intelligent tutoring system (ITS) is a computer-based educational system that provides individualized instructions. A traditional ITS, which is based on the learner’s pedagogical state, decides on how and what to teach. However, it has been demonstrated that an experienced human tutor manages the emotional state (besides the pedagogical state) of the learner to motivate him and improve the learning process. Therefore, the learner model structure needs to be augmented to include knowledge about the affective state.

Neji, Ammar, and Gouarderes described an intelligent tutoring system based on an architecture of a pedagogical agent, called “EMASPEL” (Emotional Multi-Agents System for Peer-to-peer E-Learning), where they integrated five kinds of agents (interface agents, emotional agents, emotional embodied conversational agents, curriculum agents, and tutoring agents) in order to promote a more dynamic and flexible affective communication between the learner and affective system. EMASPEL combines peer-to-peer topology and e-learning together, to propose the emotional framework for an intelligent affective system.

There were many studies which concluded that specific emotions are triggered through a series of stimulus evaluation checks (SECs) (Scherer & Ekman, 1984; Scherer, 1999). The EMASPEL links the SECs system in order to generate the appropriate emotion and, in parallel with the emotional agents, for recognizing the suitable expression given by the learner. As a consequence, the SECs are used in the input and output of the EECA. In their system, the purpose of the emotional agents consists in extracting the learner’s facial expressions (acquisition and facial alignment) and subsequently categorizing them. For an effective communication between an EECA and a learner, the EA needs to be able to identify the other’s emotion state through the other’s expression, and this is called task emotion identification, established by the emotional agents. The integration of emotional agents in the learning environment has been a success and this aims at capturing and managing the emotions expressed by the learner during a learning session.

It is difficult to develop a system that interprets facial expressions. There are two kinds of problems that have to be solved: facial expression feature extraction and facial expression classification. Facial feature extraction uses a standard webcam and requires no specific illumination, while emotional classification is based on the variation of certain distances from the neutral face, managing the six basic universal emotions of Ekman (Ammara, Neji, Alimi, & Gouardères, 2010; Ekman, 2007). After the emotion recognition, the emotions were modeled by the Kort’s Learning Spiral.
The tutoring agent achieves pedagogical expertise on the learner due to the knowledge acquired on the field (theoretical knowledge and practical skills). The diagnoses are based not only on the session’s learning courses, but also on the learner’s historic actions.

Ammara et al. (2010) proposed the following actions of the tutoring agent, after the agent has received the information about the current emotional state of the learner from the EECA emotions, modeled by Kort’s Learning Spiral. The tutoring agent must:

- provide the proper educational materials that raise the student’s curiosity level, if stopping in quadrant I, avoiding boredom to arise
- reformulate the lessons in a simpler way, or give more explanations, if the learner stops in quadrant II
- change the difficulty level of the subject, if the student gets stuck in quadrant III
- give encouragement to the learner if stopping in quadrant IV.

These actions can be modeled depending on the type of learning style detected by the tutoring system. We also think that, if, in the process of designing a pedagogical agent, the learning style of the learner is taken into consideration, then the evaluation results will bring higher scores.

**ARTIFICIAL EMOTIONS**

Today, many artificial intelligence researchers accept that emotions are imperative for the functioning of an “intelligent” computer. This insight stems from the failures of classical artificial intelligence and not from a deep reflection of the topic. Therefore, eyes are now on artificial emotions, not on artificial intelligence.

The idea of an emotional computer, as well as the idea of an intelligent computer, constitutes, for most people, a threat, rather than a hopeful vision. The emotional devices have an important role in most cultures, and there is no coincidence that this concept holds a strange fascination.

Knowing that emotions are essential for flexible and rational decision-making, it is essential for the machines to have flexible and rational decision-making, as well as truly creative thought and a variety of other human-like cognitive capabilities (Ruebenstrunk, 1998).

Affective computing is the study and development of systems that can recognize, interpret, process, and simulate human affects, in order to add this extra dimension into our interaction with machines. To achieve this, it is necessary to provide artificial agents equipped with an emotional layer that
acts in the same way as human emotions. The agent’s behavior will be influenced by the emotion subsystem, either by establishing the importance of events, influencing knowledge processing, or providing the agent an emotional state that it will be able to express and that will finish influencing its behavior.

An emotion can be defined as an internal, mental, and affective state. According to this definition, for example, pain is not an emotion, because it is a physical state, and not a mental state. In the same way, aggression is not an emotion, because it is a behavioral state. Moreover, there can be no such thing as a neutral emotion; emotions being always positive or negative (Walter, 2007).

Emotions can be divided into three layers of behavior. At the top level, there are, what we call “reactions” or “momentary emotions,” which are behaviors that can be displayed briefly in reaction to events, e.g., when someone strikes at us. In the next level, there are moods. These are prolonged emotional states caused by the cumulative effect of momentary emotions. Underlying both of these layers and always present is our personality; this is the behavior that we generally display when no momentary emotion or mood overrides.

These behavior levels have different levels of prominence. Reactions or momentary emotions have the highest priority. If there are no reactions, the behavior is generated based on the current mood. If the mood level is below a certain threshold, then the behavior is generated based on the underlying personality (Figure 4.7) (Wilson, 1999).

![Figure 4.7 Affective state in time (after Wilson, 1999).](image-url)
Affect is the term used to refer to the positive or negative valence of an emotional experience, i.e., whether an emotion is positive or negative. In other words, one important characteristic of an emotion is its affective valence (Walter, 2007).

Computer models of emotions can be classified according to their objectives in systems that “have” emotions and systems that “understand” and “express” emotions or feelings. The former is of particular interest to psychologists who study feelings, but the latter is of interest to developers of pedagogical agents, expressing emotions.

On the one hand, systems that “understand” and “express” feelings are based on more or less refined models of some theories on emotions (Oatley & Johnson-Laird, 1987; Ortony et al., 1988; Ruebenstrunk, 1998). Even if such theories put technical problems first, different models of emotional systems have been implemented (Botelho & Coelho, 2014; Sloman, 1998; Velásquez, 2004; and Canamero, Simon, Toda, Wright, Foliot & Michel, Gadainho & Hallam, Staller & Petta, cited in Ruebenstrunk, 1998). Naturally, they can express different kinds of emotions, depending on their current state, goals, and their “temperament.” However, when a system consists of pure logic only, its behavior can become predictable and, eventually, controlled by people.

On the other hand, building up systems that would really have feelings or emotions means initiating an evolutionary process, which would lead to an emotional subsystem, independent from its human creator (Dertouzos, 2000). This is why many researchers accept that feelings are not imperative for an “intelligent” system to function. An unusual intelligence becomes a real threat for people indeed. Of course, there is a close relationship between consciousness and artificial feelings of agents and this must not be ignored. It is certain that at least some feelings cannot exist outside of the consciousness. For example, shame or pride imply the concept of “self.”

ARCHITECTURE OF EMOTIONAL AGENT

Sometimes, displaying numbers is not the best form of feedback that can be offered to a student who learns something and solves problems. It would be a waste of time and ink, probably, if for each correct or wrong answer, the (real) teacher gave the student a mark. Assessment during seminars, for example, can also be done by verbal encouragement or warnings. For simple questions, or after short testing periods, it is helpful for teachers to express verbally or even mime the emotion they have, regarding the student’s current results.
If a student at the beginning of an evaluation makes a few mistakes, the teacher can express dissatisfaction. If the student persists in making mistakes, despite the teacher’s efforts to help, the fear that the student will not acquire the respective skills may occur. On the contrary, if, from the very beginning, the student answers correctly, there is a feeling of hope, and if the student keeps answering correctly to all, or almost all, questions, the teacher will be satisfied.

The mimics of the teacher’s face, as well as certain laudatory, encouraging, or critical statements can be ways of partial grading and, eventually, there will be a mark from 1 to 10, a number of points, or a certain general grade given to the student. Therefore, the final assessment and grading system should also exhibit human-like behavior similar to that of a real teacher. The agent should express feelings or emotions. The feelings can differ from one agent to another, as they differ from one teacher to another, according to their temperament.

The use of computers in education represents an immense issue in today’s schools, universities, and businesses, with annual global revenue of tens of millions of dollars. This global increase in e-learning expenditure has also been reflected in the quickly mounting body of research in this area, that has raised the question of whether computers as well as people can learn.

Today there are various studies on emotion-sensitive ITS or affective tutoring systems (ATS). Researchers are trying to find out how an intelligent tutoring system could adapt to the affective state of the students, and how it can offer feedback in an effective way (Alexander, 2007).

These tutoring systems can be based on intelligent agents (Patrut, Varlan, & Socaciu, 2008), which use conversational agents, with “body” and “emotions.” Unlike a normal software agent, a conversational agent is not an abstract entity; it can appear as an animated character or as a video with a real person. An animated character can speak to the human user, by using a speech bubble, or even natural or synthesized voices. A real person can speak using a normal voice. A pedagogical agent is a conversational agent which helps learners in computer-based education. In order to emulate the human teacher using a pedagogical agent, during the computer-based instruction, this pedagogical agent must be able to express human-like emotions and feelings. In the past, researchers in artificial intelligence used to focus on representing knowledge, now, their attention is directed toward developing agents that have emotions.

The emotion model of an intelligent/affective tutoring system should enable the agent to argue about emotions the way humans do, they must
be able to evaluate all situations that the character may encounter and they must provide a structure for variables influencing the intensity of an emotion.

In other words, it should enable the agent to show the right emotion, with the right intensity at the right time, which is necessary for the convincingness of its emotional expressions (Bartneck, 2002).

Conversational agents must transmit attitudes, identities, and emotions, and enhance inter-human communication. The users of intelligent tutoring systems expect the pedagogical agents to have socially adequate behavior, even if they are aware of the artificial factor.

We have developed a proper architecture of a pedagogical agent, in order for it to be able to have and express emotions, moreover, to have a temperament. This pedagogical agent can express emotions as a form of non-verbal communication with the learner, in order to give feedback.

Although the concepts of temperament, mood, or emotions refer to people, we will extend them to pedagogical agents. The artificial emotions of a pedagogical agent will be expressed according to the state of goals of the teaching/learning process. For example, happiness is felt when all goals have been achieved by the student; hope comes when some goals are not reached by the student, but their sum does not exceed the sum of the reached goals and there are elements which make the agent believe that eventually the goals will be reached; fear appears when there is no hope. Depending on their beliefs, goals, states, and temperament, the pedagogical agents will express such emotions.

Creating a close-to-reality pedagogical agent triggers four important educational benefits:

- A pedagogical agent that shows interest in the student’s progress can send the student the idea that it is there for him/her and it can encourage him to care more about her own progress.
- An emotional pedagogical agent who is, up to a certain point, sensitive to the student’s progress may interfere when the learner becomes frustrated and before he/she loses interest.
- An emotional pedagogical agent can pass enthusiasm onto its student.
- A pedagogical agent with a complex and interesting personality can turn studying into something fun. A student who enjoys interacting with a pedagogical agent may have a positive perception on studying in general, and, as a consequence, may spend more time studying.
Our purpose is to model the output layer of an emotionally embodied pedagogical agent, in other words, the agent’s behavior. We know that the embodied animated character has to express its emotional state through its speech and facial expressions.

The emotional model’s goal is to provide the credibility of the emotional agent that refers to the capability of the agent to express the right emotion, at the right intensity, and at the right moment. In addition, the consistency of its actions can be a good convincing feature and can be described as the personality of the character.

We started from considering the learner’s answers given in a test. Based on the student’s response, an emotion is fired in the emotional module of the pedagogical agent. The emotion that is born is one of the six basic emotions, with intensity from 1 (minimum) to 5 (maximum): disgust, happiness, sadness, anger, surprise, and fear. The agent will express according to its current emotion, but also on its mood and personality. For this to happen, we have considered a history function that keeps track of its latest $n$ emotions, and allows us to measure the agent’s mood at a certain time. We considered the following moods, extracted from the Circle of Affect (Figure 4.8) (Carmichael, 2012): disappointment, calm, boredom, irritation, inspiration, admiration, satisfaction, deference, and sadness.

![Figure 4.8 The circle of affect.](image)
Examples of how such moods can appear:

- **Disappointment**: The pedagogical agent could become disappointed if it found enough bad answers and was disgusted by the answers of the student.

- **Curiosity**: The pedagogical agent can become curious about the future evolution of the student, when it notices good answers, which activates emotions, such as happy and surprise.

- **Fascination**: Many surprises can put the pedagogical agent in a “fascination” mood.

- **Joyfulness**: Perfect answers can result in joyfulness.

In Table 4.1, these are written as mathematical formulas, depending on the temperament of the pedagogical agent.

The emotions are being stored in an array, a data structure that plays the role of the history function, helping to keep track of the mood. In our approach, the moods were thought of as a mix of the basic emotions, viewed in a diachronic manner. The proportion of emotions considered to generate a mood depends on the activation and pleasantness of the mood. The mood is also a function of the agent’s temperament; consequently, the agents will express themselves differently, according to their temperament.

We denote by: $c^k_m$ the number of occurrences of the primary emotion $e$, with at least the $m$ intensity (measured from 1 to 3), in the history of the last $k$ primary emotions of the agent (the history function, represented as an array, has 100 compounds).

The moods are dependent on the temperaments. They are represented in Table 4.1, where $S =$ sanguine, $C =$ choleric, $P =$ phlegmatic, and $M =$ melancholic.

If none of these states appears, then a default emotional state of the agent is set as the “null” or “awaiting” state.

**THE EMOTIONAL PEDAGOGICAL AGENT FOR A MULTIPLE CHOICE QUESTIONS TEST**

We proposed to implement our approach, by developing an educational program that evaluates the level of geographical knowledge of a student. This software will present a friendly interface composed of the educational content, as well as the emotional agent that will interact with the student through a video. The student will have the possibility to choose the kind of temperament the teacher has. In this way, she can establish a compatibility with the teacher. The educational software will ask the student...
<table>
<thead>
<tr>
<th>Mood</th>
<th>S</th>
<th>C</th>
<th>P</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disappointment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disappointment (S)</td>
<td>$\text{sad}^3_1 \geq 60 \land \text{disgust}^3_2 \geq 60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disappointment (C)</td>
<td>$\text{sad}^3_2 \geq 30 \land \text{disgust}^3_2 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disappointment (P)</td>
<td>$\text{sad}^3_2 \geq 30 \land \text{disgust}^3_1 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disappointment (M)</td>
<td>$\text{sad}^3_1 \geq 30 \land \text{disgust}^3_1 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Curiosity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (S)</td>
<td>$\text{surprise}^3_3 \geq 30 \land \text{happy}^3_3 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (C)</td>
<td>$\text{surprise}^3_2 \geq 30 \land \text{happy}^3_2 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (P)</td>
<td>$\text{surprise}^3_2 \geq 30 \land \text{happy}^3_1 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (M)</td>
<td>$\text{surprise}^3_1 \geq 30 \land \text{happy}^3_1 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fascination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (S)</td>
<td>$\text{surprise}^3_3 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (C)</td>
<td>$\text{surprise}^3_2 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (P)</td>
<td>$\text{surprise}^3_2 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (M)</td>
<td>$\text{surprise}^3_1 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Joyfulness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joyfulness (S)</td>
<td>$\text{happy}^3_3 \geq 60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joyfulness (C)</td>
<td>$\text{happy}^3_2 \geq 60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joyfulness (P)</td>
<td>$\text{happy}^3_2 \geq 60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joyfulness (M)</td>
<td>$\text{happy}^3_1 \geq 60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Calm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calm (S)</td>
<td>$\text{happy}^3_1 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calm (C)</td>
<td>$\text{happy}^3_2 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calm (P)</td>
<td>$\text{happy}^3_2 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calm (M)</td>
<td>$\text{happy}^3_1 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boredom</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom (S)</td>
<td>$\text{disgust}^3_3 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom (C)</td>
<td>$\text{disgust}^3_2 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom (P)</td>
<td>$\text{disgust}^3_2 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom (M)</td>
<td>$\text{disgust}^3_1 \geq 40$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Irritation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritation (S)</td>
<td>$\text{angry}^3_3 \geq 45 \land \text{disgust}^3_1 \geq 45$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritation (C)</td>
<td>$\text{angry}^3_2 \geq 30 \land \text{disgust}^3_1 \geq 45$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritation (P)</td>
<td>$\text{angry}^3_1 \geq 30 \land \text{disgust}^3_1 \geq 45$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritation (M)</td>
<td>$\text{angry}^3_1 \geq 45 \land \text{disgust}^3_1 \geq 45$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Curiosity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (S)</td>
<td>$\text{surprise}^3_3 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (C)</td>
<td>$\text{surprise}^3_2 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (P)</td>
<td>$\text{surprise}^3_2 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curiosity (M)</td>
<td>$\text{surprise}^3_1 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fascination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (S)</td>
<td>$\text{surprise}^3_3 \geq 65 \land \text{happy}^3_3 \geq 65$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (C)</td>
<td>$\text{surprise}^3_2 \geq 65 \land \text{happy}^3_2 \geq 65$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (P)</td>
<td>$\text{surprise}^3_2 \geq 65 \land \text{happy}^3_2 \geq 65$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fascination (M)</td>
<td>$\text{surprise}^3_1 \geq 65 \land \text{happy}^3_1 \geq 65$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Satisfaction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (S)</td>
<td>$\text{surprise}^3_3 \geq 35 \land \text{happy}^3_3 \geq 35$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (C)</td>
<td>$\text{surprise}^3_2 \geq 35 \land \text{happy}^3_2 \geq 35$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (P)</td>
<td>$\text{surprise}^3_2 \geq 35 \land \text{happy}^3_2 \geq 35$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction (M)</td>
<td>$\text{surprise}^3_1 \geq 35 \land \text{happy}^3_1 \geq 35$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Deference</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deference (S)</td>
<td>$\text{angry}^3_1 = 1 \land \text{happy}^3_1 = 1 \geq 30%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deference (C)</td>
<td>$\text{angry}^3_2 \geq 30 \land \text{happy}^3_2 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deference (P)</td>
<td>$\text{angry}^3_2 \geq 30 \land \text{happy}^3_2 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deference (M)</td>
<td>$\text{angry}^3_1 \geq 30 \land \text{happy}^3_1 \geq 30$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sadness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sadness (S)</td>
<td>$\text{sad}^3_3 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sadness (C)</td>
<td>$\text{sad}^3_2 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sadness (P)</td>
<td>$\text{sad}^3_2 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sadness (M)</td>
<td>$\text{sad}^3_1 \geq 70$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
questions about the geography of Europe, such as capitals of countries, neighbors of countries, rivers which pass through the countries, mountain areas, fields, etc.

Based on the answers the student gives to a multiple choice questions test, the pedagogical agent (the virtual teacher) will have a specific behavior. The answers can trigger emotions of high or low intensity that will contribute to the general mood of the agent. In Table 4.2, we illustrate the facial expressions we recorded for the six basic emotions, with the intensities from 1 (minimum) to 3 (maximum). These facial expressions were recorded as very short movies (<2 s).

The mood of the agent is the result of \( n \) triggered emotions, and is represented in the video the user interacts with. The number \( n \) can be considered depending on the time assigned to a particular question, in order to answer it. The mood is also modeled according to the agent’s temperament. In Figure 4.9, we illustrated the formation of the mood, reducing the scale to 10 emotions. The moods will be represented by movies with duration of 5–10 s.

In Figure 4.10, one can see screen captures from the quiz, when the agent is in a happy mood.

The educational software will work like this (Figure 4.11):

**Step 0:** It will start with the mood Null, i.e., with a movie of 5 s.

**Step 1:** The next question will be displayed, and the student will be invited to answer within a certain time.

**Step 2:** Meanwhile, as the student is thinking of the correct answer, the system will iteratively display the movie for the current mood of the pedagogical agent.

**Step 3:** When the student gives the answer, it can be correct 100%, or partially correct. The answer will be evaluated by the agent.

**Step 4:** Depending on the student’s answer, the pedagogical agent will adopt a corresponding emotion, with an assigned intensity. The movie for the current mood will be paused and the short movie expressing that emotion at that intensity will be displayed for a short time (<2 s).

**Step 5:** The appropriate formula from Table 4.2 will be applied, if necessary. The agent’s mood can be changed or not. For example, in Figure 4.10, the formula for the Satisfaction mood and Sanguine temperament was applied.

**Step 6:** Repeat with **Step 1** until the end of the test.
Table 4.2 Facial expressions expressing the basic emotions

<table>
<thead>
<tr>
<th>Emotions</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Happiness</td>
<td></td>
</tr>
<tr>
<td>Surprise</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td></td>
</tr>
</tbody>
</table>

Implementation of artificial emotions and moods
<table>
<thead>
<tr>
<th>Emotions</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disgust</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Sadness</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>Anger</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
</tbody>
</table>
Figure 4.9 How emotions, such as happiness and surprise generate, the satisfaction mood, for a pedagogical agent with a sanguine temperament.

Figure 4.10 Quiz with pedagogical agent expressing happy emotion, with intensity 3.
CONCLUSIONS

On the one hand, emotional intelligence is critical in performing cognitive tasks, and positive affect has been shown to assist in problem-solving and cognitive organization. However, Kort and Reilly (2002a) argue that negative affect is also part of the learning process, leading to the conclusion that tutors can help the student’s learning process by means of the student’s awareness of her affective state.

On the other hand, improving the intelligent tutoring systems with artificial emotions for their pedagogical agents will bring benefits to the educational process. An emotional pedagogical agent gives students the opportunity to be aware of their own progress in learning, interfering when a student gets frustrated or before they start losing interest. It can also make students more enthusiastic about the content transmitted and they may consider studying more enjoyable.

The implementation of artificial emotions and moods can be seen as a manner to improve proficiency and to motivate students to learn in a pleasant and human-like atmosphere.
ACKNOWLEDGMENT

We thank our friend, Elena-Paula Rusu, for her help in creating all the facial expression videos for the artificial emotions of the pedagogical agent.

REFERENCES


CHAPTER 5

Measuring Emotions: A Survey of Cutting Edge Methodologies Used in Computer-Based Learning Environment Research

Jason Matthew Harley\textsuperscript{a,b}
\textsuperscript{a}Computer Science and Operations Research, University of Montréal, Montréal, Québec, Canada
\textsuperscript{b}Department of Educational and Counselling Psychology, McGill University, Montréal, Québec, Canada

INTRODUCTION

The last few decades have witnessed a surge in the use of interdisciplinary methods (e.g., automatic facial expression analysis software, electrodermal activation measurement devices, self-report measures), to study the complex role of emotions in a multitude of learning contexts (Azevedo & Aleven, 2013; Pekrun & Linnenbrink-Garcia, 2014). Amidst the intertwining of traditional and cutting edge methods and learning environments, however, many fundamental questions have either gone unanswered or need to be revisited, in light of new empirical research and a broadening interest in interdisciplinary methods for measuring emotions. These questions include: What are the best methods for measuring emotions? Is there one optimal method, or is it best to use as many different methods as possible? What are the factors that should inform the selection of different affective methodologies? What analytical considerations are important for analyzing different types of emotion data? What theoretical considerations are important for understanding conflicting or converging results from multiple methods?

This review is selective in surveying studies and methods that illustrate the state-of-the-art in emotion measurement for the purpose of answering the above questions. In particular, this chapter is designed to introduce these interdisciplinary methods and the issues related to them to educational psychologists, by drawing primarily on a relevant field of literature where computer science and the learning sciences intersect: research on computer-based learning environments (CBLEs). CBLEs (e.g., multi-agent systems, intelligent tutoring systems, serious games) are used to help students
learn various educational and professional topics, such as science, math, computer literacy, and military competencies through computer-based learning platforms and their embedded learning tools (e.g., help options, structured notes). Many of these environments use artificial intelligence (AI) techniques to model and adapt to learners’ individual learning needs as they interact with the system and in some cases, train and foster students’ cognitive, emotional, and metacognitive self-regulatory skills (Azevedo & Aleven, 2013; Azevedo et al., 2013; Conati & Maclaren, 2009; D’Mello & Graesser, 2013; Lester, McQuiggan, & Sabourin, 2011; Rodrigo & Baker, 2011; Woolf et al., 2009). Focusing on the scope of the review on CBLEs facilitates a detailed discussion of the types of emotion measurement techniques used in this cutting edge field of research. Furthermore, although the empirical research that informs this review stems from studies using CBLEs, these methods are not exclusive to CBLE research. This chapter therefore represents a potentially useful resource for researchers interested in learning more about different methodologies for measuring emotions. Researchers (e.g., cognitive psychologists, psychophysicists) interested in learning more about how different methods for measuring emotions are being used in other fields of research and/or those familiar with the methods discussed, but less familiar with more recent studies (e.g., affective computing researchers) may also benefit from this survey chapter and its insights, in particular.

The primary objective of this chapter is to provide a synthesis of how learners’ emotions are assessed (alongside learning outcomes) in contemporary CBLEs. In doing so, the chapter adopts an analytical perspective in order to evaluate relevant literature selected to serve as examples of the state-of-the-art and to provide recommendations to assist researchers in improving their emotion measurement methodologies. As such, this review differs from other reviews, meta-analyses, and surveys of the literature that examine emotions and emotion detection methods. First, unlike reviews of emotion detection methods by Calvo and D’Mello (2010) and Mauss and Robinson (2009), this chapter does not seek to provide a comprehensive review of existing emotion measurement techniques nor a specific method or subset thereof (Porayska-Pomsta, Mavrikis, D’Mello, Conati, & Baker, 2013; Zeng, Pantic, Roisman, & Huang, 2009), nor is it a meta-analyses (D’Mello & Kory, 2012). Instead, this text distinguishes itself by providing a contemporary, critical overview of the emotion measurement methods used in research with CBLEs that yields novel insights and revisits important issues pertaining to emotion measurement, including: theoretical and analytical
considerations; the effectiveness of multimethod (e.g., multimodal) approaches; and the potential use of different types of data for informing emotionally supportive CBLEs (for a review of emotionally supportive intelligent tutoring systems, see Sottilare, Graesser, Hu, & Goldberg, 2014).

This chapter was also written with the educational psychology community in mind, where a growing interest in alternative and complimentary methods to self-report measures is being manifested through chapters published in international handbooks, such as The International Handbook of Emotions in Education (Pekrun & Linnenbrink-Garcia, 2014) and a symposia at a recent meeting of the American Educational Research Association (AERA) on analyzing data from multiple emotional channels (Harley & Azevedo, 2014). These examples highlight the timeliness of a chapter written primarily for and published within the field of educational psychology in contrast to some of the excellent reviews that have been written in the past, but primarily for other disciplines, such as computer science, psychophysiology, and cognitive psychology (Calvo & D’Mello, 2010; Mauss & Robinson, 2009; Zeng et al., 2009). As such, this chapter stands to not only add to the existing interdisciplinary discussion on emotion measurement methods, but to expand it to potentially new participants.

This chapter is therefore structured to first introduce emotions and provide the foundations required to understand how emotion-related terms are used here. Following this introduction, research measuring emotions with CBLEs is reviewed in two sections. The first provides a brief overview of the different methods used by CBLE researchers to measure emotions, organized by modality (i.e., different channels that provide information about emotions; e.g., facial expressions, physiological responses, body posture). In reviewing methods within each modality (e.g., facial expressions: human coding protocols and automatic facial expression recognition software), examples of their use in different CBLEs are provided and critically discussed, and advantages and disadvantages are summarized (e.g., whether the measurement is made on-line or off-line and if the emotion data can be analyzed in real-time). Online and offline measurements are defined in line with Azevedo, Moos, Johnson, and Chauncey (2010), who distinguished online trace methodologies as those that capture an activity that occurs during processing (e.g., eye-tracking, log-files, concurrent think-alouds, facial expression of emotions) from offline methods (e.g., self-report measures) that capture any activity that either precedes or follows learning and problem-solving. Real-time data analysis refers to a program’s or device’s ability to automatically score the data it receives as an emotional
state (e.g., anger, happiness) as it is recorded. While real-time data analysis stands to provide CBLEs with information on learners’ evolving emotional states that can be used to provide emotionally supportive interventions (e.g., through prompts and feedback from a virtual pedagogical agent), it is a technical requirement that a means exists for transmitting this data as it is being analyzed from the analyses software to the CBLE. This is a device-dependent matter and therefore beyond the scope of this chapter, which focuses on whether real-time analyses options exist for different methods of measuring emotions, rather than technical questions surrounding the implementation of this information into CBLEs. After summarizing each of the methods individually, research that examines the joint use of different emotion detection methods is summarized and the effectiveness of multimethod and multimodal approaches to measuring emotions is discussed. Theoretical and analytical considerations follow and the chapter concludes with an overview of recommendations and future directions.

EMOTIONS: A PRIMER

In this chapter, emotions are defined based on Gross (2010; Gross & Barrett, 2011), in which emotions are one of three types of affective processes. Accordingly, emotions are defined as responses to situations that are perceived as relevant to an individual’s current goals and consist of three different components, namely feelings, behaviors, and physiological responses. Different components refer to different sources from which emotions can manifest, be experienced, and therefore be classified and measured (e.g., smile or self-report of “feeling great” = happiness). Gross (2010) describes changes in emotional components as occurring in a loosely coordinated manner in response to individuals’ appraisals, defined as ways of thinking about or understanding a situation.

Debate over how many emotions exist and how they should be organized in relation to one another remain popular topics in contemporary research on emotions (D’Mello, 2013; Gross, 2010; Russell, 2012; Scarrantino, 2012). At the superordinate (i.e., dimensional) level, there is consensus that valence (i.e., positive vs. negative state) and arousal (i.e., activating or deactivating state) are critical and generalizable features of emotions that can be used to classify them in a two-axis framework or “affective grid” (Calvo & Mac Kim, 2013; Russell, 1980).

Agreement among researchers regarding how many emotions exist and what constitutes them, begins to deteriorate when one moves beyond the dimensional level. For example, Ekman’s program of research focused on
identifying six universal emotions (anger, fear, disgust, sadness, happiness, and surprise) that can be classified according to nine criteria, including having distinct universal signals (e.g., facial expressions), having a quick onset, and brief duration (Ekman, 1992). Ekman called these six emotions basic emotions because of their universality and prototypicality (i.e., each has a superordinate position at the head of six different families of emotions, e.g., delight belonging to the happiness family of emotions). In contrast, Ortony, Clore, and Collins refuted the notion that nonbasic emotions stem from basic emotions and instead organized emotions according to their valence and the relevant aspects of the individuals’ context (e.g., events, agents, and objects; Ortony, Clore, & Collins, 1988). Pekrun (2011) has also differentiated emotions based on their object focus (such as academic emotions from epistemic, topic, and social emotions). Similarly, D’Mello and Graesser have labeled additional psychological states (e.g., confusion and curiosity) as learning–related emotions because of their relevance to learning and qualities, such as valence (D’Mello, 2013; D’Mello & Graesser, 2013). D’Mello and Graesser distinguish several of these emotions as cognitive-affective hybrids because of their strong relationship to cognitive features, such as comprehension. In this review, cognitive-affective emotions refer to these emotions.

**HOW ARE LEARNERS’ EMOTIONAL STATES MEASURED IN RESEARCH WITH CBLES?**

**Facial Expressions**

Facial expressions are configurations of different micromotor (small muscle) movements in the face that are used to infer a person’s discrete emotional state (e.g., happiness, anger). Ekman and Friesen’s facial action coding system (FACS) was the first widely used and empirically validated approach to classifying a person’s emotional state from their facial expressions (Ekman, 1992; Ekman & Friesen, 1978).

An example of facial expressions being used by human coders to classify students’ nonbasic emotions comes from Craig, D’Mello, Witherspoon, and Graesser (2008) and D’Mello and Graesser (2010), who had two trained coders classify participants’ emotional states, while they viewed videos of participants interacting with AutoTutor, a CBLE designed to foster students’ comprehension of physics and computer literacy. They developed their coding scheme by reducing the set of action units from FACS (used to code facial expressions) to those that they judged relevant to classifying learner-centered emotions, such
as boredom, confusion, and frustration. The interrater reliability of trained coders using this coding framework is good for time points selected by the coders for emotions that occurred with sufficient frequency (overall $\kappa = 0.49$; boredom, $\kappa = 0.44$; confusion, $\kappa = 0.59$; delight, $\kappa = 0.58$; frustration, $\kappa = 0.37$; neutral, $\kappa = 0.31$; D’Mello & Graesser, 2010). Judges interrater reliability scores were, however, much lower for evaluations of emotions at preselected, fixed points (for emotions that occurred with sufficient frequency; overall $\kappa = 0.31$; boredom, $\kappa = 0.25$; confusion, $\kappa = 0.36$; flow, $\kappa = 0.30$; frustration, $\kappa = 0.27$; D’Mello & Graesser, 2010). Although less than ideal, these kappa values are nonetheless common (Baker, D’Mello, Rodrigo, & Graesser, 2010; Porayska-Pomsta et al., 2013) and point to the difficulty of classifying participants’ learning-centered emotions, especially at preselected intervals where little affective information is available. For this reason, most emotional coding systems that use facial expressions classify either only facial features (e.g., eyebrow movement) or basic emotions (Calvo & D’Mello, 2010; Zeng et al., 2009), or combinations of facial features and vocalizations.

One of the relatively new and promising trends in using facial expressions to classify learners’ emotions is the development and use of software programs that automate the process of coding using advanced machine learning technologies (Grafsgaard, Wiggins, Boyer, Wiebe, & Lester, 2014; Harley, Bouchet, & Azevedo, 2013; Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015). For example, FaceReader (5.0) is a commercially available facial recognition program that uses an active appearance model to model participant faces and identifies their facial expressions. The program further utilizes an artificial neural network, with seven outputs to classify learners’ emotional states according to six basic emotions, in addition to “neutral.” Harley et al. (2012, 2013, 2015) have conducted research with FaceReader to: (1) examine students’ emotions at different points in time over the learning session with MetaTutor (Azevedo et al., 2012; Azevedo et al., 2013); (2) investigate the occurrence of co-occurring or “mixed” emotional states; and (3) examine the degree of correspondence between facial expressions, skin conductance (i.e., electrodermal activity), and self-reports of emotional states, while learning with MetaTutor by aligning and comparing these methods.

Although automated facial recognition programs are able to analyze facial expressions much faster than human coders, they are not yet as accurate (Calvo & D’Mello, 2010; Terzis, Moridis, & Economides, 2010; Zeng et al., 2009). The accuracy of automatic facial expression programs varies, both by individual emotion and software program (similar to variance between studies of human coders). An important issue pertaining to automatic facial
expression recognition software, especially commercial software, is its continuous evolution, including larger training databases and the inclusion of more naturalistic (non-posed or experimentally induced) emotion data (Zeng et al., 2009).

Less sophisticated, partial facial expression recognition programs are also used in research with CBLEs, such as the Blue Eyes camera system, that is able to detect specific facial features and motions (Arroyo et al., 2009; Burleson, 2011; Kapoor, Burleson, & Picard, 2007). These programs are used differently than fully developed automated or human facial coding programs because they do not provide discrete emotional labels. Instead, the facial features they provide are combined with other physiological (e.g., electrodermal activity, EDA) and behavioral data (e.g., posture) to create sets of predictive features (e.g., spike of arousal, leaning forward, eyebrows raised) that are correlated with other measures of emotions, such as self-report instruments, to validate their connection to different emotions or emotional dimensions. Studies that investigate different emotion detection methods and the conclusions we can draw from them regarding the value of individual methods are discussed later.

In summary, facial expressions have numerous advantages as a method for measuring emotional states. For one, they are the most traditional and remain one of the best measures of emotional states in terms of their widespread use and reliability (tested with multiple raters and with other methods of emotions) that is unmatched by most other methods that are more newly developed (see Calvo & D’Mello, 2010; Zeng et al., 2009). Furthermore, facial expressions can be analyzed in real-time using software programs, such as FaceReader and the Computer Expression Recognition Toolbox (CERT; Grafsgaard et al., 2014) or after the experimental session concludes, using human coders (Craig et al., 2008). Options to detect emotions in real-time, such as automatic facial recognition software, also make facial expressions a viable channel to provide information to the CBLE about the learners’ emotional state, which can in turn be used to provide emotionally supportive prompts (these environments are henceforth referred to as emotionally supportive CBLEs). Finally, and like most of the methods discussed in this chapter, facial expression recognition measures are online measures of emotion that capture the expression of an emotion as it occurs and therefore mitigates the shortcomings of offline self-report measures (discussed in detail below).

The disadvantages of using facial expressions to measure emotions are that most facial expression coding schemes rely on the FACS system traditionally used to classify only the six basic emotions, and are very labor-intensive if done by trained human coders rather than software (Calvo & D’Mello, 2010).
Programs of research that use facial expressions to examine nonbasic emotions (e.g., D’Mello & Graesser, 2013; Graafsgaard et al., 2014; Rodrigo & Baker, 2011) require extensive cross-method validations to connect configurations of facial expressions with new emotional labels (e.g., engagement, frustration, boredom). Ultimately, facial expression research is currently best suited to examining basic emotions and most efficiently done when using automatic facial recognition programs, which are continuing to improve and approach levels of classification accuracy similar to human coders (Calvo & D’Mello, 2010; Zeng et al., 2009).

Body Posture

Body posture refers to the position of one’s body, including the arrangement of limbs (e.g., arms crossed) and positional orientation (e.g., leaning backwards or forwards). Body posture is a method of measuring the behavioral component (e.g., in addition to facial expressions) of learners’ emotions that has been used by CBLE researchers to measure their level of interest and related emotions (e.g., boredom, engagement/flow, confusion, frustration, delight). Posture evaluation methods have successfully utilized both human coding of posture (Rodrigo & Baker, 2011); automated postural sensors (Arroyo et al., 2009; Burleson, 2011; Cooper et al., 2009; D’Mello & Graesser, 2010; Kapoor & Picard, 2005); and most recently, automated computer-vision techniques using motion filtering or Kinect depth cameras (D’Mello, Dale, & Graesser, 2012; Grafsgaard, Fulton, Boyer, Wiebe, & Lester, 2012). All of these methods measure and classify participants’ seated posture (and changes to it) while they interact with CBLEs. Research with the Kinect depth camera has also examined hand-to-face gestures (Grafsgaard et al., 2012).

The advantages of postural readers include their data being online and noninvasive to collect (even for observations if done properly; Rodrigo & Baker, 2011). Automatic classification methods make this approach less resource-intensive than human coding and can provide data for classifying emotions in real-time. The major disadvantages of using body posture as an emotional channel are that, unlike facial expressions, it is a more nascent method and therefore less empirical work has been done to outline its constraints or validity. CBLE research has also examined posture as an emotional measurement method within multisensor research sets (Burleson, 2011; Cooper et al., 2009; D’Mello & Graesser, 2010; Grafsgaard et al., 2014), as well as part of a coding scheme that includes facial expressions (Rodrigo & Baker, 2011). When evaluated in multimethod studies, postural coding has had mixed results in terms of its additive contribution to
classifying emotions when combined with other methods, such as facial expression (Arroyo et al., 2009; D’Mello & Graesser, 2010; Grafsgaard et al., 2014; Kapoor & Picard, 2005).

In summary, a large enough corpus of research is not yet available to tease apart differences in findings nor confidently assess the individual or additive value of postural methods as a means for measuring emotions. Therefore, although postural techniques are online measures of emotion, are noninvasive, and can inform emotionally supportive CBLEs when real-time data is collected using automatic means, research to-date does not support its use over other methods (e.g., automatic facial recognition software, physiological devices).

**Physiological Patterns**

Physiological patterns refer to the changes and fluctuations in electrical signals produced in the brain (electroencephalography, EEG), heart (electrocardiogram, ECG), muscles (electromyogram, EMG), and skin (electrodermal activity, EDA) that have been connected through research to emotions (Calvo & D’Mello, 2010; D’Mello & Kory, 2012; Harley et al., 2015; Mauss & Robinson, 2009; Zeng et al., 2009). Most physiological methods, such as those derived from heart rate and skin conductance, measure changes in the levels and patterns of participants’ arousal/activation and are used by researchers to make inferences regarding participants’ emotional states, such as interest (for high arousal/activating states) or boredom (for low arousal/deactivating states). Recent research using various peripheral physiology measures of the autonomic nervous system (ECG, EMG, and EDA) has extended the granularity of affective detectors from dimensional categories (i.e., activation/arousal) to discrete emotional states, including boredom, curiosity, and engagement (AlZoubi, D’Mello, & Calvo, 2012).

Use of peripheral physiological measures is becoming more common in CBLE research (AlZoubi et al., 2012; Burleson, 2011; Cooper et al., 2009; D’Mello & Graesser, 2013; Harley et al., 2015; McQuiggan & Lester, 2009). AlZoubi et al. (2012) recently conducted an empirical study with AutoTutor, examining the effectiveness of user-dependent (calibrated to take individuals’ physiological variance and baselines into consideration) and user-independent affect detectors (calibrated with a pool of participants to average out individual variance and baselines), including ECG, EMG, and EDA sensors. They found that user-dependent models yielded the best results ($\kappa = 0.25$) when compared with learners’ retrospective, self-reported learner-centered emotions experienced during the session at 20-s intervals.
Less research has been done using brainwave signals, but Frasson and colleagues have conducted several basic research studies using EEGs to detect participants’ emotions as well as to model workload and engagement (Chaouachi & Frasson, 2012; Heraz, Razaki, & Frasson, 2007). Heraz et al. (2007) used machine learning techniques to predict anger, boredom, confusion, contempt, curiosity, disgust, eureka, and frustration, with an overall kappa value of $\kappa = 0.78$.

Physiological patterns are becoming a popular channel for studying emotion in CBLEs because of their potential to be used for noninvasive and non-labor intensive online measurement of students’ emotional states. For example, instruments that measure skin conductance (i.e., EDA) have evolved from gloves (Woolf et al., 2009) to wrist bracelets (Burleson, 2011). More affordable devices are also becoming available through researchers who have developed their own equipment, such as Picard’s MIT-based research group (Affectiva) and others (e.g., Ashametrics), which help to reduce costs and increase use of these technologies.

Several challenges remain, however, in using physiological devices. In addition to issues of cost (e.g., EEG, EKG), there are important issues to be addressed related to training and the human expertise required to use the devices properly, including the accurate analysis of resulting data. Specifically, physiological data requires expertise in signal processing for feature extraction and analysis to create meaningful emotional information (e.g., identification of significant increases in activation/arousal, elimination of artifacts from signal) or labels (e.g., anxiety) from fluctuations (e.g., a spike in skin conductance) in the human bodies’ response to emotion-eliciting stimuli (e.g., negative feedback on a quiz).

Context can provide additional information, which is especially helpful when classification algorithms are not available to identify specific emotional labels (e.g., anxiety, boredom) and one is working with raw data (often indicative of arousal/activation levels). In this situation, the researcher may still require the use of inferences in order to infer an emotional state. For example, a student might have a spike in arousal, as indexed from an EDA bracelet when a quiz is announced. A researcher may infer that the spike is due to an onset of intense anxiety; however, it is also possible that the student may see the quiz as an indicator that they are nearly finished with a section of content they found easy and are excited to move on to content they perceive as more interesting.

Another important consideration regarding the use of physiological data is the necessity for effective user-dependent models to be developed. These
models can quickly and accurately assess individuals’ physiological parameters (e.g., baselines), which can be done during equipment set-up or the viewing of instructional videos. In contrast, user-independent models draw on previously existing data in order to ascribe meaning to the raw physiological data generated by these devices. User-independent models have inferior accuracy rates for detecting emotions, due to the variance between individuals (e.g., baselines; AlZoubi et al., 2012).

In summary, physiological sensors offer promising directions for online, real-time, and increasingly granular measurement of emotions (individual states, such as boredom, vs. more general dimensions) that can inform emotionally supportive CBLEs, although considerations of financial cost and human expertise will slow the widespread use of this method until more cost-effective hardware and analytical solutions become available. It should be noted that many physiological sensors, such as those that measure heart rate and skin conductance, are becoming more affordable, less invasive, and consequently increasingly used to collect emotions data with CBLEs in schools (Arroyo et al., 2009). Physiological sensors are therefore a promising method for detecting learners’ discrete emotions, but require human, technical, and funding infrastructure that most research programs currently do not possess. At present, physiological sensors are thus best thought of as supporting online trace methods for detecting emotions because they do not yet have the same empirical support or detection accuracy as measures of facial expressions or self-report measures, and should not be relied upon as primary emotion detection methods.

**Self-Report Measures**

Self-report measures are the most widely used method to measure emotions and are based on participants’ self-reported (perceived) experience of emotions, rather than behavioral or physiological emotional information. Although self-reports are flexible regarding when they can be administered (e.g., before, during, or after a learning session) they are, strictly speaking, offline measures because participants are interrupted and their attention is redirected from the emotion eliciting stimuli (i.e., object focus) to the self-report measure (Zimmerman, 2008).

One example of a self-report measure administered while participants are interacting with learning software, specifically *Prime Climb* (Conati & Maclaren, 2009), was the use of an emotion reporting box for participants to report if they were enjoying themselves while using it or experiencing anxiety, as well as feelings of admiration or reproach (contempt) toward
the PA (pedagogical agent) embedded in the system. Participants either filled out this Likert-scale measure when they experienced a change in their emotional state or when prompted by the system. This administration differed from the self-report measures administered by Harley et al. (2013, 2015) to participants while interacting with MetaTutor to assess 19 different discrete emotional states on five different occasions during a 1-h learning session (at the beginning and end of the learning session and three occasions in between, spaced 14 min apart). This questionnaire was also more general than the questionnaire used by Conati and Maclaren (2009) because it did not ask students how they felt about the PAs specifically.

A less traditional approach to asking students to report their emotional states has been developed by Graesser, D’Mello, and colleagues that involves students watching a video of themselves interacting with AutoTutor. Specifically, they are presented with video recordings of their face aligned with a screen recording of their learning session (AlZoubi et al., 2012; D’Mello & Graesser, 2012a). A handout providing a definition of an emotional state was also provided to students. This approach has been validated by comparing learners’ postsession judgments of their emotions with measures of emotions and observations by judges as well as patterns of facial activity and gross body movements (Craig et al., 2008; D’Mello & Graesser, 2010).

The advantages of self-report measures mainly involve being able to measure a range of discrete emotions (e.g., basic, achievement, learning-centered). Self-report measures are also easy to administer and require little expertise in terms of coding, scoring, and analyzing. Furthermore, self-report measures tend to be seen in many disciplines as the “gold standard” for measuring psychological phenomena and as such, serve as a valuable method for both measuring emotions and cross-validating the findings of other methods (Porayska-Pomsta et al., 2013).

Although self-reports are ubiquitous in educational, cognitive, and social psychology research, there are many well-known shortcomings with this method that are relevant to measuring emotions. One of the major shortcomings is asking a learner to rate their perception of having experienced an emotion. As such, the accuracy of one’s self-reported emotional state can be undermined by the following: (1) never having experienced that particular emotion, (2) being unable to accurately remember an instance of a particular emotion, (3) having a different meaning or understanding (than the researcher) associated with the emotional term or label, (4) being reticent to espouse the experience of negative emotions during their interaction with the CBLE due to social desirability, (5) time span between experiencing a
particular emotion and being asked to report the emotion, or (6) the self-report measure eliciting a different emotion (e.g., boredom) than that experienced prior to its administration (e.g., curiosity; Greene & Azevedo, 2010; Porayska-Pomsta et al., 2013; Veenman, 2011; Winne, 2010; Zimmerman, 2008). Given the potential for self-report measures to influence the emotions they purport to assess, and the unique influence of self-report measures on emotions relative to other psychological states and processes as measured by self-reports measures, further discussion is warranted.

In addition to the abovementioned reasons as to why participants may not accurately report their emotions, there are two ways in which asking students to report their emotional states could influence their emotions. The first possibility is that the self-report measure can induce an emotion itself by drawing the students’ attention. For example, a self-report measure could surprise a learner if they are engaged in deep concentration and are not expecting or warned prior to its administration. Alternatively, a student could feel bored by having to repeatedly complete the self-report measure if they are administered too closely together or frustrated to have their learning session repeatedly interrupted. These examples illustrate the need for testing and revising the spacing and length of self-report measures administered during learning sessions.

The second way that self-report measures may elicit an emotional response is related to research on meta-emotions, defined as emotions people have about having emotions (Bartsch, Vorderer, Manggold, & Viehoff, 2008). In other words, an emotion serves as the object-focus/stimuli that is evaluated in terms of its goal-conduciveness (e.g., is experiencing anxiety helpful for studying this chapter?) and other appraisal dimensions (e.g., controllability) as opposed to an external object or event. Emotions that elicit other emotions are referred to as “primary” emotions. “Secondary” emotions are those states that arise from attending to primary emotions. In the case of self-report measures, by having the student report their emotional states, they may attend to states they might otherwise have ignored that may elicit secondary emotions. For example, a student might report a 4- on a 5-point Likert scale regarding anxiety, but then feel surprised that they felt anxious during the task.

In response to the various drawbacks of self-report measures, there are several steps that researchers can take to address some of these issues, such as providing definitions for the emotions and emotional terms learners are asked to use in reporting their emotional experiences (D’Mello, Craig, Witherspoon, McDaniel, & Graesser, 2008; Harley et al., 2013, 2015). While
definitions of emotions may differ between researchers, this approach makes the researchers’ and participants’ operationalization of emotions more transparent. Additionally, researchers can administer self-report measures while participants are interacting with CBLEs, to reduce the likelihood that they will not accurately remember the emotions they were experiencing (Conati & Maclaren, 2009; Harley et al., 2013, 2015; Robison, McQuiggan, & Lester, 2009). Accuracy of recall for retrospective self-report questionnaires may also be improved by showing learners footage of their learning session and facial responses during the session (retrospective emotion protocols; D’Mello & Graesser, 2012a). Research in educational psychology has found evidence of differences in the emotions students report when assessed during versus after an academic achievement situation (e.g., a math test; Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013). Finally, in order to decrease item fatigue and the possibility of negative emotions, researchers can also use single-item questions to assess emotions (Conati & Maclaren, 2009; Harley et al., 2013, 2015; Robison et al., 2009).

In summary, self-report measures offer a readily available and widely accepted approach to measuring emotions that, when implemented thoughtfully and recognized as an offline method of data collection, can be effectively incorporated into CBLEs. Moreover, although they are not online measures, self-report data can be provided to emotionally supportive systems to administer emotion-regulating prompts and feedback when the learners’ emotional state is known (i.e., right after the administration of the self-report; e.g., Robison et al., 2009).

Log-Files

A recent addition to the battery of methods for measuring emotions is the use of log-file data. Log-files provide rich information about learner-system interactions, including navigational patterns based on the sequences and timing of use of textual and diagrammatic representations, content of learner-agent dialogue turns, timing of administration and scores from embedded assessment (e.g., quizzes), etc. (e.g., Azevedo et al., 2013). Given the richness of information in these logs, researchers have used them in a variety of creative ways to measure emotions, including exploratory data mining (Baker et al., 2012; D’Mello et al., 2008; D’Mello & Graesser, 2010) and analyzing specific language and discourse features (D’Mello & Graesser, 2012b). Each of these approaches, the first referred to as context feature mining and the second as language and discourse feature analyses, are now described and their advantages and disadvantages briefly reviewed.
Context Feature Mining

D’Mello, Graesser, Baker, and colleagues have used log-file data in addition to behavioral observations to measure learners’ emotions in a variety of CBLEs (Baker et al., 2012; D’Mello et al., 2008; D’Mello & Graesser, 2010). In order to use the log-file data to measure emotions, Baker et al. (2012) distilled a selection of log-file data pertaining to 58 features (i.e., different types of learner-environment interactions) involving both the specific actions in which students engaged (e.g., correct and incorrect actions, asking for help) and students’ past performance (e.g., previous actions that involved help requests for a specific skill). Baker et al. (2012) used a variety of data mining procedures to examine which student behaviors correlated with the four affective states they selected based on Graesser and D’Mello’s program of research on learner-centered emotions (boredom, frustration, confusion, and engaged concentration). In a recent study, the authors reported that their detector results currently have an average kappa of over $\kappa = 0.3$ (30% better than chance) ranging from $\kappa = 0.23$ (frustration) to $\kappa = 0.40$ (confusion; Baker et al., 2012). An example of features related to boredom include working slowly (e.g., taking longer than average to respond on the current step) and repeated pausing, regardless of accuracy.

D’Mello et al. (2008, 2010) have conducted similar research with AutoTutor examining a range of features related to learners’ conversational dialog turns with AutoTutor’s pedagogical agent, including response time, answer quality assessment, tutor directness, and tutor feedback. Results indicated that conversational cues were able to significantly, yet modestly, predict learners’ experiences of boredom, confusion, flow, and frustration after individual differences were parsed out (7% of the variance; D’Mello et al., 2008). Arroyo et al. (2009) have also made use of contextual features in predicting students’ emotions, but did not report the individual contributions they made to a multimethod systems’ detection accuracy.

The advantages of this method are its low equipment cost and potential to be scaled to a variety of CBLEs, so as to detect a range of different emotions online and in real-time. It also follows that CBLEs would be able to use the emotion classifications provided by these log-file analyses to respond in real-time to students’ emotional states. Log-files also have the potential to make distinctions between emotional states that may appear similar to some methods, such as heart rate and electrodermal activation measures, by providing contextual information. The shortcomings of using log-file data to measure learners’ emotions relate to this method’s infancy, including its kappa, which is lower than the sets of research
equipment that do use physiological sensors (AlZoubi et al., 2012; Calvo & D’Mello, 2010; D’Mello & Kory, 2012).

**Language and Discourse Feature Analyses**

Reviews of emotion detection methods in the field of affective computing have highlighted the prominence of speech, including the *sentiment* (i.e., emotion) with which written or spoken words are imbued (Calvo & D’Mello, 2010; D’Mello et al., 2008; D’Mello & Graesser, 2012a; Kapoor & Picard, 2005; Litman & Forbes-Riley, 2006). Sentiment analyses have gained traction in the CBLE research community, perhaps because of the interdisciplinary nature of many research groups that include expertise in psychology computational linguistics and natural language processing. CBLEs also provide rich text-based data through log-files of learner-system interaction and transcribed think- or emote-alouds (Craig et al., 2008). A variety of methods are available to mine text-data for sentiment features (for a review, see D’Mello & Graesser, 2012a).

In a recent study, D’Mello and Graesser (2012a) examined two approaches to inferring students’ emotional states from log-files produced by *AutoTutor*. One of these tools, the Linguistic Inquiry and Word Count Tool (LIWC; Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007), analyzes transcripts of text and counts the number of words that relate to a specific category that is indicative of a psychological process. In this study, a specific subset of relevant linguistic and psychological (cognitive and affective) features were selected. D’Mello and Graesser also made use of a deeper natural language processing techniques using Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004), a program that analyzes the cohesion relationships (e.g., coreference, pronoun reference, causal, and semantic cohesion) in *AutoTutor* tutoring dialogs. By combining the most telling features from their two text mining approaches, they were able to predict an average of 38% of the variance of students’ self- and observer-reported states of boredom, confusion, flow, and frustration.

The advantages of analyzing discourse and language features to infer emotions from learners’ interactions with CBLEs are similar to those of feature mining discussed previously with regard to providing researchers with low cost and scalable options for measuring emotions online and in real-time, as well as implementing these methods in classroom research. The biggest disadvantage of this approach is that it requires a CBLE capable of using natural language processing (for analyzing log-files), which is a highly sophisticated feature not included as a key feature of many CBLEs. Further
research examining the two novel approaches investigated by D’Mello and Graesser (2012a) is also needed to validate this approach, with a larger sample size and in different contexts (learning material; Newtonian physics). However, language and discourse analysis methods do appear to be promising approaches for measuring emotions because they can transcend the challenge presented by emotionally-flat language and discourse expressed by learners during their CBLE interactions (D’Mello et al., 2008).

MULTIMETHOD EMOTION CLASSIFICATION: IS IT WORTH IT?

The present section reviews research examining the joint use of different emotion detection methods to bolster emotion detection accuracy and answer research questions, such as: How effective are such approaches at accurately measuring emotions? And, what are some of the challenges and future directions for multimethod emotion classification? Detection accuracy as discussed in this section refers to the agreement between the emotional states detected by different sets of emotion measurement equipment (e.g., multiple modalities), one of which is being used as the “grounded truth” (i.e., standard) for determining the correct emotion. These grounded truth measures tend to be self-report measures or facial expression coding and are compared with other data channels postexperiment.

The first observation upon review of the research literature in which different sets of equipment have been used to measure emotions was that few studies have been done that compare the accuracies of each assessment method used. There are, however, several methodologically sound reasons for this paucity of research. First, the accurate alignment of methods with different sampling rates (e.g., continuous, fixed intervals) and different offsets (i.e., starting and stopping times regarding when data is being collected) can be both challenging and time consuming (Azevedo et al., 2010, 2013). The second major challenge is that data from different methods provides different information about emotions in terms of granularity (i.e., emotional dimension versus discrete emotional state). Relatedly, different methods often measure different subsets of emotions (basic versus learner-centered). Therefore, it is likely that these studies are rare because they require a tremendous amount of conceptualization and methodological groundwork to do properly, as well as access to (a) multiple methods for measuring emotions and (b) a measure to use as a grounded truth comparison (e.g., self-report).
A review of the literature on CBLEs reveals a handful of influential studies from different research groups (each corresponding to a different CBLE) that have endeavored to examine the combined effectiveness of multiple emotion detection methods. Researchers examining emotions in Wayang Outpost (Arroyo et al., 2009) and AutoTutor (D’Mello & Graesser, 2010) have reported similar results regarding the channels that provide the most accurate emotion classifications, namely facial expressions and contextual features. Specifically, D’Mello and Graesser (2010) were able to accurately detect an average of 49% of learners’ emotions (boredom, engagement/flow, confusion, frustration, delight, and neutral) from facial (35%); posture (32%); and contextual (log-file information, 38%) measurement methods (D’Mello & Kory, 2012). Arroyo et al. (2009) reported being able to accurately detect more than 60% of the variance of students’ experiences of confidence, frustration, excitement, and interest; the most from combinations of facial expressions and context with the exception of frustration (which was slightly better predicted using posture and context). Unfortunately, the results reported earlier by AlZoubi et al. (2012) did not include the measures previously examined by D’Mello and Graesser (2010) but did find that including ECG, EMG, and EDA sensors contributed to accurately classifying (κ = 0.25) learners’ emotions (flow/engagement, boredom, neutral, curiosity, confusion, delight, and frustration).

D’Mello and Graesser (2010) suggest that the similarity of their findings with those of Arroyo et al. (2009) imply that gross body movements are largely redundant when facial expressions and context data are provided. Recent results from Grafsgaard et al. (2014) using JavaTutor also support these findings, where posture was not predictive of either of the emotional states that were measured (engagement and frustration; although gesture was). Moreover, D’Mello and Graesser (2010) suggest that skin conductance (though refutable from later research with AlZoubi et al., 2012) and pressure exerted on pressure-sensitive hardware (mice) may also be redundant. Kapoor and Picard (2005), on the other hand, have found conflicting evidence in their set of emotion measurement equipment, specifically that posture alone was able to accurately classify 82% of learners’ states of interest. When combined with facial, context, and body posture features, they were able to accurately classify 86% of learners’ states of interest. A later study (Kapoor et al., 2007) demonstrated that their set of emotion measurement equipment was able to accurately classify 79% of the instances of frustration self-reported by students when a pressure-sensitive mouse and skin conductance bracelet were added to the set. Recent research by Harley and
colleagues (2015) contributed to this area of research by examining the agreement between data collected from automatic facial expression recognition software, skin conductance bracelets, and self-report measures of emotions during learning sessions with MetaTutor. They found high agreement rates when the emotional states identified by facial expression software and self-report measures were compared (75.6%), but low agreement rates when these two methods were compared with skin conductance bracelets.

Given that there are disagreements and methodological differences between such a small number of studies, specific recommendations concerning multimethod research with CBLEs are limited. Additional considerations regarding the generalizability of these results are: (1) the contextual nature of each study’s findings, which are with different populations (e.g., college students vs. high school students); (2) the use of different procedures for processing physiological data; and (3) small sample sizes (ranging from 20 to 67). Therefore, the scope of this review will briefly widen now to consider findings from other research programs (not necessarily CBLEs) within the affective computing community in a meta-analysis conducted by D’Mello and Kory (2012). This review looked at the individual and collective accuracy ratings that different combinations of methods were able to achieve by averaging their performance across the subsets of emotions they classified.

Unfortunately, findings from this meta-analysis revealed a considerable amount of variance in proportional classification accuracy for each of the individual emotion detection methods (face: 0.20–0.90; voice: 0.36–0.87; text: 0.65–0.81; posture/gesture: 0.32–1.00; EEG: 0.56–0.67; and context: 0.38–0.57). It is possible that the considerations mentioned above (e.g., different analytical procedures) may play a role in these within-method classification accuracy differences. D’Mello and Kory (2012) also reported that additive gains in the accuracy scores were modest ($M=8.12\%$ improvement) when the accuracy of the single best method (e.g., facial expressions) was compared with models with additional methods (e.g., gesture/posture). Gains were even more modest when only classifiers that were trained on natural or semi-natural data were used ($M=4.39\%$ improvement) compared with acted data ($M=12.10\%$ improvement). In other words, models that were built using naturally unfolding emotion expression information (e.g., facial action units) to identify emotions (e.g., happiness, frustration) were less effective at accurately classifying the emotion being expressed than when the emotion expression information was artificial (i.e., posed). D’Mello and Kory (2012) note that the gains associated with the naturalistic
classifiers should be seen as the state-of-the-art rather than classifiers based on posed data, because of their alignment with the object of measuring authentic emotions. Their meta-analysis also revealed a considerable amount of redundancy between the various methods (e.g., facial expressions, voice features, text) where the classifiers used to analyze the different sets of features were unable to substantially improve the overall detection accuracy of emotion measures.

These results might give pause for the combining of different methods for detecting emotions. However, it is important to note that this field is still nascent and the results of D’Mello and Kory’s (2012) meta-analysis are therefore far from conclusive regarding the utility of combining multiple methods. Two important avenues for future work are research and development concerning the development of different classifiers and the use of naturalistic rather than experimentally induced emotions. Advances in classifier methods may allow researchers to leverage current emotion detection methods and close the gap between user-dependent and user-independent classifiers, which is critical for creating emotionally supportive CBLEs. More research detecting naturalistic emotions may also provide a more objective assessment of the state-of-the-art, as results with experimentally induced emotions can exaggerate the observed accuracy of sets of different emotion measurement equipment.

THEORETICAL AND ANALYTICAL CONSIDERATIONS IN MEASURING EMOTIONS

There is also theoretical work to be done. For example, it is widely agreed that emotions are multicomponential (Gross, 2010; Pekrun, 2011), but not whether different emotional expression components (experiential, behavioral, and physiological) will align in terms of expressing a common emotional state (i.e., coherence; Ekman, 1992; Gross & Barrett, 2011; Pekrun, 2011). Therefore, uncertainty exists in terms of whether divergent emotional classifications between methods and modalities that reflect different expression components (e.g., behavioral: automatic facial expression recognition software versus physiological: electrodermal activation devices) are the result of poor calibration or signal filtering or a feature of emotions themselves (see Harley et al., 2015). Moreover, expectations are not clearly defined regarding how differences in emotion classifications should be interpreted, especially in the absence of an empirically validated grounded truth measure, when encountered in measurements between methods that
Examine emotions from the same emotional expression component, but different modalities (e.g., behavioral: automatic facial expression recognition software versus automated postural sensors). Incongruous classifications from similar modalities and expression components, but different methods (e.g., physiological: electrodermal activation devices and heart rate measurement devices) can also raise the question of which one is right? It is therefore important to advance our theories of emotion (using empirical results), so that they may provide appropriate and contextualized hypotheses for comparing individuals’ emotional responses from experiential, behavioral, and physiological expression components that are measured using different modalities and methods.

In addition to other important future directions mentioned previously (theoretical and classifier advances, more research with multiple methods of emotion measurement, naturalistic data), the statistical techniques used to analyze learners’ emotions must evolve if we are to take full advantage of data with non-normal distributions, continuous, and nonparametric characteristics. These advances are critical if we are to evolve from the current best practices of time-sampling (which uses only portions of available data), averaging measures across episodes or sub-episodes together, and affective-state transitions (the likelihood of transitioning from one emotional state to another; Baker, Rodrigo, & Xocolotzin, 2007; D’Mello & Graesser, 2012a; McQuiggan, Robison, & Lester, 2010). These approaches do provide information on the temporal nature of emotions, but at a far larger grain size than we are currently able to measure with trace methods (e.g., log-files, facial expressions, physiological methods). For example, with time sampling, we can measure fluctuations of emotions over the course of a learning session, but not the moment-to-moment fluctuations that illustrate when an emotion was elicited and why. Crossing this analytical barrier will also help us answer important and informative theoretical questions about the temporal nature of emotions that many researchers have begun to investigate, but cannot presently fully answer. In other words, our analytical approaches have not yet caught up with our methodological ones (see Calvo & D’Mello, 2010, and Mauss & Robinson, 2009, for a discussion of these and other challenges and future directions related to emotion measurement).

CONCLUSIONS AND RECOMMENDATIONS

In summary, there are a variety of methods, each with advantages and disadvantages, that can be used to measure, model, analyze, and predict learners’
emotions. Not enough research has been done, however, to draw definitive conclusions about which methods are most accurate in measuring emotions, especially naturalistic emotions which make up a small percentage of emotions evaluated in multimethod studies in the field of affective computing. What is known, however, is that there are an increasing number of ways to measure emotions both within and across modalities, each in different phases of maturity. One emerging area of emotion detection is using eye-tracking devices to detect information, such as where participants are looking and how long they are fixating on specified regions of interest. For example, D’Mello, Olney, Williams, and Hays (2012) have used the Tobii T60 eye-tracker in order to infer and respond to participant disengagement from a learning session with the Guru system. Another study was conducted by Jaques, Conati, Harley, and Azevedo (2014), in which the same model of eye-tracker was used to successfully identify gaze behavior and predict learners’ experience of boredom ($\kappa = 0.33$) and curiosity ($\kappa = 0.42$).

Amongst the myriad of methods for measuring emotions the most empirically grounded approaches for measuring emotions with CBLEs are facial expression coding and self-report measures. Log-file data stands to provide scalable, low-cost, and promising options for measuring emotions. Physiological patterns are also promising and are becoming more scalable, but still have more ground to cover before they are considered empirically sound enough to be used alone as a method. The one possible exception is EEG research, which has impressive accuracy levels, but a paucity of research and replication (Heraz et al., 2007). Lastly, posture sensors are the most disputed measure in terms of accuracy and therefore might be the easiest to leave out of research using multiple methods to measure emotions, if one had to choose between them. Although this review has focused on reviewing and evaluating the available methods for measuring emotions, it has also noted that data from these methods can be harnessed for the purpose of responding (when necessary) to learners’ emotional states, so long as online data is processed in real-time (as opposed to after the experimental session). Offline self-report data can also be used, but interventions are limited to the states that learners report experiencing and when these reports are administered by the system or volunteered by the learner.

In providing this review of the state-of-the-art of emotion measurement with CBLEs, several recommendations for exploiting the affordances and minimizing the constraints of different methods have been provided. Future directions have also been outlined for multimethod research and a case has been made for advancing our analytical and theoretical frameworks.
ACKNOWLEDGMENTS

This research was supported by a Joseph-Armand Bombardier Canada Graduate Scholarship for Doctoral research from the Social Science and Humanities Research Council (SSHRC) of Canada and a Postdoctoral Fellowship from the Fonds Québécois de la recherche sur la Société et la culture (FQRSC). The author would like to thank Roger Azevedo, Nathan C. Hall, Susanne P. Lajoie, and Alenoush Saroyan for their feedback on this chapter.

REFERENCES


D’Mello, S. K., & Kory, J. (2012). Consistent but modest: A meta-analysis on unimodal and multimodal affect detection accuracies from 30 studies. In G. Castellano, & K. Masse (Eds.), *Proceedings of the 14th ACM international conference on multimodal interaction* (pp. 31–38). New York: ACM.


FaceReader (5.0) *Computer software*. Wageningen, the Netherlands: Noldus Information Technology.


CHAPTER 6

Designing Tools that Care: The Affective Qualities of Virtual Peers, Robots, and Videos

Yanghee Kim, Diantha Smith, Jeffrey Thayne
Utah State University, Logan, Utah, USA

THE INTEGRAL NATURE OF AFFECT AND COGNITION

The evidence for the integral relationship between affect and cognition has been accumulated over the last three decades, especially with the advance of research in neuroscience (Damasio, 1994). Currently, it is widely believed that affect influences social thinking and memory and facilitates effective decision-making. Also, positive affect reduces cognitive and judgmental errors. Materials associated with a users’ current emotional state are more likely to be activated, recalled, and used in constructive cognitive tasks (Bower & Forgas, 2001). People process affect-congruent materials more deeply, elaborate information from the materials more effectively, and learn from these materials more fluidly than they do with affect-incongruent materials (Bower, 1981). Moreover, our cognitive processes are responsive to environments signaled by our feelings. That is, affect signals influence the spontaneous choice of a processing style compatible with an individual’s goals and task demands. Sad individuals are likely to spontaneously adopt a systematic, detail-oriented, and bottom-up strategy; happy individuals may prefer simple heuristics, explore new procedures and possibilities, and adopt an unusual, creative, top-down strategy (Schwarz, 2002). In general, learners with a positive affect spend longer in the task and learn more when compared with learners with a negative affect.

“Affect” is a context-specific entity and is closely aligned with social contexts (Saarni, 2001). It seems natural that the social context of learning plays a significant role in shaping students’ motivation, learning behaviors, and academic outcomes in classrooms. In classrooms, the affective states of teachers, peers, and others function as social contexts and influence learners’ affective characteristics, such as emotions, self-conception, and motivation (Sutton & Wheatley, 2003). Recently, achievement gaps in challenging subject matters...
(e.g., mathematics and science) among different groups of students in public schools are often attributed to a less supportive learning context and undesirable social influences (e.g., stereotyping). Some groups of low-achieving students in conventional settings lack the instructional support that might motivate them to sustain and succeed in those areas.

Although three components of scaffolding (perceptual, cognitive, and affective) are necessary for effective learning and motivation (Stone, 1998), many conventional computer-based tutoring environments have focused primarily on assisting learners in the cognitive processes of learning. They have often neglected to implement the perceptual and affective aspects of scaffolding. A recent study reports that individuals with low self-esteem are less receptive to personal support compared with those with high self-esteem; the low self-esteem individuals become more receptive when the support validates their negative feelings (Marigold, Cavallo, Holmes, & Wood, 2014). The authors believe that minimizing a learner’s negative feelings and supporting positive affect is crucial for cognitive engagement. We have investigated various ways to create an affective learning context in technologically enhanced environments. Our experience suggests that designing tools to afford a caring context could help boost learners’ positive affect and also close the achievement gap of diverse groups of students in public education (Kim & Lim, 2013). With this goal in mind, we have used three tools: virtual peers (animated, on-screen characters), humanoid robots, and online videos to promote positive affect of learners ranging from pre-schoolers to college students. We have paid special attention to traditionally marginalized groups of students in conventional classrooms and observed how those technologies helped address their unique needs.

VIRTUAL PEERS

Virtual peers are animated peer-like characters. Human/computer interaction research argues that computer users tend to expect computers to be like social entities and build humanlike relationships with an animated character (Nass & Moon, 2000). This trend has stimulated the popular use of virtual peers in computing environments for educational and commercial purposes. The provision of simulated social presence and peer interaction may make virtual peers distinct from traditional computer-assisted tutoring, seemingly offering a unique instructional opportunity. Although the theoretical frameworks for virtual peer technology and its educational effectiveness vary, it is clear that a virtual peer plays a distinct social (even persuasive) role for
learners across age groups. For example, kindergarten children listen to their virtual peer’s stories very carefully and, afterward, mimic the peer’s linguistic styles (Ryokai, Vaucelle, & Cassell, 2003). Secondary school students prefer a virtual peer that looks similar to them over a dissimilar peer (Kim & Wei, 2011). College students expect their virtual peer to have a nice and friendly persona (Kim, 2007). As aforementioned, our affect influences our thinking processes, and we can respond more adaptively to our environments when feelings signal our cognitive processes. Therefore, a virtual peer could be designed to promote learner confidence and engage the learner effectively in challenging learning tasks. Also, the virtual peer’s affect might be adjusted to the nature of the learning challenges and designed to stimulate a learners’ positive affect, which in turn could signal the appropriate processing strategy for the task.

**Virtual Peer Affect**

In general, affective computing refers to a computer’s capabilities to recognize a user’s emotional states, to express its own emotions, and to respond to the user’s emotions (Picard, 1997). Each capability requires different kinds of technologies and resources to implement; hence, investigating each capability separately might provide more accurate information on their individual effectiveness in achieving intended goals. Given that affect recognition is engineered with auxiliary hardware, we focused on a virtual peer’s emotional expressions and empathetic responses to help achieve desired instructional outcomes. In another study, we examined the effectiveness of a virtual peer’s emotional expressions and the effectiveness of its empathetic responses, separately, in two experiments (Kim, Baylor and Shen, 2007). In the first experiment on peer emotional expressions, 142 college students in a computer-literacy course voluntarily participated in a web-based e-Learning module taught by a virtual peer for one class hour. During the instruction, the peer tutor demonstrated positive or negative emotions or no emotions at all through its facial and verbal expressions. After working at the module, the college students rated the peer that showed constant positive emotion as significantly more engaging and facilitating to their learning. In the second experiment on the peer’s empathetic responses, 56 college students in a required course participated in a web-based lesson taught by a virtual peer as a mandatory class activity. The learning task was an essential part of the curriculum the students had to master, so the students were expected to engage in the task seriously. This situation might have instigated increased emotional arousal in the students. During the lesson, they were
asked to express their current emotional states by clicking emoticons at the bottom of the screen. The virtual peer either presented empathetic messages responding to their emotional states or did not respond at all. The results showed that the students who received the virtual peer’s empathetic responses rated the peer as significantly more engaging and facilitating to their learning. Also, with the empathetic peer, the students showed significantly higher self-efficacy in the learning task than those with the nonresponsive peer.

**Affective Role Models**

Another perspective we have taken is the use of virtual peers as role models. Grounded in the classical theories of social modeling and attribute similarities, we have designed a virtual peer to serve as coping models or role models. That is, the peer responds sympathetically to the challenges the learner faces and/or demonstrates desirable behaviors that will lead to successful learning outcomes. According to social modeling theory (Bandura, 1997), attribute similarities (the similarities of personal characteristics between a social model and a learner, such as age and gender) are considered a determinant of successful modeling. The more similar a model is to a learner, the greater the probability that the learner will repeat the model’s actions, and similarity-attraction theory indicates that people are more attracted to a person who is similar to themselves (Berscheid & Walster, 1969). This attraction influences both their interpersonal relations and behaviors. More recent research also reports similarity attraction between learners and a virtual peer. High school boys and girls preferentially chose a same-gender peer and a same-ethnicity peer as their tutor in online mathematics learning (Kim & Wei, 2011); college students of color chose same-ethnicity peers significantly more (Moreno & Flowerday, 2006), and also evaluated similar peers significantly more, positively than dissimilar ones (Baylor & Kim, 2004).

It is known broadly that many adolescent females are less willing to participate in mathematics and science learning compared with their male counterparts. This is often attributed to the social influence of family, friends, and teachers who might impose gender-related social stereotypes and expectations on the girls (Clewell & Campbell, 2002). These girls lack instructional support and encouragement for their intellectual pursuit in mathematics and science classrooms. Feminist scholars argue that many females are better motivated when the learning environment affords social relations and collaborative interactions (Boaler, 1997, 2002). The provision of supportive relationships is
critical for many females’ intellectual pursuit of challenging topics and perseverance in these topics (Crosnoe, Riegle-Crumb, Field, Frank, & Muller, 2008).

Given this educational challenge, we have developed an algebra-learning environment called *MathGirls*, in which a virtual peer provides middle-grade girls with verbal encouragement and persuasion, to counteract unconstructive social influence. Research has indicated that a human-like agent’s presence primes human-to-human social interactions and motivates learners to converse with the agent and build interpersonal relations with it (Moreno, Mayer, Spires, & Lester, 2001). Thus, it was expected that the girls in the study would build social relations with the peer and thereby enhance their positive affect and sustained engagement in learning. We questioned whether virtual peer presence would influence middle-grade girls more positively than boys.

A total of 120, 9th-grade boys and girls took virtual-peer-based algebra lessons daily for 1 week. The peer presented personalized instructions and social and empathetic messages during the instruction. The results supported the expectation (Kim & Lim, 2013). Virtual peer presence had a significantly more positive influence on girls’ affect than it did on boys’. More specifically, the girls in the study evaluated the virtual peer significantly more positively than the boys, with the most positive evaluations from minority girls (here: Latinas). The girls showed significantly more positive attitudes toward learning with the peer than did the boys. Girls significantly increased their self-efficacy in learning algebra concepts after working on the peer-based lessons, whereas males did not show an increase. Follow-up interviews confirmed that girls built more developed social relationships with their virtual peer in the learning process. Another encouraging result was that both boys and girls increased their learning significantly after working in this socialized learning environment.

HUMANOID ROBOTS

Robots have been common in industrial settings for years, but during the last decade, their popularity as an educational resource has risen dramatically (Johnson, 2003). In the early 1990s, LEGO Mindstorms (http://mindstorms.lego.com) became a frontrunner in the effort to improve engagement and mastery of science, technology, engineering, and mathematics (STEM) topics for elementary, middle-school, and high school students. While some educational programs still focus primarily on “robots in education” (i.e., robots as tools for learning robotics), recently more and more educators
and researchers have become interested in “robots for education” (i.e., robots as mediators of learning in any area) (Shin & Kim, 2007, p. 1040). Today, trends in educational robotics include a wide range of robot applications in a variety of different subject matters (Rusk, Resnick, Berg, & Pezalla-Granlund, 2008). One subject where robots have become especially important is English as a Second/Foreign Language (ESL/EFL). In Japan, Korea, China, and other countries pursuing innovations in educational technology, EFL learning is the domain that has most actively used robot assistants. Due to a growing understanding of the social nature of acquiring language, robots’ involvement in language learning has sparked many questions about the ability of machines to replicate not only human speech and gestures but also human affect. There is a growing need for research that examines the value of robots as learning partners and helps developers design robots that accurately and effectively portray human affect.

Robots and Affect

When discussing robots, it is important to differentiate between robots as learning tools and robots as partners that can cooperate with, teach, and even build relationships of trust with users (Kidd & Breazeal, 2004). While humans expect to use tools, they do not expect to use partners, and this difference in expectations leads to different kinds of interaction. One major component of human-to-human interaction is affect, and many researchers have demonstrated that a robot’s display of affect can have a significant impact on the way the robot is perceived and on the robot’s ability to guide, engage, and help humans (Hudlicka et al., 2009; Leite, Castellano, Pereira, Martinho, & Paiva, 2012). Affect is especially important in robots that serve as partners or tutors because humans recognize their physical presence as evidence that robots are real, and as a result, users are more likely to perceive them as human rather than they would an animated or virtual character on a screen (Kidd & Breazeal, 2004).

The affect displayed by a robot is a major factor in its ability to elicit positive perceptions and engage users (Kanda, Shimada, & Koizumi, 2012; Leite et al., 2012). Robots that can display emotion are more likely to motivate learners (Shin & Kim, 2007) and socially supportive behavior from robots has resulted in improved performance in children learning a new language (Saerbeck, Schut, Bartneck, & Janse, 2010). All of these advantages make affect in robots more than just a nice feature; rather, it seems that affective affordances may be some of the most essential features a robot can possess.
The authors use the humanoid robot *Atti* in an ongoing project on educational robotics (http://www.create.usu.edu/projects.html). Atti’s purpose is to help ELs (English Learners) to learn English in their first few years of preschool and elementary school. The robot system is equipped with three types of sensors (optic, touch, and proximity) and the robot’s movement is controlled by Android smart-phone apps via Bluetooth technology. A phone is cradled on the robot’s head implying the robot’s visible brain. The integration of smartphone and humanoid hardware systems places Atti in an affordable range and thus solves the problem of cost that can often be a roadblock in robot development and broad use.

A sociable robot, such as Atti, is especially valuable to ELs for several reasons. First, all learners are more likely to engage in the learning tasks and perform at a high quality when they are given appropriate resources, opportunities, and environmental conditions (Brophy, Biswas, Katselberger, Bransford, & Schwartz, 1999). Instruction that attends to learners’ needs is more likely to enhance their engagement and leads to the achievement of goals (Ryan & Deci, 2009). In previous studies, students engaged more with robots that could actively interact with them and display emotions (Kanda et al., 2012). Feeling a lack of instructional support in classrooms, ethnically diverse students seem especially prone to building more developed social relationships with artificial beings (Kim & Lim, 2013). The ability of minority students to develop social skills through technology deserves special consideration, since EL children often struggle to fit in with native English speaking peers. During the design process of our project (Atti’s educational app), we tried to focus on filling the affective needs of EL students, especially the need to feel motivated and included in learning. Social robots can provide a valuable mediator between isolation and full human-to-human contact, and during user testing, we found Atti capable of providing interaction in both one-on-one and small group situations (Kim, Smith, Kim, & Chen, 2014). Furthermore, when using Atti, children are placed at the center of play; the robot pays full attention to them, and, as other studies have demonstrated, the relationship built between the robot and child can positively contribute to the children’s learning engagement, motivation, and performance (Saerbeck et al., 2010; Shin & Kim, 2007).

At first glance, Atti might not seem capable of demonstrating very much emotion. The robot cannot change its facial expression but instead relies on LED blinking to signal shifts in emotion, and its movement is limited to its
feet and hips. During user testing, it only communicated through recorded speech, noises, and flashing lights. However, research suggests that robots actually need to do very little to be perceived as social agents (Kozima, Michalowski, & Nakagawa, 2009). Toddlers perceive a robot called Keepon as a social being, even though it is limited to communicating through rudimentary movements and flashing lights (Kozima et al., 2009). Anyone who has seen the movie Star Wars and is familiar with the character R2-D2 has also seen evidence of this. Although R2-D2 is bulky and can only produce beeps and flashes of lights, this robot is one of the most beloved characters in the popular movie series (Davis & Pakowski, 2013). Additionally, Leite et al. (2012) point out that if a robot can verbalize (but not recognize human speech), this does not hinder its ability to interact verbally with users. It seems that the key to successful verbal interaction is designing a robot that can respond appropriately to both the user’s expectations and the task requirements (even with simple signals), rather than creating a robot that can perfectly process and produce speech.

**Designing for Children: Interplay Between Designers and Children**

Despite Atti’s simple design, the robot was able to encourage users to actively engage with it in various ways. None of the children who worked with Atti seemed to notice that the robot was unaware of what they were saying because all of the robots’ verbalizations fit the activities and matched the child’s performance. In fact, we used the Wizard of Oz method of user testing during the process of Atti’s design and development (Riek, 2012). This method was invaluable in creating believable affective speech and especially helpful in identifying which instructional behaviors are most relevant to their particular learning outcomes. Through role-playing in the position of the robot, our designers were able to have a much clearer idea of what the robot should say, when the robot should say it, and how the robot should say it. It can be easy to overlook the simplest kinds of human interaction (such as repeating words or changing our tone of voice), but including these relatively small details into robot applications can lead to a much more authentic and effective learning experience.

For example, during a session of user testing, the tester noticed that when she tried to elicit a response from a child, she would typically repeat the target word three times and use different variations in her voice (high and low tones, different emphasis on syllables, silly voices, etc.) to encourage the child to respond. We implemented this same pattern of repetition and variation to the robot’s dialogue. Also, after singing a song about a triangle, the
robot would say: “You say it! (pause) triangle (pause) triangle (pause) triangle (pause).” This relatively simple addition made a huge difference not only in each child’s willingness to respond, but it also added more fun to the activity as children tried mimicking the playful tones of the robot’s speech. Again, the issue was not so much how much the robot could say or understand, but how accurately it could respond in an engaging, credible, and appropriate way to the user.

Despite the incredible ability of robots to elicit or stimulate learners’ affective reactions with limited mobility and expression, our team discovered that robots have an even greater (and previously untapped) social ability: helping learners collaborate with other learners (Kim et al., 2014). One of the final sessions of field-testing with Atti occurred in the media center of a public elementary school, where students speak both English and Spanish in all of their classes. Several classes came in and out of the media center, while individual children worked through a series of activities with the robot. The team was about to wrap up for the day when two 7-year-olds from another class approached the team leader and asked if they could play with the robot. This led to the discovery of the potential for small group interaction instead of only one-on-one interaction with the robot. As the team watched, the two children voluntarily took turns solving the tasks presented in the robot application. Rather than trying to beat one another, the boy and girl collaborated to score points in the games and find the robot’s secret passcode in the book. As the design team watched them interact with each other and the robot, they became more and more excited about the ways the robot became a tool to support English language learning and, more importantly, a venue for children to collaborate and socialize, which serves as a building block for language learning. In other words, though it will take time before robots can mimic the full range of adult human social interaction, if a robot can help humans to interact with their peers—particularly those who struggle to interact because of language or other barriers—then the positive effects of social learning can be accessed both through the robot itself and through activities mediated by the robot and others.

ONLINE VIDEOS

Instructor-Learner Interaction in Online Learning

In higher education, online learning has been growing rapidly in recent years due to ease of access and scheduling flexibility (Kim, 2012; Robinson & Hullinger, 2008). One popular technology that is increasingly used to
deliver learning content is online instructional videos, which are videos pre-recorded by the instructor and placed online for convenient access by the learners. Online instructional videos might be a potentially powerful medium due to instructors’ visibility that adds a sense of social presence to, otherwise, socially bleak environments filled with text and graphical information. Social presence in online learning seems to be crucial because low completion rates in online learning are often attributed to the lack of affective support and interpersonal relationships between the instructor and the learner (Kim, 2012).

Affective experience of the learners is a valuable and vital part of the learning process (Micari & Pazos, 2012; Sakiz, 2012; Xiao, 2012). A lack of interaction between learners and instructors has a negative impact on the learner’s affective experience and can interfere with the success of online learning (Kim, 2012). Similarly, interaction with the instructor is one factor that significantly influences students’ satisfaction and perceived learning gains (Swan, 2001). Although they are limited to one-way interaction, online videos may have an advantage over the aforementioned tools (virtual peers or humanoid robots). That is, the instructor is an actual human who can present all the affective cues that humans do. With virtual peers or humanoid robots, these affective cues (such as facial expressions, tone of voice, etc.) have to be artificially defined. Adult learners might discriminate between a genuinely human presentation and an artificially programmed one. While it is clear that humans can and do attribute affective qualities to machine entities, it is also clear that no amount of clever programming can completely recreate the genuineness of a human face and voice, and all of the emotional qualities that they convey. For this reason, online videos may be used for college students as a way to support their learning processes while preserving the human element of instruction.

**Compensatory Strategies**

Indeed, there may be compensatory strategies available to help increase positive learner affect while learners are attending online video instruction. Developing a strong relational rapport between the learners and the instructor can increase positive student affect and decrease negative student affect (Angelaki & Mavroidis, 2013; Sakiz, 2012). This relational rapport can be built as the instructor demonstrates that he or she cares about the students (Teven, 2007); praises their efforts and achievements (Arghode, 2012; Komarraju, Musulkin, & Bhattacharya, 2010; Marchant & Anderson, 2012;
Xiao, 2012); and presents him or herself in an approachable way (Micari & Pazos, 2012). In particular, Micari and Pazos (2012) discovered that three variables were correlated with the rapport between the learner and the instruction: viewing the instructor as a role model, the approachability of the instructor, and the respect that the instructor shows for his or her learners.

While each of these approaches to improving instructor-learning rapport is most effective in in-person settings, it may be possible to cultivate this rapport and improve student affect in online instructional videos (Velasquez, Graham, & Osguthorpe, 2013). It is unknown yet how the learner-instructor relationship could be developed and promoted in online video-based learning (Angelaki & Mavroidis, 2013), but some attempts have been made to investigate the issue.

The authors have recently studied whether the relationship-building strategies that are effective in in-person contexts are also effective in an online context, when used in an online video (Kim & Thayne, 2015). If so, then these relationship-building strategies may serve as compensatory measures that could help alleviate some of negative aspects that result from the static nature of the video presentation. In the study, we developed two versions of a four-lesson, college-level statistics module. In one version of the instruction, the instructor used relationship-building strategies in the design and presentation, including building up the instructor as a role model, being approachable, and showing that the instructor respects the learners (Micari & Pazos, 2012; Young, 2006). The instructor used a friendly and warm tone of voice and included colloquialisms, provided anecdotes about his own experience with statistics (including stories of how he struggled with certain concepts and how he still continued to use the information), and used encouraging language toward learners. In the other version, the instructor used none of these strategies, but simply presented the curricular material in a straightforward manner.

Study participants were college students who enrolled in a required introductory statistics course offered for general education credits. The four online video lessons replaced 1 week of statistics classroom lectures in the middle of the semester. The results showed a statistically significant effect on the students’ attitude toward the instructor and the learning material. For both groups, the students’ positive attitudes toward the material and the instructor decreased over the period of time while learning the material using online videos rather than in person. However, those who watched the videos using the relationship-building strategies experienced significantly
less decrease in their positive attitudes toward the material and the instructor. If we assume that attitude is closely related to (and mediated by) learner affect, we can conclude that these strategies may have positively influenced the affect of the students. From the study, we conclude that while robust interaction between the learner and the instructor should always be encouraged, online videos that do not allow for this interaction can be enhanced—and perhaps deliberately designed to positively influence student affect—by incorporating sound relationship-building strategies into the design and presentation of the video.

**CONCLUSION**

Emotion research suggests that, to help learners with various challenges, making efforts to build learners’ positive affect should be an essential, preliminary step to any type of instructional support. Tools, such as virtual peers and humanoid robots, are relatively free from the biases in the real world, so they might provide a safer learning environment for students who can be marginalized in regular education classrooms for various reasons. Also, by adding some affective pedagogical strategies, the relatively simple tool of instructional videos can be utilized to enhance social and affective qualities in online learning.

Virtual peers, robots, and videos can be effective and affective tools when they are designed carefully with awareness of affective influence on users. When creating programs with virtual peers, designers may keep in mind the power of a peer to positively engage learners and enhance learning experiences. When designing applications for educational robots, designers should consider how affect opens up several exciting opportunities for further exploration and innovation. It seems that affect is more related to the robot’s capacity to replicate human behavior than a robot’s capacity to replicate a human’s physical abilities. Last but not least, when incorporating videos into online learning, instructors should first be aware of the affordance and limitation of the medium, as well as of the research in this area. Unless instructors take care to ensure that students are engaged on an affective level, instructors who use online videos may see their students’ positive attitudes decrease. It should be further explored how online videos can be deliberately designed to increase positive student affect. The case for doing so is strong, but the research on precisely how to do so is still weak. As our study suggests, the use of relationship-building strategies may help.
Overall, it is clear that rendering social and relational learning contexts to educational tools is feasible and crucial to success in technology-based learning. Building from our initial findings, subsequent research needs to be done on how those technologies can mediate and encourage positive learning experiences in both one-on-one and group settings.

REFERENCES


CHAPTER 7

Emotional Design in Digital Media for Learning

Jan L. Plass, Ulas Kaplan
CREATE Lab, New York University, New York, USA

INTRODUCTION

More and more evidence still points to something that practitioners in education have known for millennia: human learning and performance cannot be simply described from a cognitive or even sociocultural perspective alone. In order to fully understand how we process the world around us, we need to consider our affective responses to the information we perceive. This is especially important for the designers of digital educational materials, as these materials offer many important opportunities to incorporate emotional considerations. However, few if any theories of learning with media consider emotions, and if they do, they do so only in very limited ways.

In this chapter, we first review definitions of key terms related to emotion and learning, and summarize research on emotional design in digital media for learning. We then present a theoretical framework of learning from digital media that emphasizes the integration of emotional and cognitive processing and of related design factors, and describe a resulting research agenda for the study of emotional design.

DEFINING EMOTION, MOOD, AFFECT

There are many definitions of “emotion,” “mood,” and “affect,” and a lack of general agreement regarding this terminology among scholars. We will selectively focus on two comprehensive conceptions of emotion (Roseman, 1984, 2011; Russell, 2003). Because of their sophisticated and relatively dynamic nature, we have found that these two models have strong potential for informing instructional design.

Roseman’s (1984) initial model built on Arnold’s (1960) conception that emotion is a tendency to move toward or away from a specific object. This view highlights the motivational and behavioral aspects of
emotion (Roseman, 2011). This motivational foundation of emotion corresponds to action readiness in Frijda’s theory (1986), emotional interpretations according to Lewis (Lewis, 2000; Lewis & Douglas, 1998), and emotion schemas in Izard’s model (2007, 2009). In addition, building on Lazarus (1968, 1991), Roseman (1984) conceptualized emotion as a coping response. In this view, each emotion represents a different mechanism or strategy (Roseman, 2011) for adaptation to life events.

By incorporating various characteristics of emotions that have been emphasized by different researchers, Roseman (2011) proposed a comprehensive conceptualization of emotions as syndromes. In this view, emotions are characterized by five different components, which Roseman (2011) called response types:

1. Phenomenological component (specific thoughts and feelings)
2. Physiological component (characteristic bodily response patterns)
3. Expressive components (specific manifestations in face, voice, and posture)
4. Behavioral component (action tendencies)
5. Motivational component (corresponding goals).

Roseman stressed that the combination of these five components corresponding to a particular emotion constitutes the strategy of that emotion. Hence, each emotion syndrome is a distinct strategy to facilitate adaptation in a particular situation.

Russell (2003) provided a similarly comprehensive conceptualization of emotions. According to Russell, two basic dimensions constitute core affect: (a) pleasant versus unpleasant affect; (b) arousal. An emotional episode is a function of core affect, in addition to two other experiences, namely, “perception of affective quality and attributed affect” (p. 150). Depending on the quality, intensity, and content of the individual’s experience on these three dimensions, different types and intensities of emotion can be experienced.

Russell’s “perception of affective quality” is similar to the experience that is commonly referred to as appraisal. In Russell’s framework, affective quality is subject to the same two dimensions (pleasure and arousal) as core affect. The difference is that affective quality is experienced as located in stimuli, such as various objects that evoke emotion. For example, when an individual perceives a virtual agent on a computer screen as pleasing, she is experiencing the perception of affective quality.

Russell proposes core affect and the perception of affective quality as the two primitives of his framework. In this view, they are fundamental experiences that cannot be reduced into psychological components. As core
components, they can be used to explain the formation of more complex experiences. Hence, the third factor in Russell’s framework, attributed affect, occurs when the individual psychologically combines core affect with a particular object. In this combination, the object is perceived as the cause of one’s affect. Russell stresses that attributed affect “guides attention to and behavior directed at the Object” and “is the main route to the affective quality of the Object” (p. 149).

Russell defines mood as an ongoing and a free-floating core affect that is generally not attributed to an object. An emotional episode typically has a shorter duration but a higher psychological complexity, as it also involves perceived affective quality and attributed affect. The feelings that people recognize in themselves as fear, anger, frustration, compassion, and joy are all examples of emotional episodes. On the other hand, such emotion categories (which Russell calls prototypes) do not define emotional life. Rather, according to Russell (2003), emotional life is characterized by a continuous experience of fluctuating core affect, with frequent experiences of affective quality and attribution of feeling to a specific object (p. 152). Occasionally, when the components of an emotional experience emerge in ways that closely match a specific emotion category, then the prototype of that emotion is experienced.

According to Russell, an emotional episode is not biologically or socially determined; rather it is psychologically constructed. By implication, there is substantial interpersonal and intrapersonal variation in emotional experiences, even among those that can be categorized under the same label of emotion, such as fear or anger. Consistently, an emotional episode “is constructed anew each time to fit its specific circumstances” (p. 151). To the extent that Russell’s emphasis on construction reflects change over time (van Geert & Steenbeek, 2005), his approach can be seen as a dynamic account of emotion.

Another key aspect of Russell’s framework has to do with the connection of emotion to action. While it is commonly held that emotions are action tendencies, Russell views this tendency to be quite general. According to Russell, the specific action of an individual cannot be directly predicted merely based on—knowledge of—the specific emotional category. Rather, because action emerges contextually, the specific characteristics of behavior are variable, based on situational factors.

Russell’s emphasis on continuous core affect is generally consistent with Izard’s (1977, 2007, 2009) notion that the human mind is continuously emotional. However, according to Izard (2009), what is continuously
present is not core affect. Rather, “a discrete emotion or pattern of interacting emotions (though not necessarily labeled or articulated) in the conscious brain” (p. 4) is continuous. Despite sharp distinctions between these two perspectives as to the nature of affective or emotional experience and its process of emergence, they converge on continuous emotionality in the human mind. It is this convergence between competing perspectives that is most informative and noteworthy for educational design.

**EMOTION AND COGNITION**

Various emotion researchers have stressed that emotions are inherently motivational and interconnected with cognitions. For example, in Izard’s (1977) *differential emotions theory* “affect and motivation are interchangeable terms that refer to all motivational phenomena—emotions, drives, and affective-cognitive structures” (Izard, 1993, p. 73). In this framework, affective-cognitive structures are units of experience and motivation which include both emotion and cognition. This notion corresponds to Lewis’s concept of *emotional interpretations* (EIs; Lewis & Douglas, 1998), which represent spontaneous and repeated coupling of specific cognitions and emotions. EIs are “appraisal-emotion amalgams” (Lewis, 2000, p. 43) as organizing and motivating patterns that contribute to order in the dynamic activity of the human mind.

Izard (2007) later revised his notion of affective-cognitive structures and proposed *emotion schemas*. An emotion schema is “the dynamic interaction of emotion and cognition” (p. 265), representing “processes involved in the dynamic interplay of emotion, appraisals, and higher order cognition” (p. 261). Emotion schemas motivate both cognition and behavior. They have “special and powerful effects on self-regulation and on perception, thought, and action” (Izard, 2009, p. 9). Thus, emotion research presents a view that: (a) emotions are ubiquitous and (b) inherently interconnected with cognition. Furthermore, (c) these connections exert powerful motivational influences. These three aspects of emotions have important implications for theoretical models of learning from multimedia and for designing instructional materials and processes. The effectiveness of instructional design will depend on the extent to which it takes into account the pervasive and motivating nature of emotions and their natural interconnectedness with cognition. These qualities are important aspects of the learning model we propose in this chapter.
Advances in emotion research and affective neuroscience have led scholars to make a strong case that emotions are formative in basic cognitive mechanisms, including memory, attention, and perception (Derryberry & Tucker, 1994; Isen, Daubman, & Nowicki, 1987; Isen, Shalker, Clark, & Karp, 1978; Izard, 1993, 2007; Lewis, 2005; Lewis, Haviland-Jones, & Barrett, 2010; Tucker, 2007). These studies supported the notion that these types of cognitive mechanisms are inherently motivated by virtue of their interconnectedness with emotion. An important aspect of this motivational mechanism is that “changes in emotional state influence higher cognition” (Rutherford & Lindell, 2011, p. 337). Emotions are seen as effective retrieval cues for long-term memory (Isen et al., 1978, 1987).

Similarly, through attention, various forms of automatic and deliberate cognitive processes “are focused and motivated by ongoing emotion that is always present in consciousness” (Izard, 1993, p. 85). Research found, for example, that positive emotions support the processing of information and communication, enhance negotiation, decision-making, creative problem-solving, and similar higher level cognitive activities (Erez & Isen, 2002; Konradt, Filip & Hoffmann, 2003). The importance of this notion for instructional design becomes clear when we consider that attention is central to learning. The process by which emotional activity contributes to changes in attention may be a key mechanism to understand and facilitate the role of emotion in learning.

EMOTIONS AND LEARNING

There has been a broad gap between the ubiquitous nature of emotions in learning (Pekrun & Stephens, 2010) and the focus of most educational research and practice, which has revolved around more cognitive issues. This gap is narrowing with the recent increase in approaches that acknowledge and examine the central role of emotions in education. This increasing recognition is part of a general trend in social sciences to view emotions “as being of critical importance for performance and the productivity of individuals, organizations, and cultures” (Pekrun & Stephens, 2010, p. 238). Still, the majority of educational research and practice is subject to the enduring effects of the old paradigm of the inherent separation of cognition and emotion. Consequently, emotions continue to be either ignored or regarded as peripheral, rather than central, to learning and teaching. Consequently, “students’ emotions continue to be underresearched” (Pekrun,
Evidence from emotion research and affective neuroscience has shown that emotion and cognition are inherently interconnected (Crick & Dodge, 1994; Derryberry & Tucker, 1994; Izard, 2009; Lewis, 2005; Tucker, 2007), leading to the inference that every information processing step of the learning process is emotional as well as cognitive. Moving emotions from the periphery to the center of educational research and practice has profound implications that can change the way we design and use instructional materials. This change will substantially increase the effectiveness of pedagogical practice and interventions, and increase the prediction of significant learning outcomes (Park, Plass, & Brünken, 2014; Pekrun et al., 2006).

The need to more fully consider the impact of emotion on learning is two-fold. First, a comprehensive scientific understanding is needed about the complexity of emotions as they occur in the real lives of individuals. Theoretical and empirical interdisciplinary advances in emotion research are promising in this direction (e.g., Izard, 2007, 2009; Lewis, 2005; Picard, 2010; Russell, 2003). Second, educational research and practice must take emotions seriously as inherently important and valuable phenomena in the learning process (Lemke, 2015). Recent evidence suggesting that instructional design can facilitate learning by fostering positive emotions is promising in this direction (Um, Plass, Hayward, & Homer, 2012). Efforts in both directions will reciprocally influence each other, as emotion research will inform and be informed by educational research and practice.

This process is facilitated by an important paradigm shift in educational technology and the learning sciences. That is, we are beginning to recognize not only that “learners have needs that are different from other kinds of users (Soloway et al., 1996),” but also that the learner as the user of technology must be viewed as a complete being (Picard & Klein, 2002, p. 142). This is a recognition that “humans are affective beings, motivated to action by a complex system of emotions, drives, needs, and environmental conditioning in addition to cognitive factors” (Picard & Klein, 2002, p. 142). By implication, the learner’s efficiency and productivity, but also their emotions, wellbeing, and motivation acquire the status of primary importance in this new paradigm.

The way students respond to academic challenges is influenced by their emotional experiences, which may impede their learning and achievement (Ruthig et al., 2008). This influence of emotions on learning interacts with the effect of perceived control (Pekrun, 2006). Even though control is a
contributor to the experience of achievement emotions, educational research and practice can benefit from taking into account differing levels of both control and emotions as separate factors in learning and achievement. Ruthig’s finding points to the need to take into account important characteristics of individual students, while designing educational interventions. If educational design is flexible enough to adapt to the learner’s emotions as well as their sense of control, it will be more likely to facilitate learning and achievement for a greater number of students.

Emotions have been shown to impact learning in authentic field settings. In one example, researchers were able to link self-reported arousal to students’ regulation of their problem-solving efforts, and self-reported valence to cognitive regulation processes, although a link between the reported emotions and learning outcomes could not be established (Linnenbrink & Pintrich, 2002a, b). The importance of emotion in the context of learning is also highlighted by issues of emotional self-regulation (including regulation of emotions and their interconnected appraisals), which is considered a key process related to “the design of ‘emotionally sound’ (Astleitner, 2000) achievement environments” (Pekrun & Stephens, 2010, p. 250). This consideration must include the psychological and environmental conditions that help students experience “adaptive levels of emotions (lower boredom, lower anxiety, or higher enjoyment)” in addition to high levels of perceived control (Ruthig et al., 2008, p. 161).

An important aspect of emotional life may be the simultaneous experience of multiple emotions (Kaplan & Tivnan, 2014c; Pekrun, 2006; Roseman, 2011). This phenomenon represents substantial intrapersonal multiplicity in emotional experience within a given time and context. Consistently, it is reasonable to expect “mixtures of emotions” (Pekrun, 2006, p. 323) in achievement and learning contexts. This multiplicity has important implications for instructional design, as it presents inherent pedagogical opportunities and challenges. The intrapersonal multiplicity of emotional experience has the potential to make the learning process more engaging and highly motivating, but also overwhelming and frustrating.

EMOTIONAL DESIGN IN DIGITAL MEDIA FOR LEARNING

The emotions that learners experience in digital learning environments may not be different from those in other types of learning environments. But digital learning environments offer many more ways of influencing learners’ emotions, using a number of design features that are under the control of
We use the term emotional design to describe the use of a range of design features with the goal to impact learners’ emotions to enhance learning. Some of these design features relate to the way information is presented, and others to the way the interactions in the environment are structured (Plass & Schwartz, 2014).

Emotional Design Through Information Representation

For information representation, the visual design of the learning materials themselves can impact emotions that in turn impact learning. There is a general notion that physically attractive stimuli have a positive impact on learning—“what is beautiful is good” (Dion, Berscheid, & Walster, 1972). In fact, research has shown that children associate brighter colors with more positive emotions, and darker colors with more negative emotions (Boyatzis & Varghese, 1994). Several studies on multimedia learning have implied that different aesthetic designs can induce emotions and that these emotions affect users’ performance and cognitive processes (Harp & Mayer, 1997; Mayer & Moreno, 1998; North & Hargreaves, 1999; Szabo & Kanuka, 1998; Tractinsky, Katz, & Ikar, 2000; Wolfson & Case, 2000). Other researchers found that the design of various multimedia elements, such as the visual design, design layout, color, and sound in multimedia environments, resulted in positive user perceptions about learning (Tractinsky et al., 2000; Wolfson & Case, 2000).

Another established effect is the baby-face bias, which describes how people or things with round features and large eyes are perceived as baby-like (Lorenz & Generale, 1950). Unlike shapes featuring sharp edges, these round features induce a positive affect in the learner by evoking baby-like personality attributes—innocence, honesty, and helplessness. Anthropomorphism research, which studies the attribution of uniquely human characteristics and qualities to nonhuman beings, inanimate objects, or natural or other phenomena, has reported similar effects (Dehn & Van Mulken, 2000; Disalvo & Gemperle, 2007; Hongpaisanwiwat & Lewis, 2003; Reeves & Nass, 1996).

Based on this research on shapes and colors, we have, in our own research, investigated whether these visual emotional design elements can, in fact, impact learning. Our research with college students provided evidence that the use of round shapes and warm colors in the visual design of learning environments is able to induce positive emotions in learners that in turn facilitate comprehension and transfer of learning scientific materials.
The design of that study used a combination of two visual design elements: color and shape, to manipulate the emotional impact of the materials without adding new information. Therefore, in a follow-up study, we investigated to what extent each of these design elements separately contributed to the impact on emotion and learning (Plass, Heidig, Hayward, Homer & Um, 2014). This study was able to replicate the effect of emotional design on comprehension we found in our 2012 study, but not the effect for transfer. In addition, this follow-up research revealed that round face-like shapes induced positive emotions both alone and in conjunction with warm colors. Interestingly, we found that warm colors alone did not affect learners’ emotions. Comprehension was facilitated by warm colors, by round face-like shapes, and by combinations of both design features. Transfer, on the other hand, was facilitated by round face-like shapes when used with neutral colors.

Another interesting result of this study was that the mood induction procedures we had used, which consisted of watching funny cartoon movies, resulted in elevated feelings of excitement, enthusiasm, determination, and attentiveness, relative to those who did not view the cartoon; however, these elevations were not sustained over the course of the learning session. In contrast, the visual emotional design features resulted in elevated feelings of being inspired, interested, and enthusiastic, and were sustained during the learning session (Plass et al., 2014). We have since been exploring other promising ways of emotional design via information representation, such as through game characters, for which we found that their visual design (color and shape), but also their movement and associated sounds impacted learners’ emotion, as can be seen in Figure 7.1 (Biles, Szczuka, Plass, & Krämer, 2014; Szczuka, Biles, Plass, & Krämer, 2013).

**Emotional Design Through Interaction Design**

For learning interactions, a number of different ways have been explored that impact learners’ emotions. They can be based on the INTERACT model that describes how learners’ actions include behavioral, cognitive, and affective activity, and how these kinds of activities affect one another during learning, see Figure 7.2 (Domagk, Schwartz, & Plass, 2010). The effect of interactivity on emotion has been investigated using a number of different approaches. Below we will discuss two of them, namely a focus on the situational interest that the activity generates, and the use of guided activity, such as involving animated pedagogical agents.
Situational interest is described as an immediate affective response to particular stimuli and conditions that originate from the learning environment, a response that may be fleeting or lasting, and that directs learners’ attention to the task (Hidi, 1990; Hidi & Renninger, 2006; Mitchell, 1993; Rotgans & Schmidt, 2011; Schraw, Flowerday, & Lehman, 2001). Situational interest is

**Figure 7.1** Emotional design via game characters’ color, shape, movement, and sound.

**Figure 7.2** INTERACT model describing cognitive, behavioral, and emotional activity in multimedia learning.

**Situational Interest**

Situational interest is described as an immediate affective response to particular stimuli and conditions that originate from the learning environment, a response that may be fleeting or lasting, and that directs learners’ attention to the task (Hidi, 1990; Hidi & Renninger, 2006; Mitchell, 1993; Rotgans & Schmidt, 2011; Schraw, Flowerday, & Lehman, 2001). Situational interest is
different from learners’ individual interest, which describes their intrinsic desire and tendency to engage in a particular subject matter or activity over time. Situational interest is of importance as research has found that it is essential in the development of individual interest (Hidi & Renninger, 2006).

We have found evidence that a number of different design elements in interactive learning environments, such as games for learning, can impact the situational interest experienced by the learner. Among them are the game mechanics, the social mode of play, and the use of badges. In one study, we compared two versions of a geometry puzzle game for middle-school students, Noobs vs. Leets. The only difference between these versions was the game mechanics, i.e., the essential game play that the game afforded. In one version, players were asked to solve geometry problems of angles in quadrilaterals by computing a missing angle (numeric condition); in the other, they were asked to select the rule that needed to be applied to solve the problem (rule condition). Results showed that the numeric condition, based on the situation, was more interesting than the rule condition, suggesting that the selection of the game mechanic has an impact on learners’ affect (Plass et al., 2012).

We also compared three versions of a game on factoring for middle-school students that facilitated either individual play, competitive play of two players, or collaborative play of two players. Results of this study showed that competition and collaboration elicited greater situational interest than the individual play, suggesting that the social mode of play is able to impact affect (Plass, O’Keefe, et al., 2013). Finally, we designed a version of the Noobs vs. Leets game that awarded the learner different types of digital badges for the completion of in-game learning-related tasks. We found that the design of the badges impacted learners’ situational interest, in addition to learning outcomes (Biles & Plass, in press; Plass, O’Keefe, Biles, Frye, & Homer, 2014).

**Guided Activity Principle—Animated Pedagogical Agents**

According to the *guided activity principle* for interactive multimodal learning environments, “students learn better when they interact with a pedagogical agent who guides their cognitive processing” (Moreno & Mayer, 2007, p. 315). Animated pedagogical agents play instructional roles to support sociocognitive aspects of multimedia learning; can follow social conventions; and provide empathetic responses to learners (Hayes-Roth & Doyle, 1998). These agents are represented visually, often with human features, and with optional auditory (speech) features. Considering the inherent interconnectedness of cognition and emotion (Lewis, 2005; Tucker, 2007) and applying insights from *affective computing* (Picard, 2003), a corollary to this principle can be derived for
emotional design. That is, students are likely to learn better when interacting with a pedagogical agent who facilitates the learner’s emotional self-regulation.

A skillful tutor adjusts not only her emotional expressions (such as tone of voice), but also the content of the instruction according to her perception of dynamic emotional changes in her student. Is this a realistic expectation for a virtual pedagogical agent? The existence of systems that respond to changes in aspects of emotional experience suggests that advancements in technology and scientific understanding could fulfill this expectation. There are promising developments in the field of affective computing (Calvo, 2010; Hudlicka, 2003; Picard, 2003; Picard & Klein, 2002; Picard, Vyzas, & Healey, 2001), which suggest that emotionally intelligent human computer interactions (HCI) are not only possible, but also necessary for increasing the wellbeing of users as well as efficiency.

Research has shown that animated pedagogical agents can increase motivation (Moreno, Mayer, Spires, & Lester, 2001), and result in higher interest and lower perceived difficulty of learning materials (Mitrovic & Suraweera, 2000). When such systems induce confusion and engagement (flow), they have positive correlations with learning, but when they induce boredom, the correlation is negative (Craig, Graesser, Sullins, & Gholson, 2004; D’Mello & Graesser, 2014). Learning materials induce a state of cognitive disequilibrium in learners, which can be a result of obstacles to their goals, interruptions, anomalies, and the like (D’Mello & Graesser, 2012; Graesser & D’Mello, 2011). In such cases, there are alternative consequences. Learners may either experience negative emotions and disengage or, when they manage to overcome the source of the disequilibrium, experience positive emotion and higher engagement (Baker, D’Mello, Rodrigo and Graesser, 2010; D’Mello & Graesser, 2012; Sabourin, Mott, & Lester, 2011). Studies in which learners’ emotions were continuously monitored, and in which the agent responded to the learners’ emotional state through empathy, support, or even cheeky comments, found mixed results on learning outcomes (D’Mello, Craig, Fike, & Graesser, 2009; D’Mello & Graesser, 2014). This highlights how difficult it is to determine the best response to the detected emotions.

THE THEORETICAL FOUNDATION OF EMOTIONS AND LEARNING

The research we have discussed so far shows the importance of an integrated model of emotion and cognition for designing and studying multimedia learning, and we now propose such a model, which focuses on emotional
design for multimedia learning. Among the theories that inform our work on emotional design are theories on achievement emotions (Pekrun, 2000); theories incorporating affective elements into multimedia learning (Moreno & Mayer, 2007); and theories of affective computing (Picard, 1997). We will summarize these theories before discussing our approach to integrating them into an Integrated Cognitive-Affective Model of Multimedia Learning.

**Pekrun’s (2000) Control Value Theory of Achievement Emotions**

The Control Value Theory of Achievement Emotions is an integrative framework that describes the antecedents and effects of emotions experienced by learners, i.e., in academic settings or achievement situations (Pekrun, 2006; Pekrun & Stephens, 2010). This theory is relevant for instructional design, as it shows that learning can be facilitated through positive achievement emotions, such as enjoyment. This is possible to the extent that instructional design evokes and fosters appraisals of control and positive value for the task and object of learning. These emotions are in turn likely to increase “interest and motivation to learn” (Pekrun, 2006, p. 326), and facilitate self-regulation and performance in the learning process (Pekrun & Stephens, 2010). Consistently, positive emotions are likely to facilitate internalized motivation (Pekrun, 2006), which, according to the Self-Determination Theory (SDT; Deci & Ryan, 1985; Ryan & Deci, 2000), is operationalized as increased autonomy. This insight can be considered together with the notion that increased autonomy and internal sense of control facilitate positive emotions. It follows that there is a dynamic and reciprocal relationship between positive achievement emotions and autonomy, as they influence each other over time in a positive feedback loop.

Similarly, because emotions and appraisals influence each other over time (Pekrun, 2006), their relationships may be better characterized in terms of a dynamic feedback loop (Lewis, 1995, 2002), rather than unidirectional causation. Learning may be shaped by a similarly dynamic process of interconnected components: “Emotions are assumed to affect learning and achievement, but success at learning influences students’ appraisals and emotions. By implication, emotions, their individual and social antecedents, and their effects are linked by reciprocal causation over time” (Pekrun, 2006, p. 327).

What are some of the important ways by which emotional design can promote competence and positive value for learning tasks and materials? One of the key factors is for design features to support the learner’s sense of autonomy: “the individual has to learn how to adapt to situational demands while
preserving individual autonomy—inevitably a process guided by appraisals” (Pekrun & Stephens, 2010, p. 241). In this context, supporting autonomy requires design characteristics that make it clear to the learners that important aspects of both the process of engaging in a learning task, and its outcomes, are relatively under their control, given sufficient effort.

Another key phenomenon is the intrinsic value of the learning activity (Pekrun, 2006). Intrinsic value represents a positive motivation for engaging in the learning activity for its own sake. From a control–value theory perspective, enjoyment can be predicted to be the strongest when high intrinsic value is combined with an appraisal that the learning activity is sufficiently controllable (Pekrun, 2006, p. 323) by the learner. By establishing this combination of control and value, emotional design can facilitate enjoyment as a positive activity emotion, and reduce the likelihood of negative activity emotions, such as anger and frustration.

A related mechanism by which instructional design can facilitate positive emotions is by evoking and reinforcing specific types of goals. Considering evidence that mastery goals are positively associated with enjoyment, hope, and pride (Pekrun et al., 2006), design features that facilitate these goals can be expected to promote positive emotions in the learning process. By contrast, design features that evoke and reinforce performance-avoidance goals will promote negative emotions, such as anxiety, hopelessness, and shame (Pekrun et al., 2006).

An important implication of Pekrun’s framework is that emotional experience varies (intrapersonally and interpersonally), based on object focus. That is, an individual’s emotional experience will be different depending on whether appraisals of activity or outcome trigger emotions. Pekrun further differentiates outcomes as retrospective and prospective, to be associated with different emotions.

Furthermore, to the extent that people can experience multiple emotions (Roseman, 2011) simultaneously or successively in a short time, emotional experience can be affected by both activity and outcome appraisals. People can have various degrees of activity-focused emotions, such as enjoyment, relaxation, frustration, or boredom, combined with outcome-focused emotions, such as hope, relief, anxiety, hopelessness, pride, sadness, and disappointment. According to Pekrun, anger can be triggered both by activity and outcome-related appraisals. Emotionally intelligent designs of HCI must take into account the task characteristics and possible individual characteristics that are likely to trigger activity and outcome appraisals, which evoke multiple emotions. Considering that negative emotions such as anxiety,
boredom, and hopelessness reduce performance (Pekrun, 2006; Pekrun & Stephens, 2010), an important function of emotional design is to minimize the likelihood of such emotions and promote self-regulation when they appear. What are some of the key factors of emotional design, which could make a difference toward this aim?

Pekrun and Stephens (2010) presented a list of important qualities about instructional design that are likely to facilitate performance and learning through their positive impact on achievement emotions: (a) cognitive quality of instruction; (b) motivational quality of instruction; (c) autonomy support; (d) goal structures and expectations; (e) feedback and consequences of achievement. Among these factors, motivational quality of instruction and autonomy support are most directly related to emotional design. From a Self-Determination Theory perspective, these two factors are interconnected because autonomy support improves the quality of motivation. According to Pekrun and Stephens (2010), the motivational quality of instructional design can be enhanced to the extent that “supervisors, teachers, and peers deliver both direct and indirect messages conveying achievement values” (Pekrun & Stephens, 2010, p. 245). For example, from an SDT view, the design of learning tasks and environments can support the fulfillment of basic psychological needs of relatedness, autonomy, and competence.

**Moreno and Mayer’s (2007) Cognitive Affective Theory of Learning with Media**

Moreno proposed an addition to Mayer’s (2005) Cognitive Theory of Multimedia Learning (CTML) into which she incorporated motivational as well as metacognitive factors that mediate multimedia learning (Moreno, 2007; Moreno & Mayer, 2007). In learning with multimedia, learners first select relevant information and build verbal and visual mental representations of what was presented. They then organize this information in working memory, connect the verbal and visual representations with one another, and integrate them with prior knowledge. Because working memory can only hold a limited amount of information (Baddeley, 1986; Cowan, 2001), processing of multimedia information is performed under these memory constraints. Multiple factors impact the amount of cognitive load a learner experiences, such as the difficulty of the materials and the demands of the learning task (Plass, Moreno, & Brünken, 2010).

Moreno’s additions to CTML recognize that motivation as well as metacognition mediate the cognitive processing of multimedia information. By doing so, she expands the CTML to include the first noncognitive element.
Both motivation and metacognition have a large body of research in support of their impact on learning, and Moreno added them as supplemental factors to a system that has cognitive processing at its core.

**Picard’s (1997) Affective Computing**

Affective computing describes computational approaches to the detection and deliberate induction of affect (Picard, 2003). Applying the emphasis of Picard and Klein (2002) on the importance of emotional needs, an important function and purpose of emotional design emerges. That is, emotional design in the learning process can contribute to the wellbeing of the learners by taking into account and helping to fulfill the learners’ emotional needs. According to Picard and Klein (2002), these needs include both *emotional skill needs*, such as empathy and self-awareness, and *emotional experiential needs*, such as feeling connected to others and understood and accepted by them. Emotional design of instructional processes can help learners fulfill both types of needs.

For example, by using physiological measures, such as face recognition, heart-rate variability, and skin conductance, a computer can recognize key emotional states and related changes in the user. The communication of such information to the user can help increase emotional skill needs by contributing to self-awareness and experiential needs by helping “the user to feel as if his or her strong affective state has been effectively communicated” (Picard & Klein, 2002, p. 145). A key factor that Picard and Klein (2002) emphasize toward meeting emotional needs is “to consider the closest human analogy to the human-computer interaction being designed, and ask how that interaction might make a person feel” (pp. 150-151).

This consideration implies that human-to-human interaction and human-to-computer interaction (HCI) share fundamental features about the nature and importance of human emotions. This idea may be of utmost importance for instructional design that is centered on computer technology. Instructional designers do not have to see an inherent split and dichotomy between these two contexts of interaction. Rather, instructional design may benefit from understanding their common features.

As Picard and Klein suggest, addressing the user’s emotional needs does not require the computer to have the emotional qualities of a human being. The emotional needs of the human user can still be addressed and at least partially fulfilled during the interaction with the program, even when the user is aware that the agent with whom she is interacting is not a human being.
Significant developments in this direction can be seen as functions of “the recent shift from computers as tools to computers as partners and socially intelligent agents, and intelligent decision-aids” (Hudlicka, 2003, p. 21). Consistently, the field of affective computing is flourishing, with many projects toward “building machines that have several affective abilities, especially: recognizing, expressing, modeling, communicating, and responding to emotion” (Picard, 2003, p. 56).

As affective computing can serve toward “minimizing user frustration and maximizing user satisfaction” (Picard & Klein, 2002, p. 167), so can emotional design in education substantially improve learning and wellbeing. Similar to the role of affective computing, emotional design in education can utilize and facilitate the process by which “emotions aid in intelligent interaction and decision making” (Picard, 2003, p. 63). Therefore emotional design can make learning and teaching more compatible with the actual emotionally grounded process of how the human mind works and changes. As a result, through emotional design, digital learning environments can serve the needs of the whole person.

**TOWARD AN INTEGRATED COGNITIVE-AFFECTIVE MODEL OF MULTIMEDIA LEARNING**

To the extent that HCI can be designed to address user affect (Hudlicka, 2003), the principles of affective computing can be applied to education in order for computer-based instructional design to fulfill important educational needs. In particular, if computer systems can recognize the user’s affect and respond accordingly, instructional materials using these systems can address emotions effectively in the learning process. Increasing the accuracy of affect recognition is a worthwhile and important endeavor. Upon recognition, whether or not, and how and when the system will respond to the learner’s affect are critical decisions. Instructional designers must make such decisions depending on the specific characteristics of a given task and the aims of the particular application. Accumulation of empirical findings about the dynamic, temporal relationships between learner’s affect, learner’s motivation, and various forms of system response will be highly valuable for more informed and effective design decisions.

Given the knowledge of the user’s affective state and its likely effects and the user’s desired state for the objectives of the HCI, we must decide whether or not the system should respond to this state, and how or what
affective behavior, if any, the system should display to induce a desired user state or behavior (Hudlicka, 2003, p. 8).

Addressing a learner’s affect can be an important aspect of emotional design, enabling scholars and educators to both understand the complex role of emotions in learning and facilitate this role in order to improve the efficiency and effectiveness of the learning process. Furthermore, just as the benefits of affective computing increased efficiency and productivity, this approach may also improve the motivation and wellbeing of learners.

These outcomes will be initiated as emotional design reduces the negative emotional states, such as frustration, boredom, stress, and anxiety, and increases the positive emotions, such as hope and enjoyment, and indeed, facilitates self-regulation of all possible (negative and positive) emotional experiences throughout the learning process. It is important for this general vision to be adjusted based on the specific demands and characteristics of each learning application (e.g., a particular educational game or simulation). For greater benefit, this contextualization can be extended to take into account interpersonal variability, namely, the variable emotional and motivational needs and demands of individual learners. For example, moderate and manageable levels of stress and anxiety can be adaptive as part of the motivation process of certain individuals in some learning tasks. In this context, Hudlicka’s (2003) characterization of the “broad aim of affective HCI” directly applies to and informs emotional design in instructional technology:

Regardless of whether or not affective factors will ultimately be considered in a particular human–machine context, it is critical that the system designers accurately assess the range of possible affective states the users may, or should, experience during interactions with the system, and that they understand their effects on the user, and thus on task performance. Such understanding then allows informed decisions regarding which affective considerations must be addressed, when and how. (p. 7)

The necessary shift to consider the user or the learner as a complete human being with affective needs (Picard & Klein, 2002) is accompanied by a parallel shift in how we view the computer and the individual interacting with it. That is, there is a tendency to characterize the user and the computer together as a whole unit in terms of “collaborative systems, integrated human–machine systems,” reflecting “a deep and significant shift in underlying design philosophy and objectives, and indeed in the expectations we now have of computer systems” (Hudlicka, 2003, p. 2). This holistic view resonates with the developmental vision of Kurt Lewin (1946), an early pioneer of the dynamic systems (DS) perspective (for a discussion, see Thelen &
Smith, 1994). According to Lewin (1946), human behavior and development must be viewed “as a function of the total situation” (p. 791), in which the person and the context constitute a unified whole.

A direct implication of this dynamic developmental view is the characterization of human mind and behavior in terms of tendencies or probabilities instead of rigid rules or fixed properties. To the extent that learning is a developmental process, it involves “the emergence and consolidation of new possibilities and tendencies for behavior to coalesce in real time” (Lewis & Douglas, 1998, p. 160). The emerging mental and behavioral tendencies are various strategies and motivations that are both cognitive and emotional at the same time (Lewis, 2005). Consistently, according to Lewis and Douglas, human behavior and development can be explained through emotional interpretations (EIs), as both products and building blocks of mental and behavioral organization. As spontaneously and repeatedly occurring patterns of interactions between cognition and emotions, EIs inform and guide decisions and actions.

Relatedly, Izard (2007, 2009) proposed emotion schemas as similarly dynamic structures with a motivating role in judgment and behavior. Therefore, in order to understand and facilitate the learner’s actions and decisions in a learning environment, it is useful to consider the possible combinations of emotions and cognitions that serve as dynamic motivating forces. Because of the soft-assembly of the human mind (Kaplan & Tivnan, 2014a,b; Kloos & Van Orden, 2009), emotion–cognition interactions will occur in ways that are context-specific. Specific features of instructional design may evoke and support specific combinations of emotions and cognitions, which in turn will influence the learner’s interactions with the instructional materials. This view implies that the learner’s intrapsychic motivational dynamics (ongoing combinations of emotions and cognitions) and the features of the learning context are interconnected. In summary, (a) there is a continuous and inherent interconnectedness between emotion and cognition, and (b) dynamic cognition–emotion interactions (IEs or emotion schemas) emerge and operate in ways that are highly contextualized (and hence sensitive to contextual factors). Furthermore, (c) dynamic cognition–emotion interactions serve as motivating forces that guide human adaptation and learning in specific contexts. These three insights, coming from both emotion research and the DS perspective, contribute to, and inform, the new learning model we propose.

Figure 7.3 shows the Integrated Cognitive Affective Model of Learning with the Multimedia (ICALM) model we propose for the purpose of emotional design in multimedia learning. The system incorporates Moreno and
Mayer’s (2007) multimedia learning via selecting, organizing, and integrating visual and verbal information, and combines these processes with Russell’s (2003) notions of core affect and attributed affect, and Izard’s (2009) concept of emotion schemas. The main thesis of this model is that affective processes are intertwined with, and inseparable from, cognitive processes, and that the cognitive-affective processing of multimedia stimuli involves affective processes that make demands on cognitive resources, and vice-versa.

The multimedia environment induces affective responses, which we describe as “core affect.” This core affect is experienced as the learner perceives auditory and visual information from the environment. Some of the experienced emotions may be attributed to specific sources, but they may also persist unattributed, as mood. This attribution is impacted by the information learners select from what is presented, but also impacts that selection process. As learners organize visual and verbal mental representations in working memory, affect that involves appraisal is experienced by the learner as interest and motivation, impacting the organization of these mental representations, but also being impacted by what is being processed. The integration of the different mental representations, which traditionally only involve verbal and visual mental representations, also includes the experienced affect, forming emotion schemas that are stored as long-term memory. Depending on the type of information the learner processed, this can either be semantic or episodic memory.

Research summarized by Mayer and Moreno (2003) suggests that effective instructional design must minimize cognitive load that arises from non-essential processing, i.e., processing that does not serve learning outcomes. Similarly, with this model we propose that effective instructional design
must evoke and observe efficient emotional experience. This includes design that is compatible with individual’s emotional self-regulation skills. Emotional overload will occur, for example, when the task evokes emotions that exceed the user’s capacity for emotional self-regulation. Effective instructional design minimizes emotional overload. Furthermore, emotions that lead the user’s attention away from learning objects or activities will not be beneficial.

An example for the deleterious impact of emotion on learning is the stereotype threat (Steele & Aronson, 1995). As research has shown, the stigmatizing of specific individuals, such as African-Americans (Steele & Aronson, 1995), women (Spencer, Steele, & Quinn, 1999), or White males (Aronson et al., 1999), results in emotional responses that interfere with cognitive processing, consume executive resources, and, as a result, negatively impact learning and performance (Schmader, Johns, & Forbes, 2008).

Learning is also compromised when the individual’s diversity and intensity of emotional experience are below optimal levels. In such cases, particularly when positive achievement emotions, such as enjoyment and hope (Pekrun, 2006; Pekrun & Stephens, 2010) are not experienced strongly enough, the learner may lack the motivation and engagement for optimal performance and learning. Such a lack of motivation is likely to reduce creativity, and prevent the learning process from fulfilling its developmental potential for the individual. For example, when learners lack motivation, they may fail to engage in generative processing, even when cognitive capacity is available (Moreno & Mayer, 2007, p. 315). According to Moreno and Mayer, generative processing involves constructing a new understanding, which may integrate new information with existing knowledge.

This model makes several assumptions that are based on the literature described above, but also leads to research questions that should be investigated empirically. We describe some of these questions in the following section.

**RESEARCH AGENDA FOR THE STUDY OF EMOTIONAL DESIGN**

On the most fundamental level, research on emotional design is concerned with questions, such as: Which design elements of learning environments impact emotion? How do emotions impact learning outcomes? How can emotions be measured? What are the appropriate responses to the detected emotional state of the learner?
Research on Design Factors Impacting Learners’ Emotion

Research has identified several design elements in digital media that impact learning. However, more research is needed to systematically document which emotions are induced by each of these elements. Based on our earlier differentiation, the design to be investigated should be related to information representation and interaction design. Information representation, for which we have begun to investigate factors, such as color, shape, and movement, should focus on interactions of these factors as well as on additional formats of information representation, such as sound. Interaction design has primarily been investigated via factors related to situational interest generated by the mode of play, mechanics design, personalization, and badges design, and by animated pedagogical agents. Research on the impact of interactions on emotion should focus on additional factors, such as narrative, and on the use of personalized messages to the learner. The latter has been shown to impact learning (Moreno & Mayer, 2000), but this effect has not yet been clearly linked to emotion.

The study of emotional design in specific contexts or for specific media raises additional questions. For example, for the design of video games, it is of interest that Russell’s model (2003) describes emotion episodes as phenomenologically including and transcending core affect. One of the key distinctions between these two phenomena is that an emotional episode involves behavior directed at an object. This distinction is particularly relevant to people’s experience with video games or virtual reality. Russell stresses that while virtual reality (e.g., a scary and threatening movie character) evokes core affect (e.g., fear) and corresponding physiological changes, it does not trigger the purposeful behavior (e.g., flight) associated with an emotional episode. Thus, according to Russell, experience with virtual reality usually involves core affect of a particular feeling, such as fear or sadness, and not its “full-blown emotional episode” (p. 155).

On the other hand, this behavioral criterion for the distinction between core affect and emotion episode becomes controversial, considering that many interactive video games simulate behavior. As virtual reality, a video game shares a common ground with Russell’s example of a film, in that the individual knows that the stimulus is imaginary. However, playing a video game is very different from watching a film because of the interactive experience. Unlike a movie, a video game enables the user to simulate behaviors. Is this a sufficient condition for games to be evoking emotional episodes beyond core affect? In what significant ways do the control and
manipulation of a game character’s behaviors contribute to the player’s emotional experience in a game? The exploration of such questions may be useful in order to explicate the nature of emotional experience in digital games and simulations in ways that could inform emotional design.

Research on the Impact of Emotion on Learning Outcomes

A number of questions need to be investigated in order to validate specific elements of the model we proposed. In particular, the interconnectedness and interplay of emotion and learning provides a number of challenges that research should address. One of these important questions for research on emotions and learning relates to the insight that emotion and learning do not have a linear and uniform relationship. In order to explain real-life complexity, new models and studies must take into account nonlinear relationships between emotions and learning, including the uniqueness of individual developmental pathways. For example, as Pekrun (2006) emphasized, “positive emotions are not always adaptive, and negative emotions not always maladaptive” (p. 327). Idiosyncratic factors, such as the individual’s psychophysiological arousal response, self-regulation skills and related developmental history may moderate the learning effects of various emotions. For example, based on their self-regulation skills, some learners may use moderate degrees of negative emotions, such as anxiety and fear as motivators for high achievement; whereas smaller degrees of such emotions can be detrimental for other individuals who have low levels of emotional self-regulation.

Similarly, van Geert (2002) proposed that “the causes of human actions, performance, skills, and knowledge lie in the process of a temporal interplay between” (p. 320) the properties of three factors that are inherently interconnected: (a) the transcendental object(s); (b) the subject (e.g., a student); and (c) the context (e.g., the immediate educational context in which specific learning occurs). Applying this model to the learning process, transcendental objects are the specific learning materials as well as the tools of instructional design (such as a virtual pedagogical agent). The properties of such design materials and tools (e.g., taking into account changes in the learner’s motivation and emotions) enable the learner “to intentionally relate to the object (in the broad sense of the word) in question” (p. 320). According to van Geert (2002), there are two key principles that characterize the relationships between these three sources of experience and understanding. First, the boundaries differentiating these three sources are fuzzy and dynamic,
rather than absolute, exclusive, and static. “Because they come about as a result of the time-governed interplay among the factors, it is impossible—not only in practice but also in principle—to draw a sharp line between the factors or to specify their properties in a ‘crisp’ way” (pp. 320-321).

Research should explore, for example, how in an educational game, the developing cognitive and emotional skills of the learner are dynamically related to motivational properties of the specific tasks that are to be completed. It is neither necessary nor realistic to define the learning skills in a way that is completely independent of and distinct from the motivational properties of the game. This example also illustrates van Geert’s (2002) second principle: “if we measure a person’s psychological properties, we must, by necessity, also invoke the objects—again in the broadest possible sense of the word—and the contexts in which these psychological properties make sense, so to speak, and which are characteristic of the person in question” (p. 321). By implication, our definitions and operationalizations of learning skills, learning tasks, and design features must take into account and reflect the inherent mutuality (van Geert, 2002, p. 321) between the developing characteristics of learner, learning objects, and the learning context. These principles are useful for designing effective educational materials because they could help researchers and educators become aware of and specify the dynamic relationships between the subjects (learners), objects (what is to be learned), and the tools (e.g., computers and simulations) of the dynamic learning process.

**Research on Measuring Emotions**

There is evidence supporting the notion that key emotional states, such as anger, hate, grief, and joy, can be reliably differentiated and recognized, based on physiological measures, such as facial muscle tension, skin conductance, blood pressure volume, and respiration (Picard et al., 2001). This research should be extended to cover more emotions that are important in learning contexts, including curiosity, frustration, pride, anxiety, hope, and boredom (Pekrun, 2006; Picard et al., 2001). Identifying interpersonal commonalities and variability in the biological signals associated with key learning emotions can significantly contribute to effective emotional design of computer-based instructional materials.

The availability of several consumer grade tools to measure emotion offers promising examples in this direction. Some of these systems measure heart-rate variability (HRV) and heart-rhythm coherence. When heart-rate
exhibits orderly variability, which is neither too low nor too high, then heart-rhythm coherence is high. As an aspect of wellbeing, coherence has been found to be particularly high under positive emotions, such as gratitude, appreciation, and compassion. Increased coherence is an indication of biopsychological resilience through which the organism can respond to mental and emotional demands flexibly without burning out. Throughout seconds and minutes, coherence can increase to the extent that HRV approximates an optimal range.

Systems, such as Heartmath Institute’s emWave® Desktop system immediately respond to psychophysiological (including emotional) changes in the user, to the extent that such changes affect coherence. This biofeedback can in turn be taken into account by the user to alter his mental state in a desired direction. Consistently, appropriate use of mental imagery has been documented to have an immediate impact on coherence (Kaplan & Epstein, 2011). Because mental images are closely connected to and can easily evoke emotions (Kaplan, Epstein, & Sullivan Smith, 2014–2015), mental imagery is an effective process that can facilitate emotional self-regulation. In this research program, vivid mental images (involving a blue-golden light surrounding one’s heart) were associated with a positive emotional experience and rapid increases in wellbeing, as measured quantitatively (Kaplan & Epstein, 2011), and reflected in participants’ verbal self-reports (Kaplan et al., 2014–2015). This intimate and natural connection between mental imagery, emotions and wellbeing is an important phenomenon to consider while designing instructional materials that involve digital media presentations of various images.

Research on Appropriate Responses to a Learner’s Emotional State

As methods to measure emotion improve to a point where a learner’s emotional state can be detected at any time, the question arises how the system should respond to the emotions experienced by the learner.

Both learning and emotional experiences are dynamic processes, which unfold through changes over time. Therefore, emotionally intelligent instructional resources and tools must be dynamic rather than static. It is not sufficient for instructional resources to reflect what may be seen as general preferences for pleasant perceptions or mainly design to evoke positive affect. Rather, emotionally intelligent instructional resources require systems that are responsive to ongoing and changing emotional experiences
of each individual user throughout the learning process. Designing instructional resources with such emotionally intelligent virtual pedagogical agents is the direction that will fulfill the significant potential of emotional design for improving learning. Such a dynamic approach will meet an emerging need in emotion research and affective computing to address “individual patterns of emotions, not just group differences” (Picard, 2010, p. 250). To this aim, educational research, practice, and design must take into account the idiosyncratic nature of individual emotional experience as it changes throughout the learning process. For example, animated pedagogical agents with affective sensing can detect confusion, boredom, or happiness, but what is the most appropriate response to this emotional state, especially since it is expected to be changing over time? How should responses vary by learner, and what learner variables need to be taken into consideration? Can agents be designed that learn how students react to a particular emotional response by the agent, and can adapt their responses accordingly?

From a dynamic systems perspective, substantial intrapersonal and interpersonal variability is expected in the roles of emotions in learning, as well as the intensities of various emotions as inherent aspect of the learning process. Similarly, substantial contextual variability in emotions is expected, based on the unique characteristics of specific learning tasks and environments (Hudlicka, 2003). Therefore, individualization and contextualization of instructional design are important, as applications must be flexible enough to adapt to individual circumstances of specific learners and tasks. Computers can be very useful tools for the individualization and contextualization of instructional materials. Such an individual-based and context-sensitive approach can be combined with research-based information regarding “generic effects of emotional states on different processes involved in attention, perception, cognition, and motor performance” (Hudlicka, 2003, p. 17).

**CONCLUSION**

In this chapter, we reviewed basic concepts related to emotion and learning, and summarized research on emotional design in digital media for learning. We then presented a theoretical framework of learning from digital media that highlights the intertwined and interconnected nature of emotion and learning. This framework shows the importance of considering emotional design factors in addition to cognitive design factors when designing multimedia learning materials. We concluded the chapter by developing a
research agenda for the study of emotional design for multimedia learning. It is our hope that this chapter, as well as the other chapters in this volume, will influence the discourse both in academic and professional learning design communities to add considerations about learners’ affect to the cognitive and sociocultural considerations that are still currently dominant.

REFERENCES


Emotional design in digital media


CHAPTER 8

What Sunshine Is to Flowers:
A Literature Review on the Use of Emoticons to Support Online Learning

Joanna C. Dunlap\textsuperscript{a}, Devshikha Bose\textsuperscript{b}, Patrick R. Lowenthal\textsuperscript{b}, Cindy S. York\textsuperscript{c}, Michael Atkinson\textsuperscript{b}, Jim Murtagh\textsuperscript{b}

\textsuperscript{a}University of Colorado Denver, Denver, Colorado, USA
\textsuperscript{b}Boise State University, Boise, Idaho, USA
\textsuperscript{c}Northern Illinois University, DeKalb, Illinois, USA

INTRODUCTION

Finally, thank you for using emoticons. You are the first that I have seen using them in the program. Others have discouraged and basically banned the use of emoticons because they were considered ‘unprofessional.’ Though I understand to some degree, I also believe emoticons save a lot of grief and help to establish community. That’s just my 2 bits. So, thank you! Graduate student (personal communication, January 24, 2010)

What sunshine is to flowers, smiles are to humanity. These are but trifles, to be sure; but scattered along life’s pathway, the good they do is inconceivable. Joseph Addison (1672-1719)

Learning happens, whether face-to-face or online, within a social context. But this social context is very different in online courses. In online courses, communication (and thus the social context) is largely dependent upon asynchronous and synchronous electronically mediated communication (EMC). And despite some of the affordances of synchronous EMC, text-based EMC still remains the predominant way teachers and students communicate in online courses. Text-based EMC is popular largely because it supports the anytime, anywhere “promise” of online education. At the same time however, text-based EMC has received decades of criticism for being too lean. In the 1980s and 1990s, researchers studying computer-mediated communication
CMC—which at that time focused primarily on e-mail—came to the conclusion that CMC was inherently antisocial and impersonal (Walther, 1996; Walther, Anderson, & Park, 1994; Walther & Parks, 2002). Many of these researchers used social presence theory, developed by Short, Williams, and Christie (1976), to explain the limitations of CMC (Walther & Parks, 2002). Short et al. (1976) argued in the 1970s that communication media differ in their degree of social presence and that these differences influence how people interact, communicate, and perceive others as being “there” and “real.” However, as educators started using CMC for educational purposes, they realized that even though nonverbal and relational cues were filtered out, CMC could be very social and interpersonal (Gunawardena, 1995; Gunawardena & Zittle, 1997). This observation led researchers of online education to reconceptualize social presence theory, focusing less on communication media and its constraints, and more on how people used communication media. One way people make up for the lack of nonverbal behaviors and cues in primarily text-based environments is by using paralanguage, specifically emoticons.

“Emoticons” is short for *emotion icons*; emoticons are ways to use text to represent emotional and personality nuances present in face-to-face communication. For instance, people use “:-)” to show that they are happy or smiling. When used in text-based EMC (e.g., e-mail, threaded discussion forums, texting, social networking), emoticons function as textual representations of the nonverbal behaviors and cues prevalent in face-to-face communication, designed to convey clarity of intent and emotion in efficient, direct, and transparent ways.

In the late 1990s, researchers began arguing that emoticons are one way to establish social presence in online courses (Rourke, Anderson, Garrison, & Archer, 1999). Since that time, emoticons have become a conventional method of expressing emotion and establishing social presence in the online classroom. However, to-date, there has not been a comprehensive literature review on emoticons’ role in the online classroom. Therefore, in this chapter, we present a review of the literature on emoticons used in support of online learning, with the goal of improving future practice in and research of online teaching and learning. We also provide instructional recommendations for online educators.

---

1 “Computer-mediated communication” is a dated term. People now use a variety of different devices to communicate with each other. Thus, it is perhaps more useful to think in terms of electronically mediated communication (EMC). However, the majority of EMC is still text-based and therefore lacking nonverbal cues.
METHOD

There are multiple types of reviews of the literature (Jackson, 1980). For our literature review, we were specifically interested in summarizing the research on emoticons, in an effort to inform the research and practice of online education. Therefore, we conducted an integrative review because “the goal of an integrative review is to summarize the accumulated state of knowledge concerning the relation(s) of interest and to highlight important issues that research has left unresolved” (Cooper, 1982, p. 292). We began by searching ProQuest and EBSCO using the keywords “emoticons,” “online,” and “learning” in an effort to identify literature on emoticons related to online learning and online education.

Initially, 58 articles were identified by searching ProQuest and EBSCO. All duplicate articles, as well as articles not reporting on empirical studies, were immediately removed from the list. This left 46 articles to be reviewed. Each article was randomly assigned to three different reviewers; the reviewers then independently reviewed each article, recording key findings specific to online education and EMC. The reviewers’ notes were then analyzed for emerging themes. Some additional studies were identified through the process of reviewing the 46 articles. Finally, we integrated literature we knew from our previous work that addressed the value of emoticon use in online education. This resulted in a total of 67 articles reviewed for this study. Our aim was to synthesize the existing work in this field, as well as to offer new perspectives on the literature related to the use of emoticons in EMC. Observed gaps in the literature are noted and recommendations for further research and instructional application are discussed at the end of this review.

EMOTICONS AND ELECTRONICALLY MEDIATED COMMUNICATION

In face-to-face interactions, nonverbal behavior communicates quite a lot about intent. Those behaviors—such as facial expressions, the placement of head and shoulders, the use of hands—can deliver information, regulate the interaction, and express feelings and intimacy. In online communication, emoticons may be used to help achieve the same thing by serving as “nonverbal surrogates” (Derks, Bos, & Grumbkow, 2007, p. 843).

Emoticons are “graphic representations of facial expressions” (Walther & D’Addario, 2001, p. 324), which deliver emotional, rather than task-oriented information (Ganster, Eimler, & Kramer, 2012) and index a user’s affective stance (Park, 2007). Most emoticons are well known and
commonly recognized symbols among users of EMC. They often act as substitutes or surrogates for nonverbal cues, which are usually absent in text-based EMC. Sometimes they are used as a compliment to a text message (Stapa & Shaari, 2012). Smiling is a common human reaction mostly used to indicate happiness; hence it is not surprising that it has found a symbolic representation in EMC in the form of emoticons and smileys. There are two types of smileys, the icon (or emoji), which pictorially represents a smiling human face 😊 and keystroke-based, symbolic emoticons such as :-). They have the same impact in terms of how a message is interpreted. However, some argue that the emoji smiley has a stronger impact on the personal mood of the viewer than a keystroke-based emoticon (see Wortham, 2011); this may be due to the wide range of emoji icons now possible in EMC, or because emoji are a more realistic portrayal of human expressions. A study examining “Facebook” conversations of Malaysian college students, between the ages of 18 and 24, demonstrated that almost all sentences contained smileys or some other emoticons (Stapa & Shaari, 2012), thus showing how dominant a role these forms of nonverbal communication have begun to play in the social exchange of college students.

EMOTICONS AND GROUP DIFFERENCES

Since the emergence of emoticons, researchers have been interested in whether group differences exist in emoticon use (see Brunet & Schmidt, 2010; Locke & Daly, 2007; Wolf, 2000). Differences between men and women’s use of emoticons remain the most popular groups to compare. However, in many ways, the research remains inconclusive. In one early study, Wolf (2000) found that both men and women used more emoticons in mixed gender groups than within same gender groups. In another study, Brunet and Schmidt (2010) found that women, under certain conditions, may use emoticons more than men. They found women who were visible through a webcam used significantly more emoticons than those who had their webcams turned off (Brunet & Schmidt, 2010). But when the webcams were turned off, there was no significant difference in emoticon use between men and women. This is surprising because one would expect emoticons to be used more when the webcams were off than when they were on. Brunet and Schmidt suggested that this may have been due to the fact that women felt more “societal pressure” (p. 203) than men, to appear friendly in an online conversation. Wolf’s (2000) and other researchs, though, also suggest that it is not simply a matter of whether or not women use emoticons more
than men, because men and women actually might simply use emoticons in different ways. For instance, Wolf (2000) found women used emoticons more for humor, whereas men used them more for sarcasm and to tease. In another study, Huffaker and Calvert (2005) found men used more flirty emoticons than women.

Researchers have also investigated how emoticon use differs and/or manifests in different ethnic groups. For instance, Locke and Daly (2007) found that Chinese participants use emoticons more than non-Chinese participants. In another study, Kanayama (2003) focused on elderly people in Japan’s participation in virtual communities. He found that elderly people enjoyed using emoticons and sharing stories with others as they connected and built supportive relationships online.

Researchers have also studied how age influences emoticon use (Fullwood, Orchard, & Floyd, 2013; Kanayama, 2003; Krohn, 2004). For instance, Krohn (2004) argued that people of different ages use emoticons differently (if at all) and therefore emoticon use—at least in business settings—should be based on one’s age or generation. In fact, Krohn (2004) recommended that emoticons be freely used with Millennials (those born after 1980 and coming of age after 2000), but used progressively conservatively with Generation Xers (those born between 1964 and 1980); Baby Boomers (those born between 1946 and 1964); and Traditionalists (those born before 1946). Kanayama (2003) and Fullwood et al. (2013), on the other hand, did not find age influenced emoticon use. Fullwood even questioned whether a convergence of communication styles happens with age—with older people adapting a younger style of communication. At the same time, research has shown that (regardless of age) some EMC may not contain any emoticon use (Pillai, 2009).

**EMOTICONS AND SOCIAL CONTEXTS**

Overall, the research as a whole suggests that the social context or environment (i.e., both the application as well as the context) possibly influences how people use emoticons more than any single variable, such as gender, nationality, or even age (Derks et al., 2007; Fullwood et al., 2013). For instance, depending on the nature of the interaction—e.g., whether or not the interaction is task-oriented or socio-emotional—people may or may not feel the need to textually express nonverbal behaviors. In an experimental study, Derks et al. (2007) put students in one of two groups: either in a “socio-emotional” tasked group or in a “task-oriented” group.
Students responded to text messages significantly more often with an emoticon in the socio-emotional group than in the task-oriented group. Derks et al. (2007) speculated that these results were reflective of societal norms in which it is more appropriate to express emotions with friends and family in social contexts, than with colleagues in professional contexts. They concluded that social context matters in online communications and that social context influences whether or not emoticons are used. They gathered additional support for this claim in a later study, where they found that participants used more emoticons while communicating with friends than with strangers and used more emoticons in positive contexts than in negative ones (Derks, Bos, & Grumbkow, 2008).

All of this research suggests that even though emoticons are an effective way to make up for many of the cues absent in text-based EMC, people use emoticons in different ways, most of which appear to be dictated by one’s personal preference, experience using emoticons, and immediate context.

EMOTICONS AND ONLINE LEARNING

Online education is a unique social context. The following sections are structured around emoticons use for: (a) improving communication; (b) enhancing social presence; and (c) building community in an online education context.

Improving Communication

Moore (2013) and others posited that there is a transactional distance in online education—that is, a psychological and communication distance between an instructor and students. This distance needs to be overcome if “effective, deliberate and planned learning is to occur” (Chen, 2001, p. 459). Overcoming this transactional distance can also help improve students’ overall satisfaction with their educational experience (Stein, Wanstreet, & Calvin, 2005). One way to address this transactional distance is through improving EMC with the intentional use of emoticons. Emoticons can make communication more efficient, effective, clear, and fun (Huang, Yen, & Zhang, 2008; Kindred & Roper, 2004; Varnhagen et al., 2009).

People use emoticons in three main ways to improve communication. First, people use emoticons to indicate emotion by reflecting facial expressions (Dresner & Herring, 2010). For example, :-( means sadness—which in this case is used to reflect an emotional state. A second way people use
emoticons is to indicate nonemotional sentiments that are tied to facial expressions (Dresner & Herring, 2010). For instance, :-) indicates sarcasm or irony. A third way that people use emoticons is to indicate illocutionary force (Dresner & Herring, 2010). For instance, “What’s wrong with you? :-[]” sends a different message than “What’s wrong with you? :-{” and makes the author’s intent clearer (Dresner & Herring, 2010).

Emoticons, used in any of these ways, are very helpful at clarifying text-based messages (Derks et al., 2008). Emoticons can make the intention of a message clear (Lo, 2008) as well as strengthen the intensity of a message. A positive message, for instance, with a smiley-face emoticon can be perceived more positively than a positive message without a smiley-face emoticon (Derks et al., 2008). At the same time, however, an emoticon does not carry more communicative weight than the main (text-based) message. For instance, emoticons are not central or vital enough to change the valence of a message; that is, a positive message accompanied by a frown is still perceived as mostly positive, and a negative message accompanied by a smiley-face is still perceived as mostly negative, although, as part of a politeness strategy, emoticons can also be used to soften a negative tone of criticism, disapproval, or sarcasm (Locke & Daly, 2007; Stapa & Shaari, 2012). Research suggests that these different uses of emoticons can counter the ill-effects of absent social context cues specifically in educational settings (Tu & McIsaac, 2002).

Researchers have found that emoticons are also helpful at improving communication for second language learners (AbuSa’aleek, 2013; Beatty, 2003; Crystal, 2001). But Halvorsen (2012) did find that although students in his study used emoticons pervasively in their writing, the pattern of usage varied by individual and was influenced by things, such as the individual’s previous experience with EMC.

Some researchers have also tried to analyze how or when people use emoticons within a given message to improve communication. Research has shown that emoticons are usually placed as closers, openers, or interjectors in written conversations. In fact, Provine, Spencer, and Mandell (2007) found that emoticon placement aligned with the punctuation effect—that is, occurring at pauses, phrase boundaries, and the beginnings and ends of questions and statements.

It is important to note, though, that emoticons do not always improve communication. For instance, emoticons can lead to miscommunication and misunderstanding (Derks et al., 2008). This happens in part because of a lack of agreed-upon definitions of emoticons (Averianova, 2012;
Chen, 2006; Loewen & Reissner, 2009). But emoticons can also be used to deceive or hide meaning. For instance, people can use emoticons in text-based EMC to hide how they are really feeling. In other words, a “participant might frown at the keyboard but strategically decide to type a strategic smile” (Marvin, 1995, para. 13). But despite these possible drawbacks, the research reviewed in this chapter as a whole suggests that emoticons can improve communication even in educational settings.

**Enhancing Social Presence**

Social presence was originally defined as the sense that another person is “real” and “there” when using a communication medium (Short et al., 1976). Over the years, online educators have found that social presence is important in online education because it sets the climate for learning to take place (Caspi & Blau, 2008). Research also suggests that there is a positive correlation between students’ perception of social presence and perceived learning and learner satisfaction (Richardson & Swan, 2003; So & Brush, 2008). The lack of nonverbal and relational cues in EMC, though, can make it difficult to establish one’s own social presence or perceive another person’s social presence (Lowenthal, 2009).

Research suggests that emoticon use can enhance students’ perceptions of social presence in online learning environments that rely predominantly on text-based EMC (Aragon, 2003; Lahaie, 2007; Tu, 2002). Gunawardena and Zittle (1997) were one of the first to research social presence in an online learning setting. They were interested in participants’ use of emoticons in an online education conference. Gunawardena and Zittle found that students with higher levels of social presence, “enhanced their socio-emotional experience by using emoticons to express missing nonverbal cues in written form” (p. 23). Garrison and his colleagues later identified emoticon use as an observable indicator of affective/emotional expression and therefore an indicator of social presence in their Community of Inquiry model (see Garrison, Anderson, & Archer, 2000; Rourke et al., 1999).

In one study of social presence across three different computer-mediated communication systems, Tu (2002) found that “students used emoticons and paralanguage to compensate for the lack of social context cues” (p. 15). Tu also noticed that students tended to use smiley- and frown-face emoticons the most in the communication systems. Early research on social presence and online education, though, suggests that some people are not familiar with using emoticons in EMC (Tu, 2002; Weiss, 2000). As a result, Weiss (2000) recommended explicit encouragement in the use of emoticons
and possibly even including a list of various emoticons one could use in text-based communication. Similarly, Tu (2002) argued that instructors should model the effective use of emoticons in online courses. In fact, Tu and McIsaac (2002) later found that most students respond positively to the use of emoticons (p. 143), thus supporting the need to help students effectively use emoticons in EMC.

Yamada and Akihori (2007) found in a later study that students’ use of emoticons heightened their sense of social presence. In their study, a student’s use of emoticons often led to more responses from other students to their posts. In another study, Cobb (2009) found that 70% of students in an online nursing program used emoticons. Using the same instrument as Gunawardena and Zittle (1997), Cobb found students in an online nursing program (with high use of emoticons) actually had a higher overall social presence score than participants in Gunawardena and Zittle’s foundational study. Cobb suggested that this difference could be due to users’ increased use of EMC over the past decade. It is reasonable to expect, as students use EMC more for personal as well as educational purposes, that they will become more adept at using paralanguage and emoticons to establish social presence and make up for the cues filtered out of EMC. However, this does not discount the need of instructors and instructional designers to intentionally find ways to design for social presence in online courses (Aragon, 2003; Dunlap & Lowenthal, 2014; Greyling & Wentzel, 2007).

**Building Community**

Online students may feel isolated and alone in online courses (McInerney & Roberts, 2004); online students report missing the social presence—specifically the sense of being perceived as real and perceiving others as real—that they more easily establish in face-to-face courses (Stodel, Thompson, & MacDonald, 2006). Students often perceive the lack of a community as an impediment to their success in online courses (Song, Singleton, Hill, & Koh, 2004). Research has shown a relationship between students’ sense of community and their actual success in online courses (Conrad, 2005; Sadera, Robertson, Song, & Midon, 2009; Swan, 2002). In addition, various learning theories stress the importance of social context, collaboration, and discourse in the construction of knowledge (see Lave & Wenger, 1991; Lowenthal & Muth, 2008). For reasons such as these, online educators strive to build community in online courses.

There are various types of learning communities (Zhao & Kuh, 2004). The online learning communities that educators try to build in online
courses have been described as “bounded learning communities” (Wilson, Ludwig-Hardman, Thornam, & Dunlap, 2004). A bounded learning community is a learning community formed within a formal course. In bounded learning communities, students often do not choose their instructor or fellow students. These communities take place over a fixed period of time (e.g., a semester); and participation in these communities is often required in some way. A bounded learning community though rarely simply forms on its own (Wilson et al., 2004). It takes careful upfront planning on how best to engage students with the course content, their peers, and their instructor as well as how best to use EMC (Swan, 2002).

Building community in an online course begins and ends with learner interaction. In other words, learners must first login to their online courses and interact with each other, their instructor, and the course content for a community to even possibly form. Research suggests that frequent interaction alone is not enough. Instead, it is the quality of the interaction that matters (Goertzen & Kristjansson, 2007). The cues filtered out of EMC, however, can make it challenging for learners to effectively interact and communicate with each other online. The “Community of Inquiry” model suggests that affective, interactive, and cohesive communication are needed to build social presence and a community of learners (Rourke et al., 1999). More specifically, paralanguage in general, and emoticons in particular, can help facilitate community building by clarifying EMC, establishing social presence, and building cohesion (Huang et al., 2008; Rourke et al., 1999). Goertzen and Kristjansson (2007) found, in one study, that paralanguage and emoticons enable people “to project a sense of personality, familiarity, and closeness, along with various degrees of solidarity and alignment …” (p. 220) and that “social presence is essential to increasing a sense of belonging and social cohesion in the community as well as facilitating collaboration” (p. 213).

Members of a learning community must be able to disagree with others, however, when needed. Goertzen and Kristjansson (2007) pointed out, “reviewing and potentially critiquing the work of peers is risky business” (p. 223). As a result, learners often engage in a variety of face-saving acts. For instance, Goertzen and Kristjansson found that learners surround requests for help, clarification, and change with positive comments to improve group cohesion. Paralanguage and emoticons are also often used to avoid potential conflicts in a conversation that has a chance of getting acrimonious, or merely to soften the serious nature of a conversation (Stapa & Shaari, 2012). For instance, in a study of college classroom discussions,
Vandergriff (2013) found emoticons, nonstandard/multiple punctuation, and lexical surrogates were often used as an avoidance strategy when a participant did not want to disagree openly. Emoticons can also be used in a humorous way to politely disagree as well as to convey complex meanings, such as sarcasm and frivolousness.

Communities are constructed and maintained in part with language (Street, 1984). Participation in a community—even an online bounded learning community—requires knowing the specific language and literacy skills of the community (Gee, 1990, 1998, 2000; White & Lowenthal, 2011). Thus, acceptance within a community requires one knows and employs the language of that community. In one study, Tu (2001) found that once students became comfortable with their classmates and learned any commonly used “Netspeak” or emoticons, they reported feelings of belongingness and a sense of themselves as full-fledged members of the learning community (Tu, 2001). In another study of a gaming community, Peña and Hancock (2006) found that more experienced gamers used more emoticons and other Netspeak conventions in their communications than less experienced gamers. Their usage was tied to their existing membership in the gaming community and the established shared understanding of the Netspeak conventions used by the community. Similarly, while research on emoticons suggests that people use emoticons more with friends than with people they do not know (Huang et al., 2008), as people become more comfortable with each other, research suggests that they often feel less of a need to clarify every comment with positioning (Goertzen & Kristjansson, 2007).

People also have to become comfortable using various types of EMC media to feel a part of any given learning community. Meyer (2003) suggested that one’s comfort level with text-based EMC will likely depend on one’s ability to “to create a realistic ‘self’ in written responses” (p. 57). The constraints of text-based EMC have led some to argue that synchronous EMC is more effective in improving communication and interaction than asynchronous EMC (Fadde & Vu, 2014; Wang & Newlin, 2001); consequently, synchronous EMC may be better at helping participants develop social presence and possibly lead to a greater sense of community (McInnerney & Roberts, 2004, p. 75). Learning communities are, however, complicated (Zhao & Kuh, 2004). Therefore, emoticons alone are not enough to help a learning community form and persist. Furthermore, it is possible that only certain types of courses and learning audiences benefit from bounded learning communities.
LIMITATIONS AND GAPS

Our review of the literature is limited in scope, in part due to the keywords selected. After initially testing a few different keywords as search terms, we settled on using: “emoticon,” “online,” and “learning.” Additional keywords (e.g., “e-learning” or “education”) could possibly have resulted in additional articles that we missed. However, our initial search was used simply as a starting point; many of the articles we read pointed us to additional studies. A bigger limitation, though, is due to the research itself. First and foremost, there are very few studies that primarily focus on emoticons and online learning or online education. Therefore, the majority of the studies we reviewed did not research emoticon use in educational settings. Educational settings have their own norms and ways of communicating and being. Further, learning management systems, and specifically threaded discussions, differ in important ways from other communication platforms, such as chat rooms or instant messaging applications. Fullwood et al. (2013) questioned to what degree communication platforms, such as chat rooms, are a unique “genre” that influences how people communicate; they continued to argue that “there are recognized conventions or etiquette that guides our online behaviors in specific environments, encourage a particular style of communication” (Fullwood et al., 2013, p. 658). Another shortcoming of general research on emoticons is that it tends to focus predominantly on issues of, e.g., gender (Fullwood et al., 2013) and not on other important questions (e.g., its use in education).

There is still so much we do not know about emoticon use. For instance, it is very possible that emoticons are not always used to convey emotion; people could simply be influenced by the way others use them. Others could just be habitual emoticon users (Lowenthal, 2012), in much the same way that some people simply use their hands obsessively when they speak. Emoticon use may sometimes be more of a generative rather than a communicative act, in the sense that it serves the writer more than the reader (Walther & D’Addario, 2001, p. 343). Also, the writer may simply be feeling too lazy to use words—opting instead for the use emoticons. Hence, the overuse of emoticons due to the lack of effort on the writer’s part may wane its effectiveness, such that readers start to ignore their presence (Walther & D’Addario, 2001, p. 342). Hence, the question arises: How, and in what context, can emoticons be used most effectively to bring about maximum learning and optimal information exchange?
Research on emoticons also does not adequately acknowledge how emoticon use might be changing over time. Emoticon use, or more generally the effective use of EMC, can be viewed as a type of literacy. It is likely that peoples’ emoticon use is influenced by a host of factors, one of which being their prior experience using different types of EMC (see Fullwood et al., 2013). Emoticon use will change as people become more literate with EMC; in other words, it is probable that emoticon use is changing as people change (Huffaker & Calvert, 2005). However, some people are still “turned off” by the use of emoticons (Provine et al., 2007), which can often cause problems when studying emoticon use in educational settings where faculty prohibit the use of emoticons or Netspeak of any kind (Pratt, 2010).

**FUTURE RESEARCH**

Our review of the literature on the use of emoticons to support online learning has revealed new lines of possible inquiry. First, researchers need to examine how emoticons can be used to maximize student engagement and achievement in online courses. For instance:

- Does emoticon use in instructor feedback reduce transactional distance between the students and the instructor in an online class?
- How can paralinguistics enhance the online learning experience for students?
- Do emoticons have a more positive effect on improving communication in online courses when combined with other strategies? If so, what other strategies, and why?
- Is there a relationship between emoticon use and student persistence in online courses?
- How does the overuse of emoticons limit their usefulness in the online classroom?

Second, researchers need to focus more specifically on how emoticons are used to establish and maintain social presence. For instance:

- Are emoticons more effective for enhancing social presence for some learners than others?
- How can the intentional use of emoticons reduce the transactional distance and increase sociability between students in online courses?
- Do emoticons have a more positive effect on enhancing social presence in online courses when combined with other strategies? If so, what other strategies, and why?
Is developing an effective use of emoticons in online courses a good use of an instructor’s time, or are there other strategies (e.g., the use of video) that may have a more consistent positive effect on social presence in online courses?

Last but not least, many questions remain on how emoticons can help build and maintain effective learning communities. For instance:

- How does the background (culture, language proficiency, internet/IM use experience) of students affect their use and reception of emoticons in online courses?
- Why do some people respond positively and others respond negatively to emoticon use?
- How are emoticons used in courses with a strong sense of community?
- How does an instructor’s use, modeling, and encouragement influence students use and perceptions of emoticons?
- How does emoticon use change over time in online courses and online programs?

As people’s use of EMC increases, the platforms people use become more sophisticated, and people’s own comfort level with EMC as a vehicle for communication, collaboration, and expression increases, people’s use of emoticons in online education is likely to change. As such, research on emoticons should continue. At the same time, online educators should keep in mind that emoticons are just one of many ways to express emotion and intent in the online classroom and that emoticons cannot magically solve all of the problems of distance and isolation in online courses.

**INSTRUCTIONAL RECOMMENDATIONS**

Given the ubiquitous use of EMC in business, as well as professional communities of practice, helping students learn to communicate and collaborate well using EMC technologies (i.e., becoming literate with EMC) is an important instructional goal. Helping students understand the role of emoticons, and Netspeak in general, is an appropriate element of professional preparation. The following are some instructional recommendations that emerged from our review of the literature on emoticons:

1. **Enhance teaching presence.** Online educators should use emoticons when communicating with students to increase teaching presence. Emoticons are one way instructors can express emotion, as well as clarify the expression of emotion or intent. This may help students better understand their instructor’s approach to the course and the content (i.e., teaching
presence) while at the same time getting a better sense that their instructor is “real” and “there” (i.e., instructor’s social presence).

2. Provide personalized feedback. Personal, individualized feedback can help establish social presence in online courses (Dunlap & Lowenthal, 2014). Instructors should strive to use paralanguage and emoticons to help personalize and humanize feedback.

3. Soften critical feedback. Emoticons may be used to soften the tone of critical feedback so that students are more open to receiving and processing critical feedback; emoticons can essentially have a similar effect as audio feedback, which has been shown to help students hear the nuances in an instructor’s voice (Ice, Curtis, Phillips, & Wells, 2007; Wilson, 2009).

4. Establish clear expectations for emoticon use. Students are often unsure how best to communicate in online courses. Even when students are well-versed in EMC for social purposes, they are often unsure of the appropriate way to communicate for an academic/professional context. Therefore, instructors should establish clear expectations about the use of emoticons, as well as other paralanguage, in their online courses. When establishing expectations for emoticon use, instructors should keep in mind that students from different countries, from diverse cultural contexts, and with different levels of experience with EMC (i.e., with different levels of digital literacy) might need additional support (Vrasidas & McIsaac, 1999). Instructors should also reinforce these expectations through modeling the appropriate and effective use of emoticons (Vrasidas & McIsaac, 1999; Woo & Reeves, 2008) and possibly even holding students accountable for their appropriate and effective use of emoticons.

5. Go beyond emoticons. Emoticon use does not always address the instructional goals of improving communication, enhancing social presence, and building effective learning communities in online courses. Emoticons are one strategy that should be used in conjunction with others to achieve these goals.

CONCLUSION

The effective use of emoticons to improve communication, enhance social presence, and build community is a digital competency, one aspect of a person’s digital literacy. As such, effective emoticon use has the potential to enhance a person’s ability to accurately and appropriately use EMC. As the literature reviewed in this chapter demonstrates, the interpretation of
emoticon use may not be universal. We identified a tension between the usefulness of emoticons and some people’s perceptions of emoticons as unprofessional, in both the literature as well as our own personal experiences. Loewen and Reissner (2009) described an exasperated teacher who expressed disapproval over students’ use of emoticons by chastising them with comments such as: “What do you mean? and What language are you speaking?” (p. 111). In our introduction, we quoted a student who had similar prior experiences with instructors who disapproved or even prohibited emoticon use; this student expressed relief in finding an instructor who permitted the use of emoticons because the student believed they could effectively “save a lot of grief and help to establish community.” Reaching a universally compatible understanding of or standard for how emoticons may be used in educational and professional contexts is a task that may be useful to undertake if educators are to prepare students to meet the interaction needs of the social and professional world in which they will participate. We have identified some instructional recommendations that might help online educators accomplish just this (e.g., modeling best practices; establishing clear expectations regarding emoticon use). However, educators should keep in mind that emoticon use is just one strategy to improve communication, establish social presence, and build learning communities. Emoticons may be one of the rays of sunshine that helps online educators grow healthy, hearty, and vibrant flowers.

REFERENCES


The powerful influence of emotions in learning is widely recognized in the field of education. As educators, we have witnessed a wide range of emotions from students in response to the same circumstances involving their learning. For example, we have seen students cry in sadness, celebrate in joy, or be content when earning a C letter grade. We have seen some student’s excitement or dislike for our classes. This range of emotions impact students’ learning experiences inside and outside the classroom. Consequently, it is imperative to ask: Can we predict how students will react to different learning scenarios? Which students will be excited? Which are going to be anxious and fearful? Which will have neutral feelings? In-line with the aforementioned questions asked—Is it possible to modify students’ reactions to an educational scenario? For example, can we create interventions to reduce students’ fear for learning topics that make them uncomfortable? In this chapter, we argue that we could perhaps reduce some aspects of student’s fear of learning by integrating affective technologies.

Technology is a potential medium to improve students’ learning experiences (Foster, 2004; McMillan, 2009; Olive et al., 2010; Papert, 1980). Research supports the contributions that emotions, such as frustration and respect (from a teacher), can have on students’ achievement. Within the field of robotics, the subcategory of human–robot interaction (HRI) includes the study of how robots can learn to interpret human emotions and how robots should express emotions. The ability of robots to show and read emotions has been shown to be critical when robots interact with humans. When considering further the intersection of education and HRI, there is a need to understand the interdependency of emotions and cognitive growth. Future robots cannot be solely based on a cognitive component, marginalizing the emotions that facilitate or hinder cognitive development and growth. Rather, the cognitive and the emotional components must inform each other to improve the learning of the human and the robot.
This chapter presents a brief discussion of the role of emotions in education, then provides a rationale for why one might consider using technologies in the classroom, and ends by presenting cases where the emotional context of the HRI promoted, hindered, or failed to help students in the learning process.

THE ROLE OF EMOTIONS IN EDUCATION

When James (1884) posed the question, “What is an emotion?,” he proposed that emotions were physiological responses. Although not all researchers agreed with him, his idea was followed by the development of several competing theories of emotions. Even today, researchers have not agreed on a universal definition of an emotion. Yet, the general population is able to perceive and name different emotions. We all seem to know when someone is happy, sad, angry, or scared. James speculated that researchers studying the brain, in the 1900s, disregarded emotions because they were far too complex to study. Perhaps, at the time, it was more sensible for physiologists and psychologists to concentrate on other characteristics of the brain, such as its sensory and motor centers. Yet, James along with past and current educators believed in the influence that emotions play during decision-making. Through time, educators have sensed that learning is more than a set of cognitive processes. There is growing evidence that emotions influence the decisions people make (Cytowic, 1993; Damasio, 1994; LeDoux, 1996). Additionally, emotions are starting to be viewed in a more positive light. Emotions are not only part of the irrational decisions humans make, they are now starting to be recognized as integral to the rational decision-making process.

Why Should We Value Emotions Along with Cognition in Education?

Although researchers in psychology, cognitive science, and artificial intelligence have concentrated predominantly on the cognitive processes of the brain, recently attention has been given to understanding how emotions and cognition collectively influence our decision-making and our thinking (Gadanho & Hallam, 2001). For a long time, educators have sensed that learning is more than a cognitive process. There is more to learning than memorizing facts and processing information. Shih, Chang, Chen, Chen, and Liang (2012) found that although students in Taiwan, one of the high achieving countries in the TIMSS (2003, 2007), excelled in mathematics and
science in the international exams, they did not necessarily like mathematics and they reported less confidence in mathematics than students from other lower performing countries. For instance in 2003, when Taiwan was ranked fourth among the 49 countries that participated in the TIMSS exam, 34% of the 4th graders who participated in the TIMSS exam disagreed with the statement “I like to learn Mathematics,” and only 41% of the 4th graders reported high self-confidence in mathematics. It is possible that as the knowledge gains of the students were maximized, some of the students’ positive emotions for learning mathematics were hindered or perhaps overlooked altogether. Hence, demonstrating mastery of a topic does not translate into enjoyment of or a liking for the topic.

Sometimes students lack both a love for learning and competence in a subject area. One reason for this outcome could be a consequence of overlooking the emotions involved in learning by educators (Leonard & Martin, 2013). Several educational researchers have found that when students perceive their classroom teachers as supportive of their academic pursuits, students’ interest in their academic studies and consequently, their academic performance improves (Felner, Aber, Primavera, & Cauce, 1985; Goodenow, 1993; Midgley, Feldlaufer & Eccles, 1989; Wentzel & Asher, 1995). Hence, teacher support, defined as “the extent to which students believe teachers value and establish personal relationships with them,” is critical in the learning process of some students (Ryan & Patrick, 2001). It is possible that students who perceive that teachers are supportive, the students are more inclined to persevere academically and socially because they do not view the challenges from a negative perspective, as they know that their teacher will give them the support that they need after they explore on their own.

Wentzel (1997) found that students are indeed more likely to be interested in classroom activities when they feel supported and valued by their teacher. From the student perspective, a teacher that makes them feel valued and who cares about them displays the following characteristics: they model a caring attitude toward their work; they have individual expectations based on individual differences; they have democratic interactions with the students; and they are nurturing. Additionally, these teachers show concern for the student beyond the classroom walls, such as approaching students to inquire about their life when they notice something abnormal in the student’s behavior. Other researchers have reported similar findings. Students’ perceived support from adults has resulted in positive academic outcomes (Cauce, Felner, & Primavera, 1982; Felner et al., 1985; Phelan, Davidson, & Cao, 1991). Hence, when a learner perceives support and care from a
more knowledgeable adult, the perception could contribute to positive academic outcomes.

One concern regarding perceived care among researchers is the correlation between parenting and perceived care. Experts in African-American studies have argued that teacher care is interpreted and expressed differently. For example, the parenting dimensions described in Wentzel’s study did not result in positive academic outcomes for minority students (Steinberg, Dombusch, & Brown, 1992). Furthermore, Ware (2006) provides an example in which an African-American educator scolded her class for not doing their homework. Students perceived her demands as a demonstration of teacher care. Hence, culturally relevant pedagogy could expose different dimensions of teacher care that are critical for ethnic groups.

The emotions that a teacher promotes in the classroom impact the learning experiences of students and their engagement with the topic. After all, many of us have heard more than once a child or an adult say something along the lines of, “I need to learn this because I want to make my teacher happy” or “because my teacher loves me” or “I don’t want to disappoint my teacher,” or “because my teacher believes in me.” Assuming that in the near future, we manage to identify most of the dimensions of teacher care that nurture students’ learning, these dimensions should be used to improve the learning experiences of students.

WHAT ROLE CAN TECHNOLOGY PLAY?

Unfortunately, given some of the adverse realities in the teaching profession, it is unrealistic to expect that every student will have access to a supportive and caring teacher. Due to lack of support, among other reasons, the attrition rate for teachers within their first teaching year is over 40%, at the K–12 level (Alliance for Excellent Education, 2011). As budgets continue to shrink, while the student population continues to grow, educators face larger classroom sizes (Alliance for Excellent Education, 2011; Cuseo, 2007; Ingersoll, 2003; Miles & Darling-Hammond, 1998). Hence, even if we manage to place a supportive and caring teacher in each classroom, the teacher might not be able to reach out to every student. Large classroom sizes obstruct educators from giving students the individualized attention that engages them in the learning process (Cuseo, 2007). In the absence of increased budgets to reduce classroom size, more attention should be given to the allocation of instructional resources in our educational system, especially technological resources (Miles & Darling-Hammond, 1998).
Technology could improve the learning experiences of students in large classes and allow a supportive, caring teacher to reach more students. Kenneth Bowles, in the late 1960s, had a vision that 1 day, students will have their own personal computer with a “personalized system of instruction” (McMillan, 2009). He envisioned a future in which interactive computing could improve education by creating better, targeted learning experiences for students. Through his eyes, microcomputers had the potential to transform the big lecture halls of introductory programming classes into more interactive environments that would enable students to write programs, run them, edit them, and try them again during class time when the teacher could help students (Foster, 2004). The vision was beyond using technology to help students memorize information but aimed at improving the learning experience of students by creating an environment where students would receive the support they needed in the classroom, despite large class size. The key to his vision was the fulfillment of students’ needs.

Since experts in emotions argue that emotions give meaning to our lives by creating a structure where our needs, motives, and concerns are assigned different values depending on their importance by considering the needs of students when using technology in the classroom, Bowles was integrating emotions with technology (Bower, 1992). In this sense, when humans make a decision, the decision might be fulfilling a need of the decision-maker. Hence, if educational technologies are to enhance the learning experiences of students, then the designers must keep in mind the needs, motives, and concerns of the students who will be interacting with the technology.

Unlike Bowles’ vision, in the past, technologies functioning by artificial intelligence have failed to incorporate human intuition and emotions in their decision-making process. These technologies are designed to simply use algorithms to maximize the outcome of situations, but if a human’s value system does not match the machine’s value system, then the machine will fail to fulfill the needs of the human. For example, imagine a teenager that gets accepted to Harvard, among other local colleges and universities. If the technology optimizes the teenager’s future based on school’s rankings without taking into account the teenager’s emotions, the machine may select Harvard. Yet, the teenager may choose to attend a less prestigious local college because they value proximity to their family significantly more. The teenager still used cognitive processes to make a decision, but the internal maximization algorithm involved the emotional ties to family. The decision was based on the type of happiness valued.
In short, for technology to have a chance at truly impacting the learning experiences of students and aiding the supportive and caring teacher to reach more students, it will have to treat each student uniquely. It will have to know the needs, motives, or concerns of each student. It will have to embody the characteristics of a supportive and caring teacher. Otherwise, the technology will treat every person the same (Cowley & MacDorman, 1995; Hinde, 1988). This will hinder its performance in an educational setting because we know that humans do not treat everyone the same. Social interaction is a highly unique experience. When we meet people, we like some more than others, based on their personality, what we share with them, how they make us feel, etc. In terms of promoting learning, it is even more important for students to be treated uniquely. After all, from students’ perspectives, having individual expectations based on individual differences is one of the characteristics they value in supportive and caring teachers.

**CASES WHERE EMOTIONS INFLUENCE STUDENTS’ LEARNING IN A TECHNOLOGY ENVIRONMENT**

The main role of a teacher is to help students. In this vein, considering that technology strives to be an effective instructional tool, an educational technology’s main role ought to be to help students. By coupling the human capacity of supportive and caring teachers with an effective technology, the needs of more students might be fulfilled. This, in turn, might result in an increase of positive emotions associated with learning experiences and a reduction of bad emotions, such as fear.

Today, robotics technology seems to possibly fulfill a void, in the area of technology enhanced learning, for contemporary learners. There is a growing number of after school programs, school classes, and competitions based on robotics (Cho, 2011). However, according to Cho (2011), the majority of the existing curricula involving robotics are biased toward engineering-related topics. This is limiting the potential of robots in K-12 education because children with diverse interests outside of engineering or engineering-related fields are excluded from robotics activities and curricula. Other researchers have voiced the need to have more entry points to robotics (Rusk, Resnick, Berg, & Pezalla-Granlund, 2008). By moving away from a robotics-based curricula to a robotics-enhanced education curricula, people from different interest areas will be able to engage with robots (Cho, 2011). At this point in time, their inclusion is important because as the technology continues to gain momentum and support in education circles, it is
important to empirically identify the needs of the learners and how those
needs might be fulfilled by robotic technology to assure that the technology
continues to foster positive emotions among learners. Below, we share
examples of technologies, which either excelled or failed at fulfilling the
needs of the learner, as examples of the potential emotions robotics technol-
ogies could evoke or displace.

Fixing Work Instead of Fearing Learning
In the 1970s, Seymour Papert (1980), an educational theorist, promoted the
use of computing technology as an “object to think with” that children
could use to explore, discover, and construct their own knowledge. He
believed that computing technology could serve as a medium to eradicate
math phobia, which he interpreted in two ways. First, “math phobia”
can be described as fear of mathematics. But, much more interesting, is
Papert’s second description of math phobia, which stems from its Greek
meaning: Math in Greek relates to “learning.” Hence, he presented math
phobia in the broader sense, as a “fear of learning” (Papert, 1980).

Fear of learning is not an unknown phenomenon. How many times have
we heard a friend, relative, or student mention that they are incapable of
excelling at something (e.g., math, writing, dancing, cooking, gardening)?
Yet, inquiring further about the inability exposes the fact that they shy away
from the topic or activity when faced with the possibility of failure. Perhaps,
it may be more accurate to say that at some point, they decided to avoid
failure. It may not be the actual math, writing, cooking, etc. that limits a
person’s achievement in an area, but a prospective failure. What may be
impeding them from pursuing the learning opportunity or activity might
be a lack of “courage” to persevere in an area where they have developed
the emotionally safer approach of giving up as soon as difficulties arise, rather
than daring to spend time in an activity that does not guarantee success
(Wertime, 1979). Preliminary research in neuroscience is starting to reveal
that certain parts of our brains are activated when deciding to tackle or to
avoid a problem. The left and right sides of the dorsolateral prefrontal cortex
activate. Whether a person chooses to tackle or avoid the problem corre-
ponds to which side experiences greater activation (Wu, 2014). Although
this finding is in its infancy, it is valuable to make the learner aware that their
brain considers both possibilities: those of tackling or avoiding a problem and
it is up to them to make the final decision. Of great importance is to find out
how students approach learning opportunities that evoke uncomfortable
emotions, when aware of the possibility that they are unconsciously
sabotaging their potential. If students are asked to approach a learning opportunity every time that their brain chooses otherwise, will their emotions toward that type of learning opportunity change over time? Papert would conclude “yes,” as he believed that a rich learning environment could alter students’ emotions toward learning (Papert, 1980).

In an attempt to eradicate fear of learning, Papert (1980) used technology to create a learning world, where children learned geometry in a more organic way. He introduced “Turtle geometry.” The Turtle was an animal-like robotic toy that moved over ground. The Turtle was also represented digitally on the computer, as a cursor. Children controlled the movements of the Turtle, whether physically or digitally, via a computer using the programming language LOGO. Initially, the Turtle can only perform some basic movements, such as a move forward, a turn, and then puts the pen down to create a trace of its traveled path when given the computer commands of FORWARD, RIGHT, and PEN DOWN, respectively. Next, children were given the learning opportunity to teach the Turtle new words by creating new computer commands, such as TRIANGLE, SQUARE, or any other command the children desired. Through this process, the Turtle exposed children to a learning world, where their mathematical thinking developed by exploring their own ideas and fantasies. Thus, children became creators of mathematics.

In the world in which the Turtle lives, children were not expected to get it “right” the first time they tried to create a new word for the Turtle. In this world, children had to keep on correcting their work until the Turtle worked properly. With this new model for learning, children are not fearful of getting it “wrong.” Instead, they focused on how to correct their work or figured out if their work was even fixable.

Similar to Papert (1980), Fernandes, Ferme´, and Oliveira (2006) observed middle-school students focused on fixing their working out, not on simply accepting that they were wrong, while discovering the definition of a function via a robotics activity. Students were given two graphs of the distance traveled by a robot versus time (one depicted the graph of a function and the other one did not). Students were asked to program their robot to reproduce both graphs. The class was set-up in small groups. The researchers reported the conversations arising from two groups, which were close to each other. Rui, a member of group no. 1, noticed that there was something wrong with one of the graphs and told the teacher that the graph was not good because it required the robot to walk backwards and he thought the robot had to always move ahead. Ricardo, a member of group no. 2,
who was listening to the conversation, added that it “can’t be because the robot cannot walk backwards in time.” Then, the teacher asked them, “What will happen if the robot walks backwards in time?” A few minutes later, Rui answered, “the robot had to be at two places at the same time.” Students recognized this situation as impossible and decided that the second graph could not be achieved.

The study created a situation where the students would fail to achieve the task of the robot’s activity. Usually, students who fail to accomplish a task in a classroom setting tend to blame themselves, feel like a failure, blame their lack of ability, or simply get frustrated and give up. Nevertheless, here, these students kept on thinking about the movements of the robot until they arrived at the conclusion that the task was impossible. At no point, did they express negative emotions toward the learning activity. It is possible that physically seeing the movements of the robot kept on fueling their imagination of what they could do next to fix their work to reach their goal. Once they realized that it was physically impossible to accomplish the goal of having the robot reproduce the graph, they moved on. By allowing students to keep on exploring and fixing their code, students had a tangible avenue to explore their ideas without fear of being wrong. In a way, their needs for comfort and security were met.

**Displacement of Emotions**

Educational robots could also serve as a platform to displace students’ emotions. Goh and Aris (2007) conducted an 8-week study of six secondary school students who were participating in a robotics competition, the Robot Transporter Event. They used the RCX LEGO Mindstorms, a robotics kit consisting of sensors, motors, and Lego pieces. During the study, the researchers interviewed and observed students to determine whether the competition had an effect on the students’ interest in math, science, engineering, and technology as well as their social skills and teamwork. One observation made during the study involved conversations of students, expressing their emotions or feelings toward the robot. For example, one student said that his robot was too tired when the robot failed to perform the expected task.

The scenario in which students verbally assign an emotion, feeling, or state to the robot they are working with, instead of internalizing it themselves, is promising for education. It is possible that as long as students project their emotions of happiness, frustration, etc. onto the robot, they will continue to explore the challenge at hand, since it is the robot’s inability and not their inability. In future studies, it will be of particular interest to identify the
emotions that are internalized by the student versus those that are externalized by the student onto the robot. Ideally, students should internalize emotions that will continue to nurture their intellectual growth and externalize those emotions that might hinder their learning. On the other hand, from the educators’ perspective, how the emotions are projected is a concern. Even when these emotions are projected toward the robot, they can disrupt the classroom. Perhaps, future work might involve what to do with the projected emotions. Is it more beneficial for the student to learn to process some of the emotions instead of simply displacing the emotion toward the robot? If so, which ones?

**Affection and Gratitude Display Toward a Robot Versus Disengagement**

The same way students value individual attention from a supportive, caring human teacher; they value the same from a robot teacher assistant. Han, Kim, and Kim (2009) took a humanoid-like robot, Tiro, to serve as a teaching assistant in a 4th grade music class in Korea. They presented their classroom observations of one of the classroom lessons. The robot helped the teacher take attendance; it presented the goals of the lesson; it performed the music activities along with the teacher; and it selected students to practice the activities with him. Three male students who were selected by the robot to perform the activity together showed their gratitude to the robot before going back to their seat for selecting them. The authors shared a picture where a boy is giving the robot a kiss. The authors also noticed that male students showed more affection, paid more attention, and were more motivated with the robot than female students. There were also some students who ignored the teacher because they preferred the robot.

The expression of gratitude of the students toward the robot is in agreement with the findings in Wentzel’s (1997) study. In her study, when the teacher “asks if I need help, takes time to make sure I understand, calls on me,” it is perceived by the student as a demonstration of teacher care. It is possible that when the robot picked students to work with on the activities, the students perceived the gesture as a form of teacher care. It also highlights the importance of giving individualized attention to students in the classroom setting. By doing so, instead of a child showing frustration for not been able to follow class activities, the child shows gratitude for getting attention and help.

The identified failure of the robot to engage females needs further exploration. Although the factors that led to their disengagement are unknown,
given that the classroom is a dynamic learning environment, robots should be programmed to deal with as many scenarios related to student learning as possible. If the technology lacks the intelligence to read human emotions or comprehend the impact of its processes, the educator should maintain an awareness of these deficiencies and have the ability to control the technology to assure that the robot meets the needs of all students, not only a subgroup.

**Pity—a Caring Emotion from Student to Robot**

When the robot excels at developing a relationship with students, some students display caring emotions toward the robot. Otherwise, students can become indifferent to the robot. Kanda, Hirano, Eaton, and Ishiguro (2004) performed an 18-day field trial in an elementary school, in Japan, with 1st and 6th grade students. The researchers were interested in exploring whether robots could form relationships with children and if children can learn from robots as they learn from their peers. Two “Robovie” humanoid robots, one for each grade level, were placed outside the classroom; in a corridor close to the students’ classrooms. Four cameras and two microphones were installed along the corridor to capture the interaction of the students with the robots. The physical infrastructure was the same for the 6th graders. There were 119 1st graders and 109 6th graders. Each student physically carried an ID tag with him/her. The robot used the ID tag to identify the child and to keep track of the interaction with the child. The robot’s aim was to improve the English proficiency of the students. Hence, researchers administered an English quiz three times during the field study: pretest, test no. 1 after week 1, and test no. 2 after week 2.

They found that students interacted with the robot frequently during the first week. However, during the second week, student interaction with the robot declined. The two students who continued interacting with the robot into the second week shared that they continued interacting with it because they felt pity for the robot, as there were no other kids playing with it. That is, the students who continued to engage with the robot did so due to their feelings of concern for the robot, not because the robot was able to connect with them. Consequently, while ensuring that a relationship exists between the students and the robot is important, the nature of their relationship is equally important.

**Existing Emotions Will Emerge even in Robotics Technology**

Silk, Higashi, Shoop, and Schunn (2010) found that the sole presence of mathematics within a robotics activity does not lead to mathematics learning
nor does it diminish the negative emotions students have toward the topic. For 3 years, their research team conducted research in middle-schools, after school programs, and with teachers using Lego Mindstorms NXT 2.0. They designed a robotics activity, which required the use of ratios to make robots with different size wheels to synchronize dancing. They found that for mathematics to be salient, lesson plans needed to be purposely designed to highlight the mathematics present in the activity. Otherwise, the technology would not result in mathematics outcomes gains for the students. Additionally, they found that students have internalized negative feelings toward mathematics. For instance, the researchers shared an experience in which a student was engaged in the robotics activity until he realized that he was doing mathematics. After the realization, he dropped out of the discussion. This does not mean that robotics activities cannot empower students with math but what it may mean is that the design of the robotics activity was not solid enough to preserve the interest of the student and the “coolness” of the robot wore off before the student had developed a deeper engagement with the robotics activity. To divert the existing negative emotions toward math while the students learn math with the robot, the researchers redesigned the activity based on four principles. The robotics activity must be able to sustain students’ engagement, it should focus on the key content to be learned, it should generalize understanding, and it should require that students explain their work and how it works. Silk et al. (2010) conjecture that by properly designing a robotic activity, students who lack interest or have strong negative feelings toward a topic can be motivated to learn the topic using a robotics activity, as long as the activity is properly designed.

CONCLUSION

In 1990, Lemerise (1990) contemplated whether Papert’s (1980) Turtle geometry would survive in our educational system and if so, what would it take to assure its survival? Would it have to compromise its spirit and conform to the traditional school system? Or could it survive while preserving its spirit? Unfortunately, for those of us who like Turtle geometry, the technology did not survive over time in its original form. However, its spirit did survive through Lego robotics in afterschool programs, and more recently in technology classrooms. Today, we are still wondering: What is it exactly that students are learning or can learn by playing with robots? Can we use robots to promote or enhance the learning of school topics?
The conversations about the potential of robotics are going beyond the perception that a robot is an “object-to-think” with (Papert, 1980). Humanoids are being developed to serve as peer tutors and teacher assistants. Whether the robot is used as an “object-to-think-with” as Papert envisioned, or as a humanoid that is assisting the educator, the cases presented in here give reason to believe that what we have learned in the classroom regarding teaching and learning, should be taken into consideration when designing lesson plans and robotics technology.

Students’ emotions that were present in the technology-free classroom still seem to be present in technologically enhanced classrooms. Students still react positively to a caring teacher (whether human or humanoid); there may still be negative feelings toward learning certain topics which become an obstacle in the learning process; and there may still be gender differences, etc. Therefore, for robots to help educators improve the learning experiences of students, designers need to be aware of the advances in psychology, cognitive science, neuroscience, and their interdependencies with emotions. In the entertainment industry, they are already taking into consideration the emotional basis for HRI, to assure that their products are enjoyed over long periods of time (Arkin, Fujita, Takagi, & Hasegawa, 2003). This should also be the goal of the educational community.

REFERENCES


CHAPTER 10

Virtual Avatar as an Emotional Scaffolding Strategy to Promote Interest in Online Learning Environment

Sanghoon Park
Instructional Technology Program, College of Education, University of South Florida, Tampa, Florida, USA

INTRODUCTION

Online distance learning has been highlighted as a reliable alternative to face-to-face education (Brady, Holcomb, & Smith, 2010; Veletsianos & Navarrete, 2012). According to Allen and Seaman (2010), one in four undergraduate students in the US higher education institutions takes at least one online course during their academic programs. In US higher education, the issue of “student engagement” has been an indicator of educational quality (Sener, 2007). Although online learning allows students to have greater control over their learning in regards to time and place, as well as management of their learning process (Bai, 2003), there are still students who experience frustration and even drop out of online courses due to the physical separation from instructor and peers (Naidu, 1994). Previous studies have reported higher attrition rates in online classes than in traditional classes (Thompson, 1998), which leads to the non completion of course requirements and results in academic non-success (Phipps & Merisotis, 1999; Ridley & Sammour, 1996). According to Garland (1993), the dispositional barrier (e.g. personal problems that influence the student’s persistence behavior, such as motivation) has proved to be one of the primary causes of attrition in distance education.

Recent research on academic emotion has gained its attention in supporting the dispositional barrier in a learning context. Studies conducted within a self-determination framework posit that external supports are indeed necessary for students to internalize the value of engaging in activities that are particularly compelling or intrinsically interesting (Deci & Ryan, 1985; Reeve, 2002; Ryan & Deci, 2000). Pekrun (2006), based on his
control–value theory, provides a framework of emotions related to achievement activities by emphasizing subjective control and values over learning activities/outcomes. Both self-determination and control–value frameworks imply that educators need to attempt to influence students’ emotions by modifying their perceived value of the task and further assist them to develop self–regulation skills and emotional control skills (Goetz, Pekrun, Hall, & Haag, 2006). Often, online learning environments allow students to learn at their own pace when they read learning materials or interact with course content. However, most students do not have properly developed self-regulation strategies (Azevedo & Cromley, 2004; Graesser & McNamara, 2010). Consequently, they need to be guided through substantial cognitive and affective scaffolding to be productively involved in complex learning environments (Graesser & D’Mello, 2011).

In the classroom, teachers and peer students serve as such external supports. In online classes, however, there is a great need to provide online students with additional support to help them initiate and persist in maintaining the necessary positive emotions during the course. In Passig (2001), students participating in various distance learning courses reported that there is a greater need to experience emotional scaffolding in a collaborative manner. With the recent popularity of online learning in higher education, the need to support online learners’ academic emotions has never been greater.

In this chapter, a virtual avatar (VA, also known as a pedagogical agent) is suggested as a strategy to provide online learners with emotional scaffolding, specifically to promote interest in an online learning environment. First, the concepts of the VA and emotional scaffolding are discussed. This is followed by two design considerations for emotional scaffolding using Vas: visual persona and verbal messages. Lastly, a case study implementing two VAs that are designed to provide online learners with emotional scaffolding, especially to promote learning interest, is described. This chapter is expected to have value for both researchers and practitioners in the fields of instructional technology, human computer interaction, and VAs/pedagogical agents.

EMOTIONAL SCAFFOLDING AND VAs

Emotional Scaffolding

Scaffolding has been a popular research topic in education over recent decades. In its original definition, Wood, Bruner, and Ross (1976) adopted the sociocultural theory of Vygotsky to explain how adults help children solve problem activities. They focused on one of the key aspects of children’s
learning when guided by adults (skilled others) and found that adults help children concentrate on the task elements that are within his/her competency and support their efforts eventually until they gain sufficient skills (Belland, Kim, & Hannafin, 2013). In learning contexts, scaffolding is often used as a metaphor to explain the adults’ temporary support provided for children to complete a certain element of problem activity (Van de Pol, Volman, & Beishuizen, 2010). The emotional dimension of scaffolding was not clearly recognized in the original definition, although it was inherent in contingency management and frustration control.

Emotional scaffolding was then introduced and conceptualized by Rosiek (2003) based on a series of collaborative projects at the Standford Teacher Education Program. The focus of the project was to gather insights on how teachers provide scaffolding for student learning by documenting pedagogical representations that they use. Rosiek (2003) found that scaffolding designed to guide students’ emotional response emerged in considerable discussions and called it “emotional scaffolding.” He then defined emotional scaffolding as the way teachers make use of analogies, metaphors, and narratives to influence students’ emotional response to certain aspects of the subject matter, so that student learning is promoted (Rosiek, 2003; Rosiek & Begetto, 2009). Rosiek (2003) viewed emotional scaffolding as an emotional dimension of pedagogical content knowledge. He stated that “… emotional scaffolding requires teachers to have clear knowledge of their subject matter and knowledge about the various influences on students’ emotional experience of the subject matter” (Rosiek, 2003, p. 406).

With respect to distinct emotional scaffolding approaches, Rosiek (2003) claimed two patterns were found from the study: (1) the type of emotion teachers are concerned with; and (2) the way teachers chose to address students’ emotional experience. He differentiated emotions into “unconstructive emotions” and “constructive emotions.” Unconstructive emotions refer to emotions that distract students from the subject matter content or in some other way, inhibit their learning, whereas constructive emotions serve to focus student attention more closely on the salient aspects of the subject matter being taught. He further explained emotional scaffolding approaches as shown in Table 10.1.

Through the emotional scaffolding approaches described in Table 10.1, students learning-associated emotions can be intensified, or emotions that inhibit students learning can be minimized. This chapter uses those four approaches of emotional scaffolding as a framework to address how to design and integrate VAs to promote learners’ interest in online learning.
Virtual Avatars

VAs, also known as an animated pedagogical agents, are defined as animated life-like characters designed to facilitate learning in computer-mediated or online learning environments (Johnson, Rickel, & Lester, 2000). Many studies have reported positive effects on learner’s attitudes toward learning and performance when VAs are embedded in online learning (Baylor, 2002a, 2002b; Baylor & Ryu, 2003; Moreno, Mayer, Spires, & Lester, 2001). VAs stimulate learners’ interest and academic emotion, and further promote learning. VAs are particularly beneficial in providing emotional scaffolding to learners in that the learners can maintain particular emotional responses to the subject-matter content during the learning process.

Figure 10.1 shows how the VA-provided scaffolding differs from the instructor-provided scaffolding in online learning. Students learn through interacting with online learning contents (texts, images, animations, videos, interactive games, etc.) and through participating in online course activities (discussions, quizzes, hands-on activities). Students are also engaged in individual or group interactions through synchronous or asynchronous communication. In the instructor-provided scaffolding, generally the course instructor is responsible for providing students with emotional scaffolding by means of asynchronous communications. For example, when a student

<table>
<thead>
<tr>
<th>Table 10.1 Implicit and explicit emotional scaffolding (Rosiek, 2003)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach to emotional scaffolding</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Implicit</td>
</tr>
<tr>
<td>Explicit</td>
</tr>
</tbody>
</table>
shares his/her anxiety or frustration on a certain part of the online course, the course instructor provides possible interventions to reduce students’ negative emotions. If a student reports pride, satisfaction, or gratitude, the course instructor further reinforces the positive emotions. The drawback with this approach is that emotional scaffolding is often not provided in time when a student experiences unconstructive emotions because the student has to wait until the course instructor responds to his/her emotional states with emotional scaffolding strategies. If emotional scaffolding is provided by VAs that are present in each of the course contents and each of the course activities throughout the course, students will be able to foster constructive emotions by coping with unconstructive emotions while completing the course requirements.

**Interest Development in Online Learning**

When individuals are intrinsically motivated, they participate in activity because of learning interest and activity enjoyment (Eccles & Wigfield, 2002). Interest has been conceived as an important part in the learning process because it affects the use of specific learning strategies and the way of allocating one’s attention, as well as promoting one’s emotional engagement in a task, and the extent to which one engages in deeper processing (Hidi, 1990; Schiefele, 1996, 1999; Schraw, 1998; Wade, Schraw, Buxton, & Hayes, 1993).

According to Krapp (2002), interest is a relational construct that consists of a more or less enduring relationship between a person and an object. This relationship is recognized by specific activities that may comprise concrete actions and abstract mental operations. Hence, the meaning of interest may range from a single, situation-specific person-object relation (called situational interest) toward the development of value beliefs in particular domains (called individual interest) (Figure 10.2).
**Individual Interest**

Individual interest is topic-specific; it persists over time (Hidi, 1990; Krapp, Hidi, & Renninger, 1992) and incorporates relatively enduring preferences for different topics, tasks, or contexts (Krapp, 1999; Tobias, 1994). Schiefele (1991, 1999) suggested two subcomponents of individual interest; a feeling-related valence and a value-related valence, as illustrated in Figure 10.3.

Feeling-related valences refer to the feelings that are associated with an object or an activity, such as involvement, stimulation, and flow. It occurs when an individual experiences positive affect and emotions in conjunction with a particular topic or activity. It is presumed that these positive feelings provide a strong motivational incentive to engage in an activity (Schraw & Lehman, 2001). Value-related valences refer to the attribution of personal significance of importance to an object and activity. It is presumed that value-related interest promotes engagement because an activity or body of knowledge is judged to be salient to one’s goal (Sansone, Weir, Harpster, & Morgan, 1992; Schraw & Lehman, 2001). Those two aspects are highly correlated with each other, and it is hard to determine what aspect contributes more to individual interest (Schiefele, 1991, 1999). Students prefer to learn a certain topic or participate in a certain activity because they

![Figure 10.2 Three approaches to “interest” research (Krapp et al., 1992).](image)

![Figure 10.3 Individual interest and situational interest.](image)
intrinsically like it or because they think the activity is worth doing to achieve their future goals.

**Situational Interest**

Situational interest is elicited by aspects of an object or a situation, such as novelty or intensity, or by the presence of interest-inducing factors contributing to the attractiveness of the situation (Krapp, 1999; Tobias, 1994). Most of the research on situational interest has focused on the characteristics of academic tasks that create interest (Hidi & Baird, 1986). The features that have been found to arouse situational interest and promote text comprehension and recall are: personal relevance, novelty, activity level, and comprehensibility. Situational interest is often considered to precede and facilitate the development of individual interest (Krapp et al., 1992).

Regarding situational interest, the differentiated concepts of cognitive interest and emotional interest were proposed by Kintsch (1980). Cognitive interest adjuncts, such as explanatory summaries, influence learner’s cognition by promoting the learner’s structural understanding of the explanation. On the other hand, emotional interest is believed to occur when adding interesting, but peripherally relevant material to a lesson. Learners are expected to be energized and pay more attention to learning (Harp & Mayer, 1997). As cognitive interest and emotional interest are derived by specific features of a learning material, both need to be counted as two aspects of situational interest, as shown in Figure 10.3.

Building upon previously conducted research on interest and development, Hidi and Renninger (2006) proposed a four-phase model of interest development that describes phases of situational interest and individual interest in terms of affective and cognitive processes. The four-phase model provides a rationale for identifying early phases of interest development in terms of affect or liking. The model offers a description of each phase, information about the type of support that a person in each phase of interest typically needs, and possible ways to design educational or instructional conditions to support interest development from situational interest to individual interest. The four phases consists of: Phase 1: Triggered situational interest; Phase 2: Maintained situational interest; Phase 3: Emerging individual interest; and Phase 4: Well-developed individual interest. Suggested support descriptions for interest development in each phase of the model offer implications for designing VAs’ emotional scaffolding strategies in online learning. Table 10.2 presents summaries of each phase condition for learning environment design and conditions for VA scaffolding design.
Table 10.2 Four phases of interest development, instructional design, emotional scaffolding using VAs

<table>
<thead>
<tr>
<th>Interest development phase (Hidi &amp; Renninger, 2006)</th>
<th>Definition</th>
<th>Instructional design</th>
<th>Implicit and explicit emotional scaffolding strategies using VA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase 1:</strong> Triggered situational interest</td>
<td>Psychological state of interest that results from short-term changes in affective and cognitive processing</td>
<td>Design instructional activities that trigger situational interest (group work, puzzles, computer activities, etc.)</td>
<td>Provide scaffolding strategies that emphasize environmental or text features, such as incongruous, surprising information; character identification or personal relevance; and intensity</td>
</tr>
<tr>
<td><strong>Phase 2:</strong> Maintained situational interest</td>
<td>Psychological state of interest that is subsequent to a triggered state, involves focused attention and persistence over an extended episode in time</td>
<td>Design instructional activities that provide meaningful and personally involving activities (project-based learning, cooperative group work, and one-on-one tutoring)</td>
<td>Provide scaffolding strategies that hold and sustain meaningfulness of tasks and/or personal involvement (situational interest-inducing strategies)</td>
</tr>
<tr>
<td><strong>Phase 3:</strong> Emerging individual interest</td>
<td>Psychological state of interest that is the beginning phase of a relatively enduring predisposition to seek repeated re-engagement with particular classes of content over time</td>
<td>Design instructional activities that share ideas and viewpoints from peers or experts, that offer challenging tasks related to students emerging individual interest with opportunity to re-engage them with a choice</td>
<td>Provide scaffolding strategies that help maintain students’ self-generated curiosity; that provide support to help students increase understanding of the content; that offer encouragement to persevere when confronted with difficulty</td>
</tr>
<tr>
<td><strong>Phase 4:</strong> Well-developed individual interest</td>
<td>Psychological state of interest that is a relatively enduring predisposition to re-engage with particular classes of content overtime</td>
<td>Design instructional activities that facilitate the well-developed individual interest by providing interaction and challenge which leads to knowledge building</td>
<td>Provide scaffolding strategies that help students generate and seek answers to their own curiosity questions; that support long-term constructive and creative endeavors; that offer deeper levels of explanation for tasks; that help preserve work, even in frustration</td>
</tr>
</tbody>
</table>
In online learning, where VAs are present, students’ interest is affected by VA design considerations within the four-phases of interest development. In the next section, such VA design considerations are discussed.

**VA DESIGN AS AN EMOTIONAL SCAFFOLDING STRATEGY**

VAs can be designed by employing two communication modes; visual persona and verbal messages. The key characteristics that determine a VA’s persona include its propensity to be engaging, person-like, credible, and instructor-like (Baylor & Ryu, 2003). Creating such a persona is as important as designing a VA’s appearance. VAs also communicate with learners via verbal messages to facilitate academic emotion and learning. The content of messages could be directly related to the learning content and delivered in the form of visual representations and/or narratives of subject-matter concepts to foster learners’ emotional reaction to the content.

**VA Persona**

Engaging persona of a VA facilitates the learner-VA relationship and motivates the learner to be involved in the learning task. Person-like persona of a VA forms a viable relationship with the learner. Credible persona of a VA makes learners feel confident in learning with the VA and helps them to recognize the VA as trustworthy, competent, and consistent in behavior. Finally, instructor-like persona of a VA serves as a pedagogical mentor to effectively represent the content and pedagogy (Baylor, 2000). Media features, such as voice, emotional expression, gesture, image, and animation are integrated to create such different persona.

Therefore, media features are essential in terms of constructing a VA’s different persona and indicating social presence of a VA to promote learning interest. “Voice” has been suggested as a key aspect for enhancing VAs’ persona. In fact, the voice feature has been found to be a critical element for designing VA-embedded learning. Prior research has indicated that voice has a superiority effect to visual appearance for communication in computer-based media (Mayer & Moreno, 1998; Moreno & Mayer, 1999), and with VAs in particular. While there are consistent results that voice (in conjunction with text) is a key aspect for enhancing a VA’s persona, the role of VA image and animation is not proven (Baylor & Ryu, 2003). However, given the importance of emotional expression by a VA, animation would be necessary for a VA to demonstrate facial expressions. Animation with emotional expression increases the VA’s persona of credibility and
also improves social presence of the VA by increasing non verbal behavior. Baylor (2003) found that animation provides the most positive impact for a VA to be perceived as engaging, because it contributes to the VA’s expression of a personality through non verbal behavior, leading it to be more likable, and thus more enjoyable to learn with, and also more emotionally expressive.

**VA Message**

According to Deci and Ryan (1991), emotion regulation involves learning to reflectively interpret incoming stimuli in more integrated ways so that one can make a decision with respect to the behavior with full awareness of the goals and values relevant to it. Pekrun (2011) contended that emotions can be regulated and changed by addressing any of the elements involved in the cyclic feedback processes between emotions, their effects, and their antecedents. According to Pekrun (2011), emotional regulation can be achieved by targeting the emotion itself, the control and value appraisals underlying emotions, the competences determining individual agency, and tasks and learning environments. For example, if a student feels confused or frustrated while working on a learning task, it would be gradually changed to anxiety if proper reappraisal process does not occur. By providing external support to influence the reappraisal process, the student can re-examine his/her perception of control and value over the learning task and begin to regulate emotions into positive ones so that he/she can concentrate on the learning task again. Indeed, online learners’ negative emotions were found to be mainly associated with the given workload, the time demanded, the level of difficulty presented in a course task, and a sense of isolation and loneliness (Zembylasa, Theodoroub, & Pavlakis, 2008).

VAs deliver messages to convey information. The information delivered by a VA can be designed to scaffold the learners’ perception of the situation or perceived value and feeling of the situation, and further changes their learning interest. Therefore, VAs delivering messages must be designed, based upon the four-phase model of interest development, so that learners’ interest is induced differently depending upon the learners’ interest development stage. Specifically, if learners are in the Phases 1 and 2 of the interest development model, situational interest-inducing messages would stimulate the learners’ learning interest directly targeting the emotion itself. If learners are in the Phases 3 and 4, individual interest-inducing messages targeting the learners’ control and value appraisals would be preferable.
An example of such situational interest-inducing messages can be found from the distinction between cognitive interest and emotional interest. According to Kintsch (1980), interest can be promoted in two different ways by adding cognitive interest or emotional interest. On the one hand, the cognitive interest influences learner’s cognition by promoting the structural understanding of the content. On the other hand, emotional interest is increased by adding interesting, but irrelevant content to instructional material. Emotional interest does not help understand the content, yet it energizes a learner’s arousal, so that they pay more attention and learn more overall (Harp & Mayer, 1997). Researchers have used the term “seductive detail” to refer to interesting but irrelevant points that are added to a passage to make it more interesting in the reading education field. In multimedia learning, seductive detail is often called “seductive augmentation.” This term refers to not only text but also graphics, narratives, voice, animation, and text accompanied in a multimedia learning environment, with the purpose of increasing learner’s situational interest.

In summary, both a VA’s persona and the message information delivered by the VA play a critical role in promoting learning interest. Media features are important factors to create the VA’s persona, while the information type delivered by VAs is critical in determining the target of emotional regulation.

Virtual Agent Scaffolding Model for Interest Development

To describe the common characteristics of scaffolding, Van de Pol et al. (2010) presented a conceptual model of scaffolding, emphasizing three components: Contingency, fading, and transfer of responsibility. Contingency is a tailored support that the teacher adaptively provides to adjust the current level of the students’ performance at the same or a higher level. According to Van de Pol et al. (2010), the students’ current level of competence needs to be determined prior to offering proper contingent support. Fading is the gradual withdrawal of such support, depending on the students’ level of competence in completing a task. The amount of scaffolding is decreased over time. Consequently, the responsibility for the performance of a task is transferred to students when their level of control over the task increases. In this model, scaffolding is viewed as an interactively participating process that occurs between teacher and student (Stone, 1998a, 1998b). Applying his conceptual model of scaffolding into the four-phase development model of interest, we can think of four different phases of scaffolding and each phase has its unique scaffolding purpose (Figure 10.4).
In Phases 1 and 2, where a student has not developed any level of interest in learning yet, the goal of emotional scaffolding must be on triggering and maintaining situational interest by providing stimuli, such as cognitive interest inducing messages or emotional interest inducing messages. As the student develops and internalizes interest, the goal of the emotional scaffolding needs to be shifted to affect the students’ perception of feeling and value of learning tasks, so that individual interest emerges. While progressing from Phases 1 to 4, the amount of scaffolding for interest is gradually reduced, while the amount of learners’ situational interest and individual interest is increased. When interest is finally internalized in Phase 4, no more emotional scaffolding will be necessary.

**THE CASE STUDY**

**Aim**

In order to empirically examine the VA scaffolding model for interest development, a simple case study was conducted. In this study, the first phase of interest development, triggering situational interest, was tested by designing and implementing different types of interest-inducing scaffolding through VAs. The main research question leading this case study was, “What is the effect of interest-inducing scaffolding by VAs on students’ perceived learning interest?” This study hypothesized that the mean score of learning
interest for the VA conditions as a source of interest-inducing messages is significantly higher than the scores for the non-VA conditions.

Participants

A total of 127 college undergraduate students enrolled in “Computer literacy” classes in a large public university in the southeastern United States took part in this case study. This course was one of the required courses for undergraduate students. All participants were recruited from 10 sections of the course and offered extra credits as compensation. A total of 136 students voluntarily participated at the beginning of the study. Of these, 127 participants were included in the final data analyses because 9 students did not complete the post questionnaire. The average age of the sample was 19.72 years (SD = 1.96). Among the 127 participants, 63.0% were Caucasian; 19.7% were African-American; 11.0% were Hispanic/Latino; 2.4% were Asian/Asian-American; 1.6% were bi-racial; and 2.3% were other ethnicity groups. There were 60.6% male students and 39.4% female students. The majority of the participants were sophomores (43.3%), with 21.3% freshmen, 20.5% juniors, and 15.0% seniors. The required sample size was estimated by using pre-determined alpha, effect size, and power. The sample size was determined using Cohen (1988), with a power level of 0.9, and large effect size of 0.50 at an α-level of 0.05.

Study Design

This study employed a randomized post-test only experimental design. In order to explore the research question, participants were randomly assigned into one of the four conditions (A-D), based on the sequence of their entry to the research lab, as shown in Table 10.3. The independent variable in this study was the source of interest-inducing messages. Therefore, four study conditions were created. Group A received interest-inducing messages from

Table 10.3 Research design

<table>
<thead>
<tr>
<th>Group</th>
<th>Random assignment</th>
<th>Group</th>
<th>Treatment</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer literacy class students</td>
<td>R</td>
<td>A</td>
<td>X₁</td>
<td>O₁</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>X₂</td>
<td>O₂</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>X₃</td>
<td>O₃</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D</td>
<td>X₄</td>
<td>O₄</td>
</tr>
</tbody>
</table>

1. Companion role of VA; 2, instructor role of VA; 3, text message without VA; 4, no message.

*aIndependent variable: the source of interest-inducing messages.*
a companion role of VA; Group B received interest-inducing messages from an instructor role of VA; and Group C received interest-inducing messages in the form of text without a VA. Last, group D received no interest-inducing messages.

VAs with two personas, one instructor VA and the other companion VA, were used in this study as shown in Figure 10.5. The roles of VA were designed through the general persona represented by image, voice, animation, and affect. The companion role of VA named “Mike” was designed with a friendly and energetic voice, youthful appearance, and expressive emotions. The instructor role of VA named “Dr. Handricks” was designed with a dry and straightforward voice with little inflection, older appearance, little animation, and informative and directive characteristics. Figures 10.6–10.9 show the four experimental conditions used in this study.

Figure 10.5 The example of a companion VA and an instructor VA.

Figure 10.6 Interest-inducing message delivered by the companion VA.
Measures

Learning interest in this study was defined as the psychological status within a person as a combination of individual interest and situational interest. Therefore, learning interest was measured in terms of individual interest and situational interest. Prior to beginning the lesson, participants were informed
that the topic of the instructional material is “intellectual property.” Then, the level of individual interest was measured from two aspects: feeling-related interest and value-related interest. To measure feeling-related interest, the participants indicated the feeling they expected to have while studying the instructional material. A 5-point Likert scale was provided with response choices ranging from “not at all true” to “very true.” The participants were asked to read the following five adjectives in estimating their expected feelings: (“While studying the instructional material on ‘intellectual property’/I expect to feel …”)—“bored,” “stimulated,” “interested,” “indifferent,” and “engaged.” The reliability of the feeling-related interest survey for this study was 0.74. In order to measure the value-related aspect of individual interest, participants were asked to read the terms “meaningful,” “unimportant,” “useful,” “worthless,” and again asked to describe the value of the topic “intellectual property” to them personally. A 5-point Likert scale was again provided with response choices ranging from “not at all true” to “very true.” For each participant, a score of individual interest was computed by adding feeling-related interest scales and value-related interest scales. Previous studies have shown that this measure of interest is unidimensional and highly reliable (Schiefele, 1996, 1998; Schiefele & Krapp, 1996). The reliability of the value-related interest survey for this study was 0.75.

Situational interest was measured from three aspects: arousal, involvement, and attention. In order to measure arousal, five items from the activation-deactivation adjective check list (AD-ACL) were used (Thayer, 1985, 1986). Only the energetic dimension of AD-ACL was selected in this study. The reliability of the arousal measurement for this study was 0.91. Participants’ involvement was measured using two items: “I was completely caught up in what I was studying,” and “When learning from the material, I was concentrated.” A 5-point Likert scale was provided with response choices ranging from “not at all true” to “very true.” In order to measure attention, Keller’s IMMS (Instructional Material Motivation Survey) was employed (Keller, 1993). The IMMS was intended to be a situational measure of students’ motivational reaction to the lesson. The original statements of the attention measurement were adapted to reflect the context of this study. For example, an original item, “There was something interesting at the beginning of this lesson that got my attention” was revised to “I found something interesting at the beginning of this instructional material that got my attention.” The reliability of the original attention measurement was 0.89 (Keller, 1993), but the response reliability of the survey for this study was 0.83.
Study Material

Simple web-based course material was designed for this study. The topic of the course material was “Introduction to intellectual property,” consisting of three subconcepts of intellectual property: patents, trademarks, and copyright. The topic of intellectual property was selected in this study for several reasons. First, the topic “intellectual property” had been one of the topics in the target class. Therefore, students were aware of the topic, but not familiar with detailed information, specifically regarding the areas of patents, trademarks, and copyright. Second, as the primary goal of this study was to investigate the effect of interest-inducing messages by VAs, instructional text for which students presumably had low levels of prior knowledge was needed. Third, this topic was related to the students’ attitude in everyday life, regarding how to use computer applications without violating any legal and ethical issues. Instructional material on intellectual property consisted of three learning phases:

- **Introduction phase**: Students were given a brief introduction about intellectual property and basic information containing history, related regulations, and examples.
- **Learning phase**: Students were given a detailed explanation in regard to the three subconcepts: patents, trademarks, and copyright.
- **Test phase**: After completing instructional material, students were given an opportunity to actually test what they had read in the instructional material.

For this particular case study, the lesson material was developed in a web-based format that can be easily inserted into online courses. Students were forced to go through all screens in the order, from the introduction phase to the test phase. Only navigation moving onto the next screen was allowed in order to verify that students were not given a second chance to refer to the text information.

Procedure

Participants were guided to log in to the computer and open the web-based course material. Completing the online consent form, participants studied the material in their normal studying speed. Each participant was presented with the instructional material corresponding to his/her treatment group and told to begin the material. Participants were not allowed to take notes or refer to other resources. When individual participants finished the instructional material, they were guided to the post-measure instruments.
on screen. We ensured that no participants were distracted by other computer activities, such as instant messenger programs, playing online games, or e-mail checking, etc.

Data Analysis

The learning interest data were analyzed by two sets of measures: individual interest and situational interest. Individual interest data were collected prior to the main experiment to be compared among conditions as a prior interest. The scores for individual interest were added to the scores for situational interest that were collected after the main study and together with constructed learning interest scores. Prior to the main data analysis, the equivalence of treatment groups in terms of pre-interest and prior knowledge was verified. Additionally, a missing value analysis, a case analysis, and a detection of violations of assumptions for the dataset were conducted. The descriptive statistics for each of the interest measures are presented in Table 10.4.

An ANOVA indicated that significant differences occurred in arousal, $F(3, 119) = 4.98, p < 0.01, \eta^2 = 0.11$, and attention, $F(3, 119) = 7.01, p < 0.001, \eta^2 = 0.15$. To assess pairwise differences among the four levels for the main effect for arousal, the Tukey HSD (honestly significant difference) follow-up procedure ($\alpha = 0.05$) was used to provide additional family-wise protection. The result indicated that arousal score for the companion role of VA ($M = 2.19, SD = 0.57$) was significantly higher than for both the text only condition ($M = 1.76, SD = 0.61$) and the no message condition ($M = 1.77, SD = 0.60$). Also, arousal score for the instructor role of VA ($M = 2.19, SD = 0.68$) was significantly higher than for both the text only condition ($M = 1.76, SD = 0.61$) and the no message condition ($M = 1.77, SD = 0.60$). For attention, pairwise differences among the four levels for the main effect were assessed using the Tukey HSD follow-up procedure ($\alpha = 0.05$). The result indicated that the attention score for the companion role of VA ($M = 3.30, SD = 0.65$) was significantly higher than for both the text only condition ($M = 2.90, SD = 0.47$) and the no message condition ($M = 2.76, SD = 0.66$). Also, attention score for the instructor role of VA ($M = 3.30, SD = 0.59$) was significantly higher than both the text only condition ($M = 2.90, SD = 0.47$) and the no message condition ($M = 2.76, SD = 0.66$). The attention score for text only condition ($M = 2.90, SD = 0.47$) was higher than no message condition ($M = 2.76, SD = 0.66$) but the difference was not statistically significant. The results are summarized in Table 10.5.
Table 10.4 Means and standard deviations for learning interest by conditions

<table>
<thead>
<tr>
<th>The source of interest-inducing messages</th>
<th>Companion VA (n = 30)</th>
<th>Instructor VA (n = 32)</th>
<th>Text only (n = 31)</th>
<th>No messages (n = 34)</th>
<th>Total (n = 127)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest measures</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Feeling-related</td>
<td>2.96</td>
<td>0.69</td>
<td>2.96</td>
<td>0.70</td>
<td>2.82</td>
</tr>
<tr>
<td>Value-related</td>
<td>3.90</td>
<td>0.64</td>
<td>3.96</td>
<td>0.75</td>
<td>3.93</td>
</tr>
<tr>
<td>Arousal</td>
<td>2.19</td>
<td>0.57</td>
<td>2.19</td>
<td>0.68</td>
<td>1.76</td>
</tr>
<tr>
<td>Involvement</td>
<td>2.67</td>
<td>0.94</td>
<td>2.50</td>
<td>0.92</td>
<td>2.18</td>
</tr>
<tr>
<td>Attention</td>
<td>3.30</td>
<td>0.65</td>
<td>3.30</td>
<td>0.59</td>
<td>2.90</td>
</tr>
</tbody>
</table>

Note. The possible score range for learning interest was 1-5.
In order to examine the difference between both VA conditions and the text only condition, a contrast analysis was performed. A contrast of 0.5 was assigned to the companion role of the VA condition and also to the instructor role of the VA condition. And $C_0$ was assigned to the text only condition. Consistent with the ANOVA analysis, students who received interest-inducing messages from the companion role of VA and the instructor role of VA rated the learning interest score significantly higher than did students in the text only condition in terms of arousal score, $t(90) = 3.13, p < 0.01$; attention score, $t(90) = 3.15, p < 0.01$. Therefore, the study hypothesis was supported by the findings.

Findings

The students who were given interest-inducing messages delivered by the VA showed significantly higher arousal and attention, in terms of learning interest than the students who were given either the text type of interest-inducing message or no message. Students in the control group in which no messages were provided showed lower interest than those in the other study conditions. An interesting finding to note is that there was significant difference of the arousal score and attention score between the two VA delivered interest-inducing message groups and text-based interest-inducing message group. The result of contrast analysis confirmed that interest-inducing messages are more effective to promote students’ interest when a VA presents and delivers the message. The findings are explained below.

First, the presence of a VA could affect students’ interest. Previous research on the effects of VA showed that users’ interactions with the computer are much smoother when likeable animated VAs are present. VAs also help students develop an emotional connection and facilitate their

<table>
<thead>
<tr>
<th>Variable</th>
<th>Feeling-related interest</th>
<th>Value-related interest</th>
<th>Arousal</th>
<th>Involvement</th>
<th>Attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest-inducing messages</td>
<td>$F^b = 0.66$</td>
<td>$F^b = 0.37$</td>
<td>$F = 4.98^*$</td>
<td>$F = 1.83$</td>
<td>$F = 7.01^{**}$</td>
</tr>
</tbody>
</table>

A Bonferroni adjusted $\alpha$-level $= 0.01$ was used.

*ANOVA, univariate analysis of variance.

$^b$Univariate df for the source of interest-inducing messages $= 3119$.

*$p < 0.01$.

$^{**}p < 0.001$. 

Table 10.5 Univariate analysis of variance $F$ ratios for learning interest

220 Emotions, Technology, Design, and Learning
enjoyment of the learning situation (Dehn & Mulken, 2000). Therefore, although identical interest-inducing messages were delivered in both VA conditions and the text only condition, students with VAs perhaps felt higher interest in the learning content. A second reason to possibly explain the findings is that the source of the interest-inducing messages was different between the VA conditions and the text only condition. In other words, interest-inducing messages were delivered by a VA via audio voice, while the same messages were delivered in text. Therefore, it is possible that students perceived the audio voice more user-friendly than the text, based on Mayer’s multimedia principle and voice principle (Mayer, 2009).

CONCLUSION

This chapter has presented a case of emotional scaffolding by VAs, in which the theoretical concepts of interest and the interest development model provide implications for the scaffolding strategies. Also emphasized was that VA needs to be designed considering “persona” and “delivering messages” to offer such emotional scaffolding strategies. A case study was conducted to empirically test the effectiveness of the VA delivering interest-inducing messages, as a strategy of emotional scaffolding, and the results showed that interest-inducing messages increased students’ situational interest, especially when accompanying the presence of VAs. However, no situational difference was found between the two different roles (persona) of the VAs.

At the same time, it should be noted that the case study focused only on the first phase of the interest development model, which is triggering situational interest. Therefore, future studies will be needed to examine the other three phases of interest development and corresponding emotional scaffolding strategies using VAs. Especially, how to design VAs’ persona and verbal messages differently in each phase to promote interest development in online learning will be the main agenda of future studies. In this study, the distinction between emotional interest and cognitive interest served as a source of VA persona and verbal message design for Phase 1. As learners develop interest into different phases, VA persona and verbal messages will need to be designed accordingly.

Also addressed here were theoretical foundations and practical design issues that only involve developing emotional scaffolding strategies in online learning with VAs. A comprehensive model of scaffolding will need to be explored to include cognitive, metacognitive, and motivational scaffolding strategies, so that online learners’ positive emotion experience can be holistically supported.
REFERENCES


Virtual avatar to promote interest in online learning


Veletsianos, G., & Navarrete, C. (2012). Online social networks as formal learning environments: Learner experiences and activities. The International Review of Research in Open and Distance Learning, 13(1), 144–166.


CHAPTER 11

Animated Pedagogical Agents and Emotion

Enilda Romero-Hall
University of Tampa, Tampa, Florida, USA

INTRODUCTION

As cognitive, developmental, and educational psychologists have continued to contextualize their inquiry within the education system, they have found that emotions are an integral part of education (Schutz & Lanehart, 2002). Due to their findings, researchers interested in teaching, learning, and motivation transactions cannot ignore emotional activity settings (Eynde & Turner, 2006; Meyer & Turner, 2002; Schutz & Lanehart, 2002). Craig and Rebolledo-Mendez (2009) believe that there are empirical and theoretical questions that need to be answered in order to understand the relationship between cognition and affect. The next generation of education researchers must look beyond cognition; they must investigate educational strategies tailored to restore the balance between cognition and affect (Craig & Rebolledo-Mendez, 2009; Picard, 1997).

This transition into the affective domain requires innovative approaches to construct models of the emotional dynamics of learners and the efficient use of these models to optimize face-to-face instruction, but also equally important for web- and computer-based instruction. For example, it is important to know how researchers can develop and evaluate systems to automatically detect learner-centric emotions in real-time and how computer-based learning environments can be responsive and reactive to the learner’s affect. Additionally, there are gaps in understanding the social rules and design and development elements that animated pedagogical agents serving as tutors and peer-learning companions should employ in order to synthesize affective expressions and yield more naturalistic emotion-like communication.

This chapter focuses specifically on the role, benefits, and design of animated pedagogical agents in web- and computer-based instruction. It addresses the role of emotion in the design of animated pedagogical agents and its effect in human-agent interactions in a learning environment. The
chapter also discusses recent research efforts on the effect of the emotional dimension of animated pedagogical agents on learning and how the research helps us understand which emotion elements are needed for the design of animated pedagogical agents.

WHAT ARE ANIMATED PEDAGOGICAL AGENTS?

Animated agents are life-like computerized characters designed to facilitate learning in interactive environments (Craig, Gholson, & Driscoll, 2002; Johnson, 2001; Shaw, Johnson, & Ganeshan, 1999). Agents have also been defined as computer programs that simulate a human relationship, by doing something that another person could otherwise do (Laurel, 1990; Selker, 1994). The most common agent interfaces consist of an animated face, a cartoon character, or a human-like virtual agent (Moreno, 2001). They draw their strength from the naturalness of the living-organism metaphor in terms of both cognitive accessibility and communication style (Laurel, 1990).

The idea of animated agents is not new. Over the previous decades, numerous researchers have studied the design, development, and implementation of animated pedagogical agents (Laurel, 1990; Picard, 2000; Picard, 1997; Riecken, 1994). There are varying views on the use of agents in computer-mediated environments. Even the terminology is inconsistent, with various names being used to describe an animated agent, e.g., intelligent agents, software agents, guidebots, relational agents, embodied conversational agents, and responsive virtual human, to name a few (Riecken, 1994). Agents occupy a strange place in the realm of technology, generating fear, fiction, and extravaganza (Norman, 1994). The concept of an agent, especially when modified by the terms intelligent, animated, or conversational, brings forth images of human-like androids, working without supervision on tasks thought to be for our benefit but not necessarily to our liking (Gratch et al., 2002; Norman, 1994).

THE ROLE OF ANIMATED PEDAGOGICAL AGENTS

In 1994, Selker categorized agents as assistant-style agents and advisory-style agents. The assistant-style agent builds a relationship, in which its very success creates dependency for the user. On the other hand, the advisory-style agent builds a user relationship with the explicit goal of educating the individual. Today, the use of animated agents surpasses the role of advisor or assistant. Some animated agents are support tools, such as virtual tutors,
or help aids to instruct the user to perform a certain task (Hubal, 2008; Johnson, 2001; Kim & Baylor, 2006; Laurel, 1990; Moreno, 2001; Woo, 2009). Others present the instructional content via a virtual instructor training the user (Hubal, 2008; Johnson, 2001; Kim & Baylor, 2006; Moreno, 2001). Some animated pedagogical agents play an acting role to demonstrate examples of concepts and skills, or to engage the user in conversation (Hubal, 2008; Kim & Baylor, 2006; Moreno, 2001).

Overall, the role of animated agents is to accurately model the kinds of dialogs and interactions that occur during apprenticeship learning and one-on-one tutoring (Shaw et al., 1999). Additionally, animated agents play a significant part in stimulating social interaction that can facilitate learners to engage in the learning task and, consequently, to enhance learning in computer-based environments (Kim & Baylor, 2006; Laurel, 1990). As a result, they can be integrated into a variety of interactive media, such as web-based information spaces, interactive pedagogical dramas, virtual environments, educational games, and simulations (Johnson, 2001).

**BENEFITS OF ANIMATED PEDAGOGICAL AGENTS**

Discussions about the uses of agents in computer-based instruction suggest that they provide at least three primary types of benefits to learners (Baylor, 2011; Clark & Choi, 2005). First, agents may have a positive impact on learners’ motivation and how strongly they value computer-based learning programs (Baylor, 2011; Clark & Choi, 2005). Second, animated agents might help learners focus on important elements of learning materials (Clark & Choi, 2005). Third, agents may also provide learners with context-specific learning strategies and advice (Clark & Choi, 2005; Johnson, 2001).

There are other benefits provided by animated agents. One of these benefits is that the agent can demonstrate physical tasks, such as operation and repair of equipment (Johnson, Rickel, & Lester, 2000). Demonstrating a task may be far more effective than trying to describe how to perform it, especially when the task involves spatial motor skills (Johnson et al., 2000). An interactive demonstration using an animated pedagogical agent provides students with the ability to move around in the environment and view the demonstration from different perspectives (Johnson et al., 2000). Agents can also demonstrate procedures performed by complex devices by taking on the role of an actor in a virtual process (Johnson et al., 2000).

When a student’s work environment is large and complex, animated agents are valuable as navigational guides, leading students around and
preventing them from becoming lost (Johnson, 2001; Johnson et al., 2000). Research on training using immersive virtual reality shows that students can easily become disoriented and lost in complex environments, thus making animated agents that serve as guides can help the learning process (Johnson et al., 2000). Learning environments with navigational guides can help students develop spatial models of the subject matter and the virtual environment (Johnson et al., 2000).

Another benefit of animated agents is their ability to serve as the learners’ attentional guides, using the most common and natural methods of communication: gaze and deictic gesture (Johnson et al., 2000). Because of significant advances in the capabilities of graphics technologies in the past decade, interactive learning environments increasingly incorporate visual aids (Johnson et al., 2000). Animated agents can use gaze and gestures to draw students’ attention to a specific aspect of a chart, graphic, or animation (Johnson et al., 2000). Agents can employ deictic behaviors to create context-specific references to physical objects in a virtual world (Johnson et al., 2000).

Two behavior-related benefits of animated agents are their ability to provide non-verbal feedback and conversational signals (Hubal, 2008; Johnson et al., 2000). The ability to provide non-verbal feedback in addition to verbal comments allows an animated agent to provide more varied degrees of feedback (Cassell et al., 1998; Johnson et al., 2000). For example, a simple head nod to show agreement can assure a student without interrupting them. Non-verbal feedback is less obtrusive than a verbal comment (Johnson et al., 2000).

Animated agents can also use non-verbal signals to help regulate conversations with students and complement their verbal utterances (Bickmore & Cassell, 2000; Cassell et al., 1998; Johnson et al., 2000). For example, the animated agent can use an intentional pitch accent to highlight a salient word or phrase. Although communication can happen in the absence of these non-verbal signals, it proceeds most smoothly when they are available (Bickmore & Cassell, 2000; Johnson et al., 2000). If the bodies of the animated agents are used in ways that leverage knowledge of human communication behavior, animated agents will provide a qualitative advantage over other types of interfaces (Cassell, Bickmore, Campbell, Vilhjálmsson, & Yan, 2001).

EXAMPLES OF ANIMATED PEDAGOGICAL AGENTS

Animated agents represent a new paradigm for education and training in interactive learning environments, providing a new metaphor for
human–computer interaction based on face-to-face dialogue (Andre & Rist, 2001; Bickmore & Picard, 2003; Cassell et al., 1998, 1999; Johnson, 2001; Johnson et al., 2000; Kim & Baylor, 2006). Agents with names such as Adele, Steve, Herman the Bug, Cosmo, and Rea have been developed to serve a variety of instructional goals in computer-based instruction: to facilitate tutoring system architectures, provide assistance to trainers in virtual worlds, and act as co-learners (Clark & Choi, 2005; Elen, Clarebout, & Johnson, 2002; Johnson et al., 2000; Shaw et al., 1999).

Facilitate Tutoring Systems Architecture

In case-based clinical diagnosis application, students are presented with materials on a particular medical condition and are then given a series of cases to work through. “Adele” is used to highlight interesting aspects of the case, monitor and provide feedback as the student works through a case, and give hints or rationale for particular actions (Johnson, 2001; Shaw et al., 1999). Similarly, “Cosmo” was designed to provide problem-solving advice in the Internet Protocol Advisor environment. Students interact with Cosmo as they learn about network routing mechanisms by navigating through a series of subnets. Cosmo has the ability to dynamically combine gestures, locomotion, and speech, to refer to objects in the environment while delivering problem-solving advice.

Provide Assistance in a Virtual World

“Steve” was designed to interact with students in an immersive virtual environment located in the engines aboard a US Navy surface ship (Johnson, 2001; Johnson et al., 2000). As with some other animated agents, Steve can monitor the student’s actions, point out errors, and answer questions (Johnson, 2001; Rickel et al., 2002). He can also demonstrate actions, i.e., use gaze and gestures to direct the student’s attention (Johnson, 2001; Rickel et al., 2002). Similarly, “Herman the Bug” inhabits Design-A-Plant, a virtual learning environment for the domain of botanical anatomy and physiology (Johnson et al., 2000). As the students build plants, Herman observes their actions and provides explanations and hints. Herman also performs a broad range of actions, including walking, flying, shrinking, swimming, and other acrobatics (Johnson et al., 2000).

Act as a Co-Learner

“Rea” is an embodied, multimodal real-time conversations agent (Bickmore & Cassell, 2000; Cassell, 2001; Cassell et al., 1998, 1999, 2001). It implements
the social, linguistic, and psychological conventions of conversation to make
interaction with a computer as natural as face-to-face conversation with
another person (Bickmore & Cassell, 2000; Cassell et al., 2001). Rea has a
human-like body and uses her body in human-like ways during the conversa-
tion (Bickmore & Cassell, 2000; Cassell, 2001; Cassell et al., 1999, 2001).
She uses eye gaze, body posture, hand gestures, and facial displays to organize
and regulate conversation (Bickmore & Cassell, 2000; Cassell, 2001; Cassell

DESIGN OF ANIMATED PEDAGOGICAL AGENTS

Agents can be well or poorly designed (Laurel, 1990). The aim when
designing an animated agent is to create a life-like persona (Rickel et al.,
2002; Shaw et al., 1999). An agent should also have behavior and appearance
that helps enhance the perception of expertise in the agent (Andre & Rist,
2001; Shaw et al., 1999). Veletsianos (2007) referred to the perception of
expertise provided by the agent as “contextual relevance.” The contextual
relevance can be defined as the conformity of an agent’s visual characteristics
to the content area under which the agent purports to function (Veletsianos,
2007). It is also important that the agent has enough domain knowledge
to support the anticipated instructional dialogs with the learner (Shaw
et al., 1999).

Another valuable aspect of motivational animated pedagogical agents is
the possibility of customizing the agent to represent an ideal social model for
a particular user or group of learners (Baylor, 2011). A key characteristic for
designing motivational agents is a pleasant physical appearance, which
should refer to the age, status, attractiveness, and credibility of the agent
(Andre & Rist, 2001; Baylor, 2011; Veletsianos, 2007). Animated pedagog-
ical agents should also present good character building qualities (Gratch
et al., 2002; Heckman & Wobbrock, 2000). Character building qualities
include animations, gestures, communication, and personality (Baylor,
2011; Heckman & Wobbrock, 2000; Rickel et al., 2002). Animations
and gestures, such as gaze, body posture, and tone of voice, have a deep
impact on student’s impressions of agents (Andre & Rist, 2001; Baylor,
2011; Cassell, 2001; Cassell et al., 1998; Johnson et al., 2000; Shaw et al.,
1999). Communication in animated pedagogical agents can be verbal or
non-verbal. Thus far, all agents used in computer-based instruction are tech-
nologically able to communicate both verbally and non-verbally (Cassell
et al., 1998, 1999; Clark & Choi, 2005; Elen et al., 2002; Johnson,
Most agents use speech (Elen et al., 2002). Some use both speech and text (Elen et al., 2002). The most frequently used delivery modality is dialogue, in which the agent provides explanation and complements the interaction with questioning (Elen et al., 2002). Combined, these design characteristics give the animated agents the strength needed to make the agent more believable.

**AGENT DESIGN AND EMOTION**

In addition to the previously mentioned design characteristics, it is critically important that animated pedagogical agents capture the core of our cultural representations by providing appropriately timed and clearly expressed emotions (Barbat & Cretulescu, 2003; Bates, 1994; Baylor, 2011; Heckman & Wobbrock, 2000; Rickel et al., 2002; Romero & Watson, 2010; Veletsianos, 2009). An animated agent that can effectively convey appropriate emotional responses greatly augments the illusion of life because emotions are something that we find at the heart of what it means to be human (Barbat & Cretulescu, 2003; Bates, 1994; Heckman & Wobbrock, 2000).

Animated agents that draw from a rich repertoire of emotive behaviors to exhibit contextually appropriate facial expressions and expressive gestures can exploit the visual channel to advise, encourage, and empathize with learners (Johnson et al., 2000; Lester, Towns, & Fitzgerald, 1999). An animated pedagogical agent with a rich repertoire of emotive behaviors may simply make learning more fun (Johnson et al., 2000; Lester et al., 1999). A learner that enjoys interacting with a pedagogical agent may have a more positive perception of the overall learning experience and may consequently choose to remain in the learning environment for a longer period of time (Lester et al., 1999).

An animated agent that properly conveys and elicits emotion has the ability to motivate the learner (Johnson, 2001; Johnson et al., 2000; Kim & Baylor, 2006; Rickel et al., 2002) and may encourage the student to care about their own progress (Johnson et al., 2000). Additionally, an emotive agent may convey enthusiasm for the area of the instruction and therefore foster similar levels of enthusiasm in the learner (Andre & Rist, 2001; Hubal, 2008; Johnson et al., 2000; Rickel et al., 2002). Overall, the quality, clarity, and dramatic impact of communication can be increased through the creation of emotive movements that underscore the affective content of the message (Johnson et al., 2000). Therefore, as animation of emotions becomes more sophisticated, animated agents are better positioned to improve student’s motivation (Johnson et al., 2000).
Some assume that something so central to humanity, emotions, cannot be replicated due to their complexity (Lester et al., 1999). But it is not emotion that must be replicated in animated agents; it is the appearance of emotion (Lester et al., 1999). By carefully orchestrating facial expression, full-body behaviors, arm movement, and hand gestures, animated agents can be designed to visually augment their emotive communication (Lester et al., 1999).

**RESEARCH EVIDENCE**

Emotional behaviors can create a sense of empathy and drama and fill the agents with a rich mental life (Gratch et al., 2002; Rickel et al., 2002). The growth in animated pedagogical agents with emotion-like qualities in computer-based instruction builds on the theory that entertainment value translates into greater student enthusiasm for instruction and better learning (Rickel et al., 2002). But beyond creating a sense of engagement, emotion also appears to play a central role in teaching (Rickel et al., 2002). As a result, research on emotion models has increased in recent years (Rickel et al., 2002). This work is motivated by psychological theories of emotion that emphasize the relationship between emotions, cognition, and behaviors (Gratch et al., 2002; Rickel et al., 2002).

Bickmore and Picard (2004) conducted studies (Bickmore & Picard, 2005) based on psychological theories of emotions that focused on the human-agent relationships using a caring and relational animated agent. According to Bickmore and Picard (2004), feeling that one is cared for has profound effects on physiological cognition and emotional state in humans. Caring is expressed not only through speech content of the communication with the animated agent, but through non-verbal and paraverbal modalities including facial expression, posture, tone and timing of speech (Bickmore & Picard, 2004).

The research conducted by Bickmore and Picard (2004) consisted of an animated agent that played the role of an exercise advisor helping learners through a behavior change program designed to increase the learners’ physical activity levels. The human-like agent, “Laura,” included speech and non-verbal behaviors, such as hand gestures, eye gaze behavior, posture shifts, head nods, proximity, and facial expressions (Bickmore & Picard, 2004, 2005). In the interactive environment, Laura provided feedback on the exercise behavior of the learner; helped the learners overcome obstacles to exercise; provided educational content related to exercise; and followed up on commitments to exercise (Bickmore & Picard, 2004, 2005).
In their study, Bickmore and Picard (2004) compared an interactive environment with a relational agent and another environment with a non-relational agent. The results indicated that participants felt that the agent cared about them, was genuinely concerned about their welfare, and that the agent liked them (Bickmore & Picard, 2004, 2005). The results of the month-long study implied that a relational agent can have a significant impact on people’s perception of caring (Bickmore & Picard, 2004, 2005).

Kim, Baylor, and Shen (2007) conducted a study that integrated psychological theories of emotions using emotional animated agents. The investigation focused on the impact of an animated agent’s emotional expression on learner’s affective and cognitive characteristics. The animated agents used in this study were a male and a female agent (Kim et al., 2007). The emotional expressions of the agents were achieved through verbal and facial expressions, tone of voice, and head movements (Kim et al., 2007). The animated agents had three emotions: positive, negative, and neutral. In the positive emotion, the agents had a happy, smiling face, and an engaging posture, with eye contact and head nodding (Kim et al., 2007). In the negative emotion, the agents had a somber and rather frowning face and an aloof posture, with evasive eye gaze and less head nodding (Kim et al., 2007). The emotional states of the agents were clearly communicated to the learners to ensure the validity of the findings (Kim et al., 2007). The neutral condition of the agents did not express emotions at all (Kim et al., 2007).

The results showed that the positive emotional state of the male agent significantly influenced the learner’s social judgment of the agent (Kim et al., 2007). Also, when the male agent expressed a positive emotion, the learners showed higher interest on the content of the instruction (Kim et al., 2007). Overall, the study showed that the learners’ affective characteristics were influenced by the digital peers’ emotion, as in the case of peer-to-peer interaction (Kim et al., 2007). According to Kim et al. (2007), this investigation confirmed that “smiley faces” may make a learner smile, but may not be sufficient to increase learning. However, it is important to note that one possible reason for not having an increase in learning might be the lack of range in the animated agents’ emotions (Kim et al., 2007). The agents expressed one constant type of each condition: happy, sad, or neutral, throughout the module (Kim et al., 2007).

Veletsianos (2009) also conducted an investigation focused on the animated agents’ expression of emotion. In the investigation, Veletsianos (2009) compared two versions of a tutorial lesson that differed in terms of verbal expressiveness. In the expressive version of the tutorial lesson, the
agent emphasized certain parts of speech by including additional pauses, instances where the content was delivered in a louder voice, and instances where words were more enunciated (Veletsianos, 2009). In terms of the agents’ non-verbal behaviors, gaze was predetermined and eye and eyebrow movement were coordinated (Veletsianos, 2009). The findings of this research, as in other studies on emotionally expressive animated agents, showed that the expressive agents’ interaction ability was rated more favorably than the non-expressive agent’s interaction ability (Veletsianos, 2009). In addition, the results indicated that the participants in the expressive agent group increased their learning outcomes when compared with the participants in the non-expressive agent group (Veletsianos, 2009).

Veletsianos (2009) showed that the impact of animated agents research and advances will be minimal if agents are designed to deliver pre-recorded and dispassionate lectures. The future of animated agents needs to focus on transforming content to engage and capture student attention. According to Veletsianos (2009), instructional designers should focus on designing efficient and effective instructional learning experiences, rather than designing media.

**CONCLUSION**

What do we know about animated pedagogical agents? Animated pedagogical agents have been used in instruction for several years, but they are now morphing into more interactive characters that possess human-like qualities. We also know that they must present a certain level of credibility and believability to serve as motivational factors in the instructional process. We know that the implementation of animated pedagogical agents in web- and computer-based instruction can lead to certain benefits, including: demonstration of physical tasks or procedures; navigational and attentional guides; aiding and advising students during problem-solving activities with non-verbal feedback and visual cues.

The lesson learned from research on animated pedagogical agents is that well-designed agent-based instruction provides emotional visual cues to their users that help them understand the rules of the environment (Cassell, 2001). Humans provide very strong visual cues about the protocols that they engage in, protocols that must be integrated into the very heart of a system in order for animated pedagogical agents to respond with appropriate surface-level behaviors (Cassell, 2001). Unfortunately, many of the animated pedagogical agents used in computer-based instruction consist of an agent inserted into a
system without a specific instructional purpose. They are capable of portraying a series of communicative poses without attention to how humans actually convey their knowledge of the world and of human interactions to their interlocutors (Cassell, 2001). Such systems represent an enormous missed opportunity.

As human-animated agent interaction becomes more common, it is more important that the interactive learning environments rely on the same interaction rules and emotions that humans use with each other (Cassell, 2001). The purpose is not just to add animated pedagogical agents for entertainment, but also to leverage learners’ natural tendencies to attribute humanness to the animated pedagogical agent.

We also need to understand that realistically animated pedagogical agents are not always the best approach for a web- and computer-based instructional product. There are circumstances in which the best-designed emotive animated pedagogical agent will not influence the learner. It is the designer’s job to have realistic expectations and to understand in which circumstances these expectations will be fulfilled or not. Furthermore, there are occasions in which the effort and cost of developing a high quality emotive animated pedagogical agent will be a hurdle to implementation.

However, when animated pedagogical agents are the best instructional approach, they should be designed with important character building qualities, as well as the psychological theories of emotion in order to portray a certain level of believability. As expressed earlier in this chapter, the aim is to increase the connection between the learner and the agent, while using the emotion component to enhance the cognitive processes that occur during learning. Research has determined that emotions serve as a significant element in the learning experience because they play a critical role in attention, planning, reasoning, learning, memory, and decision-making, and their influential capability toward perception, cognition, coping, and creativity of the learner (Johnson et al., 2000; Picard, 1997; Um, Song, & Plass, 2007).

REFERENCES


CHAPTER 12

Investigating Students’ Feelings and Their Perspectives Toward Web 2.0 Technologies in a Teacher Education Course

Yungwei Hao
National Taiwan Normal University, Taipei, Taiwan

INTRODUCTION

From the printing press to the static web pages of Web 1.0 to the current interactive user-generated applications of Web 2.0 (Reilly, 2005), educators have struggled with technology in order to anchor best practices and invigorate their instruction. Web 2.0 technologies, often called by the alternative names of social media or social software, have been heralded as a tool for information gathering, communication, interaction, and social networking. These social technologies can provide learners with user-friendly platforms to enhance communication with teachers and peers (Schroeder, Minocha, & Schneidert, 2010). The present study adopted the definition of Web 2.0 brainstormed in the first Web 2.0 conference at O’Reilly Media in October 2004, which defined Web 2.0 as an interactive platform, facilitating communication, interaction, and collective intelligence through dynamic display channels (Reilly, 2005).

Web 2.0 applications are cost-effective and can facilitate collaboration (Reilly, 2005). By creating opportunity for collaboration, Web 2.0 technologies improve learning, support the creation of social relationships, and develop social presence (Schroeder et al., 2010). According to Short, Williams, and Christie (1976), different media can achieve different levels of social presence, and the social presence of a communication medium can be understood as a user’s feelings toward the medium. The significance of social presence or feelings has been documented in a corpus of empirical studies (i.e., Gunawardena & Zittle, 1997; Rourke & Anderson, 2002; Tu, 2002). When students feel that they have social presence while using
a medium, they can better represent themselves in the online learning environment (Lowenthal, 2009), thereby achieving interpersonal involvement and collaborating to accomplish tasks (Kehrwald, 2008). Based on Short et al. (1976) concepts of feelings, this study hypothesized that different types of Web 2.0 technologies would achieve different degrees of social presence for students.

To better know the characteristics of the Web 2.0 technologies, the technologies need to be comprehensively studied in the classroom context (Hartshorne & Ajjan, 2009). Therefore, this study investigated 57 student teachers’ sense of social presence and their perspectives toward Web 2.0 technologies. The purposes of the present study were twofold. One was to address how students felt regarding social presence when using different types of Web 2.0 technologies. The second purpose was to explore students’ perceptions of the Web 2.0 technologies by conducting focus-group interviews to explain the survey data. The chapter is organized in the following manner. First, the theoretical background of the study is introduced, and a discussion of methodology follows. Results are reported separately for the two types of data collected (surveys and interviews). The interpretations of the results conclude with implications, limitations, and directions for further study.

**THEORETICAL BACKGROUND**

**Social Presence**

Through the mechanism of social connection, Web 2.0 technologies can afford users a sense of their social presence (Anderson, 2005; Grant, 2006; Schroeder et al., 2010). Short et al. (1976), the earliest researchers defining social presence, specified social presence as “the capacity of the medium to transmit information about facial expression, direction of looking, posture, dress and non-verbal cues” (Short et al., 1976, p. 65). The social presence of a communication medium can provide users with a sense of intimacy and immediacy. Short et al. (1976) described social presence as personal, sensitive, and warm. In other words, social presence is an effect of the medium that allows users to feel personal, sensitive, or warm. Based on Short et al.’s definitions, other researchers have conceptualized social presence in a variety of ways (Lowenthal, 2009; Tu and Isaacs, 2002). The present study, adhering to the original Short et al.’s definition, adopted the meaning of social presence as the capacity of an information medium to create a feeling of involvement in the users’ affective dimensions of personal, sensitive, or warm feelings.
Understanding more about students’ affective reactions while using Web 2.0 tools is essential for course designers intending to create effective online instruction. Short et al. (1976) indicated that social presence was a reliable discriminator to compare the effectiveness of various media. Social presence is a construct influenced by factors, such as the type of media used, individual characteristics, and learning context (Francescato et al., 2006). Communications media vary significantly in their ability to create a feeling of social presence in individual users (Short et al., 1976). However, only a few studies have described users’ sense of the social presence in a variety of Web 2.0 tools. Some studies of social presence have focused on students’ cognitive processes and ignored socio-emotional processes (Kreijns, Kirschner, Jochems, & Buuren, 2004). More currently, Hartshorne and Ajjan’s (2009) study was based on students’ expectations or abstract conceptualization of a few tools, but was not based on students’ real experience of using Web 2.0 in their learning environment. This large gap in our understanding creates an urgent need to thoroughly investigate the educational value of social presence and Web 2.0 technologies in actual practice. Therefore, the present study examined how students perceived their social presence, while using a variety of Web 2.0 tools in the learning context.

Although there is no study specifically investigating what criteria can influence levels of social presence, it is recognized that the social presence of a medium can allow users to feel a sense of intimacy and immediacy (Short et al., 1976). Social presence is contextual, and levels of social presence need to be measured in a particular context (Crim, 2006). Short et al. (1976) measured the sense of social presence of different media by rating bi-polar scaled items with the following dimensions: personal/impersonal; sensitive/insensitive; and warm/cold. Based on their study, Gunawardena (1995) used six paired items to assess users’ feelings regarding the social aspects of the computer-mediated communication (CMC) medium in her study: personal/impersonal; sociable/unsociable; sensitive/insensitive; warm/cold; interactive/non-interactive; and immediate/non-immediate. Considering the issues of reliability, validity, relatedness, and simplicity of use, the present study employed Gunawardena’s six pairs of adjectives as the criteria to measure students’ sense of social presence in Web 2.0 tools.

**TYPES OF WEB 2.0 TECHNOLOGIES**

Although there are many types of Web 2.0 applications, not all of the tools were designed specifically for educational use (Ferdig, 2007; Hartshorne &
This study focused on the educational use of blogs, wikis, social networks, social bookmarking, and a virtual world. This selection of technologies was based on an appraisal of student learning needs, the current level of technology acceptance, the potential value to education in general, and whether or not the technology tool was free of charge. According to Schroeder et al. (2010), blogs, wikis, and social networking sites are the most frequently used Web 2.0 technologies in education. Social bookmarking sites offer an abundant repertoire of ways to collect information and resources from websites (Churchill, Wong, Law, Salter, & Tai, 2009). Because of the richness of their interactional potentials, virtual worlds are gaining increased attention from educational research and practice (Freitas & Veletsianos, 2010). Each of the selected Web 2.0 applications is briefly discussed below.

Blogs, also called “web-logs,” provide a platform for individuals to publish text materials, and audio or video files, online. People visiting blogs can leave comments or interact with the blog owner or other visitors directly but not synchronously. Blogs provide a platform for self-expression and self-reflection through text or multimedia. The default display of content is in reverse-chronological order; therefore, logs can function as personal or group journals that facilitate metacognitive and higher-order thinking (Gunawardena et al., 2009). The present study chose the popular, Google Blogger, as the blogging tool.

A wiki, from the Hawaiian word “hurry,” is a web page where users can publish content created collaboratively. Wiki pages may be accessible to the public or be limited to a group. Creating wiki pages can be a collaborative generation of content in a shared space (O’Reilly, 2004; Wheeler, Yeomans, & Wheeler, 2008), and individual contributions can be monitored (Trentin, 2009). Through wiki participation, communities of practice (Wenger, 2000) can develop, and this type of social constructivism can lead to knowledge acquisition (Wheeler et al., 2008). The present study chose PBWorks as the wiki tool.

Facebook is a website self-described as a directory, connecting users with social networks. Social networking websites are indebted to the concept of “Six Degrees of Separation” (Travers & Milgram, 1969); on average, six steps/levels of separation connect any two persons. People can interact with friends and with their friends’ links (Fletcher, 2010; Lampe, Ellison, & Steinfield, 2006). According to the cover story published in Time magazine on May 31, 2010, Facebook has had 500 million active users and, by analogy, is the world’s third largest country by population. Users can add friends, post pictures, and send messages (Ferdig, 2007). The study chose Facebook as the social networking tool.
Social bookmarking is a Web 2.0 framework that facilitates the building of an easily accessible, online bookmark warehouse. After logging into a social bookmarking website, users assign descriptions (tags) to a chosen website, and then the website becomes searchable to their community members or to the public. Tags, also called “folksonomy,” allow users to manage their researched knowledge (Gunawardena et al., 2009). Users can browse their own bookmarks by time, tag, or topic (Farkas, 2008). The present study chose “Del.icio.us” (now “Delicious”) as the social bookmarking tool.

Second Life® is an Internet-based client that is downloaded from the Second Life® website. The interface provides a three-dimensional view, where users are immersed in a multiuser virtual environment. Users create characters and objects in this virtual world and can interact with each other in real-time through text or voice chat. The simulated context provides a social atmosphere and affluent learning opportunities. In this virtual environment, users can observe, experience, and create what might be impossible or difficult to do in real life. To date, Second Life® has been the most used virtual world in teaching and learning (Warburton, 2009). This study chose Second Life® as the virtual world tool.

Meyer (2010) compared eight students’ use of blogs, wikis, and online discussions, in a doctoral-level finance class. Meyer also examined levels of learning and student perspectives of the Web 2.0 tools used in the course. The results indicated that in the technology-integrated learning environment, the levels of learning mirrored the levels of the triggering questions the instructor raised. In addition, levels of learning were affected by students’ difficulties using blogs and wikis. Students were more comfortable with blogs, had both positive and negative opinions of wikis, and, in general, preferred online discussion boards. These findings cast some light on Web 2.0 usage, but left some issues unanswered. Meyer’s case-based study used only a few tools, had few participants, and the instructor provided monotonous questions, which restricted levels of learning. All of these variables may have influenced Meyer’s results. In another study, Hartshorne and Ajjan (2009) investigated 60 undergraduate students’ use and awareness of the potential of Web 2.0 technologies for education and learning. The findings indicated that most students felt that integrating Web 2.0 technologies into classrooms could be effective in improving course satisfaction, facilitating student learning, and promoting student interaction with peers and faculty. The combination of these elements enhanced the development of a student-centered learning environment and contributed to better knowledge creation and retention (Hartshorne & Ajjan, 2009; Maloney, 2007). Regarding the actual
student use of Web 2.0 in classrooms, the majority of the respondents in Hartshorne and Ajjan’s study did not use either blogs (56%) or social bookmarking (71%), and almost half of the research participants did not use social networks (46%) in educational contexts. A small number of respondents (20%) did not use wikis. Although the students agreed that there was educational value in using Web 2.0 applications, they rarely used the tools in their learning process (Hartshorne & Ajjan, 2009). Their awareness of the potential of the Web 2.0 tools was based on perceptions instead of actual use. Furthermore, the results might have been affected by feelings and attitudes, but feelings and attitudes were not examined. The effects of individual learner characteristics were not reported. Thus, more empirical studies are needed to illuminate these affective dimensions.

The primary focus of the study was to investigate students’ feelings regarding the social presence affordances of Web 2.0 tools and students’ perspectives in the use of a variety of Web 2.0 technologies. Combining the quantitative method of survey inquiry with qualitative data derived from focus-group interviews helped to clarify students’ emotional reactions to the technologies and interpret why the students felt the way that they did. Our mixed-method study strived to answer the following questions:

1. What were students’ feelings regarding social presence in use of each of the Web 2.0 technologies?
2. What were the students’ perspectives on the different types of Web 2.0 technologies?

**METHODOLOGY**

**Participants and Context**

The research participants were 57 first-year undergraduates in a 4-year teacher education university, enrolled in “Introduction to Educational Media,” a required course. The course was conducted in a computer lab. There were no prerequisites for taking the course. The study occurred in a country in which an official IRB system did not exist; yet the research team followed research ethics. An invitation to participate in this study, and participation was voluntary and anonymous. A total of 57 (out of 60 students in the class) agreed to participate in the study. The 57 participants included 15 males (26.3%) and 42 females (73.7%). Other demographic features, such as ethnicity, were all the same.

Student groups were required to complete a WebQuest project by the end of the semester that required the use of various Web 2.0 tools either
in or out of class. This inquiry-based learning technique (Borich, Hao, & Aw, 2006) was the course instructor’s pedagogical choice for incorporating Web 2.0 technologies into the instructional context.

During the first class–day in a campus computer lab, the students, with the guidance of the instructor, opened individual accounts in Second Life®. Each student chose an avatar and dressed it as a representation of “self.” The instructor directed the students to a virtual classroom at Second Life®, where they met their fellow student avatars and formed small groups. The Second Life® classroom was a virtual place where students could meet during the semester to synchronously discuss details for the class project. In addition, students were required to open an account with Google Blogger. The blogger tool functioned as a journal for students to reflect upon their learning in the weekly face-to-face meetings. The blogger tool allowed the students a place to construct their reflections and read each other’s blogs. Initially, to motivate students to co-create their WebQuest topics, the instructor raised course-related questions in class and in blogs. Once the student groups decided upon their WebQuest topics, the instructor opened accounts for each group in PBWorks. Each group had a shared workspace to brainstorm, plan, and document their project in the wiki website. Students also used the wiki space to co-write content for their WebQuest. The instructor also created student accounts in the “Delicious” bookmarking tool to facilitate students collecting and sharing online resources with their group members and the whole class. And, to facilitate students knowing each other better, to boost their momentum, and to facilitate their collaboration, students were required to open individual accounts in Facebook. During the process of conducting the WebQuest inquiry, students could synchronously discuss their findings in a virtual classroom in Second Life®, or asynchronously share their ideas in the wikis. Online bookmarks were created collaboratively and shared among group communities. Students kept journals and reflected on their learning in individual blogs. Group spirit was cultivated through the social networking site (Facebook). Finally, the finished WebQuest project morphed from text-based wiki documents to multimedia-embedded blogs, showcasing student teams’ results for the whole class.

**Data Collection**

Data were collected through online self-report surveys and focus-group interviews during the last week of the semester. Participants joined the study voluntarily and were guaranteed anonymity and informed that they could end their participation at any time. The 57 students spent approximately
30 min in a computer lab on the last class-meeting day, completing the surveys. In total, 12 volunteers participated in a focus-group interview through semi-structured questions. Further details follow regarding the focus group interview.

**Instrument**

Questions concerning the frequency and purpose of individual technology use were excerpted from Hartshorne and Ajjan’s survey (2009) and were incorporated into this study. The surveys used in the study are described below.

*Assessment of Feelings:* Gunawardena (1995) modified Short et al.’s (1976) scale and developed six items to measure the social aspect in CMC technologies. The six items are bi-polar, with levels ranging from 1 to 5. The dimensions include perceived feelings and are considered social indicators (Gunawardena & Zittle, 1997). They are: personal (1) to impersonal (5); sociable (1) to unsociable (5); sensitive (1) to insensitive (5); warm (1) to cold (5); interactive (1) to non-interactive (5); and immediate (1) to non-immediate (5). The present study used the six pairs to assess feelings. The survey used in this study produced reliable indicators of social presence in all of the different tools (all $\alpha > 0.75$). The survey was valid due to the high construct validity and reliability (strongly positively correlated with other social presence measures, $r > 0.50$) (Gunawardena, 1995; Gunawardena & Zittle, 1997).

**Interviews**

Focus-group interviews elicited information from the 12 focus-group participants about their perspectives on the Web 2.0 tools. The 12 participants, a part of the 57 students, voluntarily took part in the study. Group interviews can facilitate personal disclosure (Farquhar, 1999). The group of participants met with the research team on a single occasion on campus, and the discussion lasted for approximately 3.5 h. The informal group discussions were audio-recorded and transcribed. Interviews were semi-structured with open-ended questions, and the interviewees reflected on their experiences using the Web 2.0 tools in the course. Sample interview questions included: “What and how did you feel while using [the technology]?”; “What do you think about [the technology]? Why?”; “How did [the technology] help you complete the course?”; and “What problem(s) did you have using [the technology] in the course?”

**Data Analysis**

Descriptive statistics were used to analyze data with respect to the demographic profile and students’ feelings toward Web 2.0 technologies. A one-way
repeated measures analysis of variance (ANOVA) was conducted to determine whether there was a significant difference in students’ feelings using the tools. The focus-group interviews were transcribed orthographically (preserving just the words spoken). The content was analyzed through the Strengths, Weaknesses, Opportunities and Threat (SWOT) framework (Glaister & Falshaw, 1999) to elicit information from the participants about their responses to the Web 2.0 tools. SWOT, originated in the business domain, is an acronym for SWOT. Its analysis technique is suitable for developing strategies, making environmental evaluations and situation analyses. SWOT can be used to illuminate the features of an organization or entity. The flexibility of the SWOT analytical framework makes it applicable in diverse fields (Fernandez, 2009; Jackson & Helms, 2008). SWOT can organize large numbers of issues surrounding complex entities, although it does not provide solutions, as Jackson and Helms (2008) commented. Web 2.0 technologies have many different characteristics and represent diverse environments. When instructors make instructional decisions on technology integration, the process is strategic and complex. Therefore, the present study used the SWOT analysis to clarify the features of the Web 2.0 tools.

By examining each type of Web 2.0 tool from the students’ perspective, the present study highlighted the SWOT of each tool. In our SWOT framework, strengths and weaknesses represented the internal characteristics of the Web 2.0 tools, as reported by students. Opportunities and threats were the external opportunities and threats that the students perceived in the tools. Through counting the number of responses falling within the four dimensions, and then summarizing the number and calculating the percentage of responses for each category, a few themes emerged. To ensure reliability, two graduate students who were familiar with the SWOT coded the content of interviews and classified codes into the framework. Afterwards, two education researchers verified the categorization within the SWOT framework. The general interrater reliability was 0.83. The combination of the qualitative and quantitative information was expected to increase educators’ understanding of the educational value of the various Web 2.0 technologies.

RESULTS

The Quantitative Results

Demographic Data

More than 60% of the participants “made friends” online. Approximately 90% of the participants occasionally used blogs, wikis, or social networks. Half rated themselves as competent in using blogs, wikis, and social
networks. Over 70% of the participants had not used the social bookmarking site or the virtual world prior to the course. Approximately 80% of the participants rated themselves as novice users of social bookmarking and virtual worlds. More than 60% of the participants indicated that they mostly used blogs in their learning, wikis next (almost 30% said wikis were their first choice), and the rest of the technologies were used very little for learning purposes (see Tables 12.1-12.3 for more details).

Research Question 1: What Were Students’ Feelings in Use of Each of the Web 2.0 Technologies?

Descriptive statistics were used to present the data. Table 12.4 and Figure 12.1 illustrate the details. In general, among the five types of Web 2.0 tools, the social networking site was rated as the most sociable, sensitive, warm, interactive, and immediate tool. The social bookmarking site was rated as the most unsociable, cold, non-interactive, and non-immediate tool. Blogs were rated the most personal, while wikis were rated the most impersonal. The virtual world was rated the most non-immediate and insensitive. For each of the social indicators, the rating details are as follows.

In the personal-impersonal dimension, blogs were rated as the most personal, with the social networking site, the social bookmarking site, the virtual world, and wikis following, respectively. In the sociable-unsociable dimension, the social network site was rated the most sociable, with blogs, the virtual world, wikis, and the social bookmarking site following, respectively. In the sensitive-insensitive dimension, the social networking site was rated the most sensitive, with blogs, wikis, the social bookmarking site, and the virtual world following, respectively. In the warm-cold dimension, the social networking site was rated the warmest, with blogs, the virtual world, wikis, and the social bookmarking site following, respectively. In the interactive-non-interactive dimension, the social networking site was rated the most interactive, with blogs, the virtual world, wikis, and the social bookmarking site following, respectively. In the immediate-non-immediate dimension, the social networking site was rated the most immediate, and wikis, blogs, the virtual world, and the social bookmarking site following, respectively. All details are displayed in Table 12.4.

The Qualitative Results

Research Question 2: What Were the Students’ Perspectives on the Different Types of Web 2.0 Technologies?

The focus-group interviews provided information about student perspectives on the different types of Web 2.0 technologies and were analyzed through the
Table 12.1 Participants’ comfort levels with Web 2.0 technologies

<table>
<thead>
<tr>
<th></th>
<th>Blog</th>
<th>Wiki</th>
<th>Social networks</th>
<th>Social bookmarks</th>
<th>Virtual world</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. (%)</td>
<td>Freq. (%)</td>
<td>Freq. (%)</td>
<td>Freq. (%)</td>
<td>Freq. (%)</td>
</tr>
<tr>
<td>Never use</td>
<td>2</td>
<td>3.5</td>
<td>1</td>
<td>1.8</td>
<td>1</td>
</tr>
<tr>
<td>Use, but novice</td>
<td>11</td>
<td>19.3</td>
<td>21</td>
<td>36.8</td>
<td>8</td>
</tr>
<tr>
<td>Use, and competent</td>
<td>29</td>
<td>50.9</td>
<td>32</td>
<td>56.1</td>
<td>28</td>
</tr>
<tr>
<td>Proficient</td>
<td>15</td>
<td>26.3</td>
<td>3</td>
<td>5.3</td>
<td>20</td>
</tr>
</tbody>
</table>

Investigating students’ feelings and their perspectives toward Web 2.0 technologies.
<table>
<thead>
<tr>
<th>Frequency of Use</th>
<th>Blog</th>
<th>Wiki</th>
<th>Social networks</th>
<th>Social bookmarks</th>
<th>Virtual world</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq.</td>
<td>Percentage (%)</td>
<td>Freq.</td>
<td>Percentage (%)</td>
<td>Freq.</td>
</tr>
<tr>
<td>Never use, and will not use</td>
<td>3</td>
<td>5.3</td>
<td>1</td>
<td>1.8</td>
<td>6</td>
</tr>
<tr>
<td>Never use, but will use</td>
<td>3</td>
<td>5.3</td>
<td>7</td>
<td>12.3</td>
<td>2</td>
</tr>
<tr>
<td>Occasionally use</td>
<td>32</td>
<td>56.1</td>
<td>26</td>
<td>45.6</td>
<td>28</td>
</tr>
<tr>
<td>Often use</td>
<td>12</td>
<td>21.1</td>
<td>14</td>
<td>24.6</td>
<td>11</td>
</tr>
<tr>
<td>Always use</td>
<td>7</td>
<td>12.3</td>
<td>9</td>
<td>15.8</td>
<td>10</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------</td>
<td>----------------------</td>
<td>------------</td>
<td>---------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>63.2</td>
<td>17</td>
<td>29.8</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 12.4 Each dimension of students’ sense of social presence using each of the Web 2.0 technologies (Mean±SD, n = 57)

<table>
<thead>
<tr>
<th></th>
<th>Blog</th>
<th>Wiki</th>
<th>Social bookmark</th>
<th>Social network</th>
<th>Virtual world</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal (1) to impersonal (5)</td>
<td>1.72±0.10</td>
<td>2.82±0.97</td>
<td>2.47±0.89</td>
<td>1.77±0.78</td>
<td>2.51±1.02</td>
</tr>
<tr>
<td>Sociable (1) to unsociable (5)</td>
<td>1.72±0.96</td>
<td>2.23±0.95</td>
<td>2.56±0.89</td>
<td>1.28±0.65</td>
<td>2.12±1.12</td>
</tr>
<tr>
<td>Sensitive (1) to insensitive (5)</td>
<td>2.49±0.95</td>
<td>2.67±0.85</td>
<td>2.93±0.68</td>
<td>2.19±0.83</td>
<td>3.11±0.72</td>
</tr>
<tr>
<td>Warm (1) to cold (5)</td>
<td>2.49±0.95</td>
<td>2.91±0.83</td>
<td>2.95±0.74</td>
<td>2.14±0.83</td>
<td>2.86±0.77</td>
</tr>
<tr>
<td>Interactive (1) to non-interactive (5)</td>
<td>1.74±0.96</td>
<td>1.98±0.99</td>
<td>2.46±0.80</td>
<td>1.39±0.70</td>
<td>1.96±0.98</td>
</tr>
<tr>
<td>Immediate (1) to non-immediate (5)</td>
<td>2.23±0.98</td>
<td>2.12±0.83</td>
<td>2.47±0.91</td>
<td>1.58±0.76</td>
<td>2.47±0.93</td>
</tr>
</tbody>
</table>
SWOT framework. The strengths and weaknesses of the technologies were organized into three themes: the information collected by the technology, the interface, and the interaction supported by the technology. Student perceptions of the technologies’ affordances and threats were indicated. The results are reported by points in Tables 12.5-12.9 and explained below.

**Blogs (Google Blogger):** Table 12.5 illustrates the results of student responses to blogs. Regarding internal strengths, blogs in which the content was distributed via the RSS (Really Simple Syndication) were considered easy to use and multimedia capable (interface), created different types of interaction and made thoughts visible (interaction). The information that the blogs provided was broad, and the content of blogs represented personal styles (information). For example, one student said, “It’s easy to have a blog site. I can record things by writing text or posting pictures anytime. Blogs are personal to me. I communicate my thoughts and feelings to my net friends through blogs. Writing blogs got me a couple of friends. Some blogs are really interesting. I can write, or I can read others’. I have lots of freedom.” In addition to expressing their own thoughts or emotions, bloggers can read others’ blogs, observe others’ work, and reflect on the thoughts, having indirect or vicarious interaction with the blog participants. Internal weaknesses of the blogs included a lack of user controls for advanced features (interface); the threat of loss of privacy (information); lack of timely responsiveness; writing-based; and the time-consuming nature of keeping a blog (interaction). A few interviewees mentioned that they did not blog when they were busy.
Table 12.5 Blogs (Google Blogger): SWOT analysis

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Percentage (%)</th>
<th>Weaknesses</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interface:</strong> easy to use; personalizable; multimedia capable; good tracking; rapid update</td>
<td>18 33</td>
<td><strong>Interface:</strong> limited file storage capacity; limited user control of advanced features; limited time-frame to gain readers’ attention (only newer posts are featured)</td>
<td>2 4</td>
</tr>
<tr>
<td><strong>Interaction:</strong> affective, sociable, cognitive, and vicarious interaction (lurking); reveals thoughts</td>
<td></td>
<td><strong>Interaction:</strong> delayed response; time-consuming; writing-based</td>
<td></td>
</tr>
<tr>
<td><strong>Information:</strong> records personal work; journaling; supporting the expression of opinion and emotion (self-disclosure); collecting topic-oriented information</td>
<td></td>
<td><strong>Information:</strong> loss of privacy</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Percentage (%)</th>
<th>Threats</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy distribution of learning materials</td>
<td>25 46</td>
<td>Time-consuming</td>
<td>9 17</td>
</tr>
<tr>
<td>Supports portfolio development</td>
<td></td>
<td>Possible exposure to cyber-harassment and cyber-bullies</td>
<td></td>
</tr>
<tr>
<td>Immediate publishing in multimedia or written format</td>
<td></td>
<td>Teachers informally blogging with students could weaken perceived teacher authority</td>
<td></td>
</tr>
<tr>
<td>Facilitates reflection and self-directed learning</td>
<td></td>
<td>Students who prefer privacy may feel anxious posting personal writing</td>
<td></td>
</tr>
<tr>
<td>Facilitates higher-order thinking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharing thoughts promotes interaction among class members</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blogs can be used as a study guide and to review class materials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support equal participation and provide opportunities to lessen social hierarchies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encourages creativity (learners create something online)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Table 12.6 Wiki (PBWorks): SWOT analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Strengths</strong></td>
<td><strong>Percentage (%)</strong></td>
<td><strong>Weaknesses</strong></td>
<td><strong>Percentage (%)</strong></td>
</tr>
<tr>
<td><strong>Interface</strong>: easy to learn, post information, collaborate, and easy tracking of individual contributions; centralized content management; individuals have equal access; updates appear immediately</td>
<td>10</td>
<td>22</td>
<td><strong>Interface</strong>: limited accessed to one person at a time; unstable tracking system; weak organization of information; no user-controls for the interface</td>
</tr>
<tr>
<td><strong>Interaction</strong>: collaborative, cognitive, and vicarious interactions (lurking)</td>
<td></td>
<td></td>
<td><strong>Interaction</strong>: no real-time communication between users</td>
</tr>
<tr>
<td><strong>Information</strong>: supports collective intelligence; a consensus-building tool</td>
<td></td>
<td></td>
<td><strong>Information</strong>: unstable information, no assurance of accuracy</td>
</tr>
<tr>
<td><strong>Opportunities</strong></td>
<td><strong>Percentage (%)</strong></td>
<td><strong>Threats</strong></td>
<td><strong>Percentage (%)</strong></td>
</tr>
<tr>
<td>Facilitates collaborative learning and reciprocal teaching</td>
<td>18</td>
<td>40</td>
<td>Effective as an information repository but lacks organization</td>
</tr>
<tr>
<td>Tracks individual contributions</td>
<td></td>
<td></td>
<td>Interface unpopular with students</td>
</tr>
<tr>
<td>Trains students to organize thoughts</td>
<td></td>
<td></td>
<td>Interface limited to English only</td>
</tr>
<tr>
<td>Encourages students to participate in group discussion</td>
<td></td>
<td></td>
<td>Editing content risks offending fellow students</td>
</tr>
<tr>
<td>Encourages students to think more deeply and practice metacognition</td>
<td></td>
<td></td>
<td>Students could participate inequitably</td>
</tr>
<tr>
<td>Provides students with opportunities to experience the value and inviolability of intellectual property</td>
<td></td>
<td></td>
<td>Concern for plagiarism</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>If students are not familiar with their collaborators’ contributions, it may be more difficult to achieve consensus than in face-to-face meetings</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Accuracy of information may not be reliable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>High-level skills for information organization are required</td>
</tr>
<tr>
<td>Table 12.7 Social network (Facebook): SWOT analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Strengths</strong></td>
<td><strong>Weaknesses</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td><strong>Percentage (%)</strong></td>
<td><strong>n</strong></td>
<td><strong>Percentage (%)</strong></td>
</tr>
</tbody>
</table>

**Interface:** multi-functional (photo album, e-mail box, data storage, etc.); immediate update; networking, searching people by names; easy to learn

**Interaction:** social, playful, communicative, instant, and vicarious interaction (lurking); facilitates inter-cultural communication; allows users to observe others’ interaction

**Information:** group, friend, and event updates; networking and re-connecting with friends who the user lost contact with; entertainment

| 9 | 31 | Interface: redundant updates filling e-mail boxes; interface embedded with too many functions; frequent changes of interface |
| 3 | 10 | Interaction: time-consuming; unknown people asking to be added to the friend list |
|   |     | Information: too many updates; loss of privacy (i.e., photo exposure in public); information overload |

**Opportunities**

A community-building tool
Better understanding of self and fellow students through games
Interaction and networking among students, teachers, or school administrators

| 10 | 35 | Possibly addictive |
|    |     | No protection from cyber-bullies |
|    |     | Network extends beyond the classroom; network complicates students’ friend communities |
|    |     | Extend classroom problems, such as school cyber-bullying and peer conflict to the cyber world |
|    |     | Mislead students into wrong values (i.e. farm games convey that theft is okay) |
|    |     | Social networking is superficial |

| 7 | 24 |
Table 12.8 Social bookmarks (Del.icio.us): SWOT analysis

<table>
<thead>
<tr>
<th>Strengths</th>
<th>n</th>
<th>Percentage (%)</th>
<th>Weaknesses</th>
<th>n</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interface</strong>: ability to save and share bookmarks; a logical tagging system; aggregation of websites; high accessibility; easy to learn. Mutual tagging exposes valuable or popular websites</td>
<td>16</td>
<td>59</td>
<td><strong>Interface</strong>: difficulty retrieving saved websites when tags were forgotten; lack of interactional interface</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><strong>Interaction</strong>: cognitive, collaborative, and vicarious interaction (lurking); efficient and instant way to search, share and collaborate</td>
<td></td>
<td></td>
<td><strong>Interaction</strong>: lacking a mechanism for interaction between users</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Information</strong>: sustainable and global; more trustworthy than wikis</td>
<td></td>
<td></td>
<td><strong>Information</strong>: lack of clear criteria for searching bookmarks</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Opportunities</strong></td>
<td>n</td>
<td>Percentage (%)</td>
<td><strong>Threats</strong></td>
<td>n</td>
<td>Percentage (%)</td>
</tr>
<tr>
<td>Facilitate self-directed and inquiry-based learning</td>
<td>8</td>
<td>30</td>
<td>Problems with locating the tagged websites</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Extend topics of targeted information</td>
<td></td>
<td></td>
<td>Only English interface available</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop a global knowledge-management community</td>
<td></td>
<td></td>
<td>Low-level of interaction between learners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop global views</td>
<td></td>
<td></td>
<td>Lack of standards or censorship</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extend learning boundaries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrich class curriculum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strengths</td>
<td>n</td>
<td>Percentage (%)</td>
<td>Weaknesses</td>
<td>n</td>
<td>Percentage (%)</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----</td>
<td>----------------</td>
<td>-----------------------------------------------</td>
<td>----</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>Interface</strong>: appealing; authentic; realistic; entertaining; delicately designed avatars; beyond text presentation (multimedia); high level of user control; easy to make friends</td>
<td>10</td>
<td>27</td>
<td><strong>Interface</strong>: high requirement for computer equipment and Internet bandwidth; installation of software required; reports of users feeling dizzy using the 3D interface; impersonal text-based message board; cross platform limitation (objects cannot be imported from outside programs); high learning curve</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td><strong>Interaction</strong>: facilitates all types of interaction; moves from one virtual place to another in a second; real-time, real-life interactions</td>
<td></td>
<td></td>
<td><strong>Interaction</strong>: requires (international) social skills and advanced technology skills in order to join into conversations</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Information</strong>: provides simulation of real-life situations; bringing imagination into play</td>
<td></td>
<td></td>
<td><strong>Information</strong>: unreliable or inaccurate</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Opportunities</strong></td>
<td>14</td>
<td>38</td>
<td><strong>Threats</strong>: technical problems; high learning curve</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>Provide a social platform; facilitate role play; enhance interpersonal and communication skills; facilitate culture exchange; promote cultural awareness; cultivate community of practice</td>
<td></td>
<td></td>
<td>Limited language availability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encourage authentic, immersive, experiential, and situated learning; promote learning motivation; support online instruction; display (art) work online</td>
<td></td>
<td></td>
<td>Lack of authority to maintain order in the virtual world; anonymity may create unsafe conditions; difficulty in monitoring students' choices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promote foreign language training; facilitate development of interdisciplinary curriculum and instruction; encourage innovation and creation of instructional materials; make on-site global or international education possible</td>
<td></td>
<td></td>
<td>Easy to become superficial learning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The external opportunities that instructor-prompted blogs afforded were the facilitation of higher-order thinking and the students felt encouraged to write or publish online. In terms of educational objectives, educators can go beyond Bloom’s (1956) traditional taxonomy of learning domains to incorporate Anderson, Krathwohl, Airasian, and Cruikshank (2001) taxonomy, which ranks creation as the highest objective in cognitive learning. Blogs encourage users to create something, which concurs with Anderson et al.’s (2001) newly added educational objective. For example, some students stated that blogs provided a learning space for them to create, reflect, discuss, and share information with each other. One interviewee communicated a common theme among the students: “I wrote blogs every week, and they helped me reflect what I learned in class. Blogs made me have a reason to write. They made a difference for my classroom learning.”

External threats that blogs brought to the learning experience included concerns over a blogger’s vulnerability to cyber-harassment or cyber-bullying. Additionally, concerns about personal disclosure were communicated. One interviewee mentioned that she hesitated to share thoughts or feelings in public, because she preferred keeping a low profile and did not like displaying her thoughts to strangers. Two interviewees commented that the instructor’s authority could be diminished by blogging with students, because of closeness.

Wikis (PBWorks): Table 12.6 illustrates the analysis of student response to wikis. Internal strengths of the wikis included ease of collaborative documentation, tracking of individual contributions, equal access and small learning curves for users (interface). Wikis were considered a good consensus-building tool by the interviewees and an easy way to create collective intelligence (information). Most of the interviewees commented that wikis were an effective tool to co-write and modify content. In the process of co-writing, the participants cooperated and brainstormed with each other. The platform created opportunities for participants to observe others’ ideas through collaboration and vicarious interactions (interaction). Internal weaknesses included an unreliable tracking system, a weak mechanism for information organization, no user-control in the interface, and a weak interface design (interface). The information in wikis was perceived as unstable and easily changed (information). A few interviewees complained that real-time or synchronous communication was not possible in the wikis (interaction).

External opportunities wikis brought to the educational experience included engaging students in the co-construction of knowledge and encouraging collaborative learning, reciprocal teaching, and group discussions. The wiki
tracking mechanism was designed to document student contributions during collaboration and to display the level of students’ contribution to their peers. One interviewee stated that the wikis encouraged her to think more deeply and reflect upon her contributions. When she posted, she said she had to organize her thoughts and evaluate how to integrate her words into the document. Another interviewee reported discomfort when her words were copied by a group member and not cited with her name, recognizing the value and inviolability of intellectual property.

*Perceived external threats of wikis* were reported. First, the interface did not encourage interaction. Interviewees mentioned that once they posted their content, they did not stay to read their peers’ contributions until when they were required to submit group work. Because of this, the students were not necessarily familiar with the content of their collaborators’ documents, and the students had difficulties achieving consensus in their collaborative writing. A further interface problem as reported by a few interviewees was that *PBWorks* was only available in English. Another concern of threat was content-related. Two interviewees expressed concern that some fellow students might be offended if their writing was edited. More than half of the interviewees commented that they were hesitant about the accuracy of the posted content. And several interviewees mentioned concern about plagiarism. They were fearful that once content was posted, unprepared group members could simply paraphrase. A few interviewees expressed that they lacked confidence in their editing skills. The final threat came from students’ perceived lack of cognitive preparedness. Due to an overwhelming amount of information coming in, wikis required a high level of information-organization skills, which many interviewees admitted that they lacked.

*Social networks (Facebook):* Table 12.7 illustrates the analysis of students’ responses to social networking. Reported *internal strengths* included multifunctional affordances (photo album, e-mail box, etc.); networking with friends; searching people by names; immediate updates on how friends are doing (*interface*); creating topics for friends’ chats; making new friends; sharing feelings; synchronous interactions; facilitating intercultural communication; fun communication (i.e., comments on Walls); releasing pressure; increasing the closeness with family and friends; allowing users to observe others’ interaction; a good way to spend time (entertainment, i.e., owning Farms); and playing games and sharing results with friends (*interaction*). One interviewee mentioned that she enjoyed chatting with an Iranian, because it was a novel experience for her. Students reported that they enjoyed the tool’s updates of friends, groups, or events connected with one’s preferences.
Friends one had lost contact with could be re-connected through Facebook (information). On the other hand, students reported that too many updates interfered with privacy (information), and therefore caused internal weaknesses. Other issues included unknown people requests to be added to friend’s list, time-consuming activities related to interacting in Facebook (interaction), and instability of the interface and messages constantly sent from Facebook filling personal e-mail boxes (interface). Several interviewees complained that Facebook created annoying and overwhelming messages. These notices jammed their e-mail boxes with irrelevant and useless information. A few interviewees complained about the frequent changes of the Facebook interface requiring extra time and effort to manage the changes. Regarding photograph exposure in public, one interviewee shared her unpleasant experience about her photograph and how she incurred harassment.

External opportunities of Facebook, brought to education, included: community-building (a bulletin board, a voting tool, and an announcement of events); promotion of interaction between students, teachers, or school administrators; and a better understanding of oneself and fellow students through the embedded psychological games. One student mentioned that through Facebook, he made a connection with a university professor, and attended academic events that the professor recommended.

External threats included: encouraging addiction to games (Virtual Farms); too much social networking beyond the class boundary; complicating students’ friend communities; misleading students that theft was okay (e.g., via Farm games); exposing students’ privacy while they were in a vulnerable situation as students; and facilitating the extension of school cyberbullying and peer conflict to the cyber world. One interviewee emphasized that she felt that the interaction in Facebook was superficial and interactions with people were on a surface level. She did not feel connected with fellow students through the online social networking mechanism.

Social bookmarking (Del.icio.us): Table 12.8 illustrates the analysis of students’ responses to social bookmarking. The students reported that the internal strengths included: the collected information was global and sustainable, and more trustworthy than wikis (information); saving and sharing bookmarks; easy to organize information by tagging (small learning curve); and mutual tagging. Mutual tagging exposed valuable or popular websites, which may have provided extra information, a tip some interviewees mentioned and then shared with the focus group (interface). Some interviewees said that, when they wanted to know what someone was doing, they could review that individual’s bookmarks. In addition, social bookmarking
provided an efficient way to search, share, and collaborate; information was updated instantly, as it was collected (interaction). Internal weaknesses included a poor interface that lacked a user-friendly design. Several interviewees mentioned that the bookmarking site did not provide a way for users to interact with one another. One interviewee said, “If I could have left feedback (i.e., a thank-you note) or comments or seen others’ comments in Del. icio.us, I would have felt warmer and more human,” using the tool. Another participant reported that one weakness depended upon the user recalling how websites were originally tagged.

External opportunities, brought to education, included: the development of a global knowledge-management community; the extension of learning boundaries; facilitating self-directed and inquiry-based learning; developing global views; and enriching classroom curriculum. “By checking some bookmarks on the topic I’m interested in, I would be able to get more information on the topic,” one interviewee stated.

External threats included an English-only interface, thereby creating inequality for students who were not strong English readers. Other participants pointed out that the lack of feedback (no interaction system) made a cold atmosphere for some learners. A few interviewees voiced concerns that violence or porn-related websites might be integrated in some collections of bookmarks.

Virtual World (Second Life®): Table 12.9 lists the analysis of students’ responses to the virtual world. Internal strengths included: an appealing interface; delicately designed avatars; a high level of user control (i.e., selection of avatars, user as builder/creator); ease in making friends; authentic real world feeling (interface); diverse ways to communicate (text, facial expression, gesture, voice); moving from one virtual place to another in a second (teleporting); real-time, real-life communication (interaction); authentic, brought imagination into play (information). The mechanism was familiar to users who played online games, and several interviewees expressed fascination with the real-time, real-life, authentic environment. Internal weaknesses included the requirement of social skills to enter into conversations (interaction) and unreliable or inaccurate information to users (information). Moreover, there were a few technical problems (i.e., required software installation, large computer memory, advanced hardware, and high-speed bandwidth). Several complained that the virtual world caused their computers to shut down repeatedly. Objects could not be exported from other programs, the 3D concepts were confusing, 3D objects caused dizziness for some users, and the program required a high learning curve for users.
Another interface problem was that, when students were in the virtual classroom, the messages posted in the message board could not be associated with their authors. Most of the interviewees complained about technical problems. One interviewee reported, “I tried to change my skin, but it lagged so seriously that I couldn’t carry out tasks as freely as I wanted to.” A few reported that the difficulty in learning how to function in the virtual world prevented them from using it.

External opportunities that were brought to education included: providing a social platform and environment for situated and experiential learning; promoting foreign language training; facilitating development of interdisciplinary curriculum and instruction; immersing learners in authentic environments; encouraging innovation and creation of instructional materials, training interpersonal and communication skills; facilitating culture exchange, creating international experience; creating an alternative way for classroom instruction/discussion, enriching online instruction; making on-site global or international education possible; and fostering self-realization. Some students mentioned they were excited with the authenticity of the virtual world, and that they became more motivated and engaged to participate in class activities in the immersive learning environment.

External threats included: technical problems; difficulty monitoring students’ choices; limited availability of languages; lack of law to maintain order in the virtual world; threat of pornography, violence, or improper content; and anonymity creating unsafe conditions for students. A few interviewees stated if they did not know who the users were behind the avatars, they felt a loss of control and unsafe with the environment. Some interviewees were concerned that learning in a virtual world could be superficial if the course design was entertainment-oriented and lacked pedagogical support. The interviewees mentioned that, when they were undertaking learning tasks in Second Life® they were distracted by scenes or objects nearby. Furthermore, the difficulty in learning the tool dissuaded several students who lacked technological confidence and who had unstable hardware equipment, from participating and engaging in class activities.

**DISCUSSION**

The survey data indicated that most of the students were novice users of social bookmarking and the virtual world, but most were familiar with blogging, wikis, and social networking. These results concur with Schroeder et al.’s (2010) finding that students were more familiar with the same three
types of tools. Data from the interviews indicated that blogging and social networking were easier to learn, and the students were more familiar with these tools. Familiarity with certain tools may help explain why the students regarded blogging and social networking as more helpful in improving course satisfaction and facilitating interaction with others. The interviewees indicated that both blogs and wikis inspired them to write and enhance their learning. Findings, in this study, may help explain why the students regarded blogs and wikis as being most helpful for their classroom learning and improvement of their writing. The result was consistent with the findings of Williams and Jacobs (2004) blog study and Matthew and Callaway’s (2009) wiki study, that students felt encouraged to express thoughts or write ideas in blogs or wikis.

AFFECTIVE DIMENSIONS OF WEB 2.0

In the personal-impersonal dimension, the results of the study concurred with the findings of Sim and Hew’s (2010) review study that students felt the most personal sense in blogs. When students blogged, they felt in control of the blog interface and its content. This study indicated the highest personal level in blogs, and the students expressed strong ownership in their blog sites. On the other hand, interviewees mentioned that their awareness that their work in wikis would likely be edited may explain why wikis were rated the most impersonal. In the sociable-unsociable dimension, the social networking site was rated the most sociable, and the virtual world was rated the least sociable. The interviewees disclosed that they easily made friends through online social networking. The instantaneous interactions and sharing of emotions made them feel connected, which may account for social networking’s highest rating. In addition to the technical problems reported earlier, students did not appreciate that, at first, they were “house arrested” by the course instructor in the virtual classroom. They were asked to form a group and conduct class activities. Since they were not afforded the opportunity to informally meet each other before being asked to complete learning activities may have influenced them to feel unsociable using the virtual world tool. In the sensitive-insensitive dimension, the interviewees pointed out that they received many invitations and updates from the social networking site, which may explain why social networking was rated as the most sensitive tool. Most of the participants, in the present study, were novice users of the virtual world, and they tended to be ignored by the other residents. One interviewee recalled, “I kept saying Hi to a lady and a guy, and they
didn’t even respond and just walked away. That was so rude. I felt neglected and distressed.” And the technical problems mentioned previously may explain why the virtual world was rated the lowest in sensitivity.

In the three dimensions, *warm-cold, interactive-non-interactive*, and *immediate-non-immediate*, the results were similar: social networking was rated the warmest, the most interactive, and immediate. Social bookmarking was rated the coldest, the least interactive, and immediate. The interviewees mentioned that, after they had expressed emotions in the social networking site, they quickly received feedback from friends or even from strangers. One interviewee commented that family and friends became closer through social networking because of the constant updates. These affordances may help reveal why social networking was rated the warmest, and the most interactive and immediate tool. This finding corresponds to Ellison, Steinfield, and Lampe (2007) finding that social networking has a sufficient amount of social resources to connect users together. The social bookmarking site was rated the coldest, the least interactive, and the least immediate, perhaps because users could not leave comments or feedback on the site. The main task users conducted was to bookmark websites and share their resources with the local or global communities. The interaction was indirect and implicit, and the immediacy was low. These variables may explain why social bookmarking sites received the lowest ratings. The virtual world was also rated the least immediate tool for receiving feedback. The students reported that they did not receive feedback or interactions. Perhaps it was because the students did not get to meet more advanced users who were willing to help, or perhaps because the students lacked the relevant social skills to interact with the virtual world residents. The variety of reasons may have contributed to the fact that users did not acquire immediate feedback or assistance to explore the virtual world.

Combining the findings concerning the six dimensions of feelings, in general, the present study indicated, in Figure 12.1, that social networking achieved the most feelings, and the social bookmarking site achieved the least. Social networking was designed to facilitate social interaction, to connect users with each other, and be part of a community (Ellison et al., 2007). That may explain the phenomenon of the highest social presence in the social networking tool. The social bookmarking site was designed to organize information, not for social interaction. A few interviewees mentioned that the social bookmarking site lacked an interaction mechanism, a way for the users to leave comments, which may explain why the feelings were the lowest using the social bookmarking tool.
A virtual world provides a social environment where students are situated in an authentic context, interacting with each other, to experience and learn (Salmon, Nie, & Edirisingha, 2010). In this study, students did not feel much social presence in the virtual world. That may be because they were generally unfamiliar and unskilled in working within a virtual world, and so were unable to realize the social potentials of the tool. As indicated in a few interviews, the dominant English language culture in the virtual world left minority groups feeling neglected, or inferior and lacking intimacy and immediacy in the virtual world. Some students indicated that they became disoriented in the virtual world and just wandered around. The virtual world was an open environment, so the learners may have needed more structured assignments in the virtual environment to experience its affordances.

Blogs achieved the second most social presence. This may be due to the students’ high familiarity with blogs. The high social presence may be explained, because blogs are designed to provide students with an environment to reflect and express their thoughts and feelings (self-disclosure), which can magnify the underlying reasoning (Sharma & Hannafin, 2007), further promote student-teacher interactions and help produce positive learning outcomes (Harper & Harper, 2006). In essence, wikis belonged to all participants and were the product of the whole community. Although wikis were rated the most impersonal, among the five Web 2.0 technologies, generally speaking, wikis produced medium levels of feelings. In the process of wikiing, participants can deepen their understanding “by externalizing and comparing knowledge” (Sharma & Hannafin, 2007, p. 43) with their peers and instructors.

This study had several limitations that are discussed in this section. First, many of the interpretations and assumptions of the research results were contextual and would require future research to confirm the accuracy and the appropriateness. Most of the participants were women. The study was conducted in the oriental context; the students’ feelings and perspectives may differ from other cultures. For example, a few interviewees commented instructor authority may be weakened because of blogging with students, which may not be an issue in the Western context. Future research is suggested to recruit a diversity of participants to confirm these findings. Second, based on course objectives, a few technologies, such as the micro-blogging application, were not included but may warrant investigation in future studies, when they are more suitable with the course objectives. Third, the study examined each dimension, instead of grouping all of the dimensions into two components of social presence, intimacy and
immediacy (Short et al., 1976). Future research in this area may investigate the two groups of feelings separately and explore other findings. Fourth, the present study showed that the social bookmarking site and the virtual world provided no or little feelings for the student participants. The main reason may be because the two technologies were new to the students, and they did not get immersed in the two types of environment long enough. It is suggested to have more advanced users, to duplicate the study, and to confirm the generalization of the findings. Fifth, as Grandey, Tam, and Brauburger (2002) commented, feelings or emotional states, which are intense and short-lived, are difficult to measure through self-report surveys. Conducting think-aloud protocols or using student reflective journals to monitor or track the feelings in a more objective way would be worthy of study. Finally, the SWOT elements were identified from the students’ perspectives, which may significantly differ from teachers’ perspectives. Although the participants were enrolled in a teacher education university, they were students, and the findings were based on their perspectives. Studies comparing the perspectives of practicing teachers with teacher education students may increase insight into the use of these technology tools for designing systems for learning.

A few implications emerged. First, each type of technology was designed for different purposes; different technologies can scaffold learning by providing procedural and metacognitive types of support (Sharma & Hannafin, 2007). On the one hand, several interviewees reported that the collaborative writing in wikis improved the quality of their work, yet others complained that co-writing made consensus hard to achieve and made the process inefficient. This is one example that shows no technology tool can meet all course objectives for all learners. Educators need to understand the affordances and limitations of each tool, so as to make the best use of technologies for enhancing student learning.

The SWOT analysis in the present study helped identify the features of each technology. This information can function as a proxy for directions of Web 2.0 technology integration. The findings indicated that students usually use the technologies with which they are familiar. To effectively integrate technologies in learning and teaching, educators must ensure that students have sufficient skills and practice with the learning tools. Otherwise, the tools may exacerbate students’ cognitive load and negatively impact their content learning. The SWOT analysis highlighted the strengths, weaknesses, opportunities, and threats of each tool. Note that, although the strengths and weaknesses of each technology were emphasized as internal
characteristics, the information remains pertinent when the technology is used in the education context (as external characteristics). The internal and external characteristics are not exclusive of each other and together, provide a comprehensive picture of the features. Finally, one other finding is noteworthy: the students regarded connecting and networking with their teacher as a positive experience while using technology, which educators need to thoughtfully consider. As Johnson (2010) stated, social networking is not equal to educational networking. There is a line educators should not cross (Deubel, 2009) to protect students and to protect teachers themselves from avoidable ethical problems.

CONCLUSION

This study revealed students’ perceived feelings related to different types of design attributes associated with Web 2.0 technologies and also provided the main characteristics of Web 2.0 technologies from a student perspective, which can guide educators in choosing the appropriate tool depending on the instructional objectives. To be literate in the twenty-first century means to be proficient with technologies that are used in the production and processing of text and non-text media (Mills, 2010). To promote the most effective learning for our students and prepare them with digital literacy for their future lives, educators need to effectively and creatively use a variety of technologies in the classroom (Tang & Austin, 2009). Educators need to train students to develop digital media literacy and skills across different curricula and develop competency in reflection and mental flexibility (Johnson, Levine, Smith, & Stone, 2010). Understanding how students feel using the technologies and how these technologies may be used in teaching and learning will benefit teacher educators (Ferdig, 2007). The present study strived to meet those goals. As we can see from the findings, some Web 2.0 technologies were not designed for educational purposes. In the long run, educators may need to transform “social media” into “educational media” (Johnson, 2010), yet not be driven by technologies that lack educational value.

REFERENCES


CHAPTER 13

Engagement, Emotions, and Relationships: On Building Intelligent Agents

Candace L. Sidner
Worcester Polytechnic Institute, Worcester, Massachusetts, USA

INTRODUCTION

The Collagen Collaboration project research centered on the nature of collaboration between people, where collaboration is defined as: two or more people agreeing on a shared goal and undertaking to perform it. The phenomenon of collaboration is valuable because it is inherently fascinating and, understanding collaboration is essential for building computers that can collaborate with people. Figure 13.1 illustrates the nature of the collaboration in the Collagen Collaboration Project. The human and the “smiley” agent can collaborate in the same circumstances as two human partners, that is, they can communicate with each other, can each interact with the shared objects, and each observe what the other is doing with the shared objects.

The study of collaboration in this chapter differs from the approach of an anthropologist or psychologist; computational models of collaboration benefit from insights from these fields, but they look toward the nature of the computation of collaboration. Computation demands designing the algorithms and architectures needed to create an intelligent computer agent that takes the place of a human and performs reasonably well. The agent is not exactly like a human (that is still too difficult), but participates in collaborations without holding up the human, getting in his or her way or doing odd things that keep the human partner from operating, as he or she normally would with a collaborator. Most critically, the agent must know about the task, what it can contribute, and how to proceed when it is missing information that is vital to the collaboration.

This chapter will focus on collaborations and what is needed to make an intelligent computer agent a useful participant in collaborations. It discusses how non-verbal behavior plays a role and delves into emotional expression as a non-verbal behavior in collaborations, and to what purpose it can and
must serve to collaborators. As will be seen, emotional expression provides much more than a signal of the cognitive state of the one who expresses emotions.

Collaboration concerns “doing” things, e.g., fixing something, cooking a meal together, making decisions in a group meeting about a company’s products, instructing a person on how to do a task or playing a game, such as soccer (Tambe, 1997). In all these collaborations people talk, but they also perform actions, sometimes together, but just as often, each participant acts on his or her own. Understanding collaboration requires accounting for both the joint and individual behaviors in the collaboration process. Some collaborations involve nothing more than conversation between the collaborators. For example, counseling sessions, lectures on a topic, and even conversations that are just to pass the time of day, are all collaborations. They each have a shared purpose among the collaborators and a known way of getting that purpose to come to pass. Each participant has certain tasks to do, they all know what those tasks are and they are committed to them (Cohen & Levesque, 1990).

Though collaborations are diverse in their goals, they have many common properties (Grosz & Kraus, 1996; Grosz & Sidner, 1986): a shared goal; shared beliefs among the actors about the state of the world and how to accomplish collaboration; individual intentions to perform actions needed to accomplish the shared goal; intentions that others do their part; and shared beliefs about their intentions.

Figure 13.1 Collaboration.
The theories of collaboration cited above are computational in nature. They offer constraints that can be expressed in computational terms. These theories can be and have been used to build intelligent computer agents, both virtual and robotic, that can participate in collaborations with people (generally with one human at a time). Rich, Sidner, and Lesh (2001) provide details on the collaboration agent “Collagen,” discussed here. Work by Allen et al. (2007) reports on an agent that can be taught by a person to explore information on the internet. These human–computer collaborations use dialogs between the person and the agent during the collaboration. Examples of collaborations built using Collagen, with the accompanying computer interfaces, are shown in Figure 13.2. They include: air travel planning; scheduling TV shows; instruction in the operation of a gas turbine engine; assistance with reading and writing e-mail; demonstrating a new laboratory invention; and guidance on the operation of home appliances. These interfaces depict the computer as an agent, shown either as a smiling face, a penguin robot (talking to a person), or with an icon of a human shape (in the home appliance interface). The depiction of the agent makes clear its role in the interaction with the user.

A related effort with a full-face virtual, collaborative agent, illustrated in Figure 13.3, was developed to counsel people on the adoption of new health-related behaviors, such as exercising more or eating more fruits and vegetables (Bickmore, Schulman, & Sidner, 2011).

Figure 13.2 Example of collaborative agent interfaces.
THE ROLE OF ENGAGEMENT IN COLLABORATIONS

Certain basic properties of interactions, which will be called here “engagement properties” must hold for a collaboration to take place and succeed, and these properties can be built into intelligent computer agents when they interact with people. These properties are necessary, but are not sufficient to ensure success in collaboration. Yet without them, the collaboration will either never take place or will fail. The collaboration will not be able to get started, or if it does, the lack of these properties will quickly de-motivate the collaborators from continuing to work together.

When people collaborate, they converse, both about the collaboration and about whatever else they deem relevant or appropriate to their partners. However, before they even start talking, they must engage with one another. Here, “engage” means to establish a perceived connection between two Interactors. Engagement begins with glances toward each other, facing one another, and then addressing some dialog from one to the other, often starting with a greeting. This process, the initiation of engagement, is fundamental to their collaborative interaction. All through their subsequent collaboration, the participants must maintain this perceived connection to one another, and when they decide to terminate their interaction, they must
terminate this perceived connection (Sidner, Lee, Kidd, Lesh, & Rich, 2005). Bohus and Horvitz (2009) have developed an agent that can interact with two or more people.

Much of what is necessary to initiate engagement and to maintain it happens without explicit mention of the perceived connection. Participants accomplish engagement by verbal and non-verbal means. The verbal means is through conversation, while the non-verbal includes looking, gazing, body stance, and making facial and hand gestures. Non-verbal cues, as well as verbal exchange, are evidence that the partners are engaged, even though these behaviors are not directly discussed or used solely for engagement. All of the non-verbal behaviors are serving double-duties, as each of them also allows a person to gather information and/or communicate on the conversational situation and their own mental state. Lack of these behaviors, especially looking, glancing, and taking part in the conversation, or indeed, mis-use of these, will allow the conversational partner to conclude that the other does not want to continue the interaction.

To understand these engagement behaviors better, the author and colleagues studied videotapes of people in collaborations. Four major activities, called “connection events,” that supported the engagement between the participants (Rich, Holroyd, Ponsler, & Sidner, 2010), were cataloged. These were: directed gaze, mutual facial gaze, adjacency pairs, and backchannels. In directed gaze (Kendon, 1967), one person (the initiator) looks and optionally points at some object or group of objects in the immediate environment, following which the other person (the responder) looks at the same object(s). Mutual facial gaze (Argyle & Cook, 1976) occurs when one person (the initiator) looks directly at the face of the other person (the responder), and the responder returns his or her gaze to the initiator. An adjacency pair is a two-part exchange, in which the second utterance is functionally dependent on the first, as exhibited in conventional greetings, invitations, and requests (Schegeloff, 2007). During a communication from an initiator to a responder, a backchannel (Yngve, 1970) event is when the responder directs a brief verbal or gestural communication back to the initiator. Typical examples of backchannels are nods and/or saying “uh, huh.”

Using a formal description of these connection events to indicate what each person does over a given time period, intelligent agents can keep track of connection events in conversation and assess both its own level of engagement and that of its human partner during collaborations. A robot was programmed to interpret connection events from a human partner and to generate appropriate ones when making a “tangram” (a puzzle figure) together with a person
(Holroyd, Ponsler, Rich, & Sidner, 2011). The robot had a set of policies for generating connection events, which it used to decide when to look at the human partner, when to look and point at the puzzle pieces, which referential description to use for the puzzle piece (the pink triangle vs. that one), and when to track the human partner’s hands because he or she was pointing to a puzzle piece. The robot also had an algorithm for recognizing connection events generated by its human partner. Figure 13.4 illustrates the robot pointing and looking at one of the puzzle pieces with a human partner.

Does the inclusion of engagement connection events make a difference in collaborations? Three of our studies indicate that it does. In the first study, people interacted with a penguin robot. Figure 13.5 shows the penguin robot, as it
collaborated with a person to demonstrate a laboratory invention (Sidner et al., 2005). The collaboration was required because the penguin did not have “hands” with which to manipulate the glass on the table surface to show to the user. Another version of a robot produced no “looking” gestures at all during the conversational interaction, while the robot with normal looking gestures (called the “moving robot”) held the same kind of conversation. Thus, the non-moving robot never looked at the objects that it was demonstrating to the user, nor did it look at objects when the person made use of them, but the moving robot did all these things, as well as gazed at the person’s face and tracked it, as the person moved about. Not only did people report that the moving robot had more appropriate gestures, but they held significantly longer conversations with the moving robot and looked back at it significantly more often.

In a second study (Holroyd et al., 2011) with robots that collaborated on tangram puzzles, people interacted with either a robot that produced connection events or one that did not. People offered significantly higher agreement to statements that the robot looked at the person in a natural and appropriate way, and that the person could easily tell what object the robot pointed at and looked at. A third study (O’Brien, Sutherland, Rich, & Sidner, 2011) discovered that a robot that looked at its human partner when conversing and pointed appropriately during collaboration in comparison with one that did not, received much higher approval as a collaborator from its human partners.

Affective Expression

Of all of the connection events, the one that is the most relevant for the remainder of this chapter is mutual facial gaze. When individuals gaze at another’s face, they each gain information not only about his or her attention to their partner, but also about his or her affect, simply because faces express emotions. Simply said, the face is an important source of this information. How well individuals can read facial expressions of emotion may vary. Furthermore, individuals choose to display more or less of their emotion state in their faces (as well as their bodies), depending on their situation. Individuals can also read the faces of others when there is no mutual facial gaze, but in mutual facial gaze, each agent is aware of the face providing that information.

Affect in the face, as well as the body, has been understood by many scientists (e.g., Ekman, 1993; Scherer, 1986b) as reflecting the internal emotional state of a person. This chapter explores the need for creating intelligent agents that can express affect and read it in its human counterparts.
This section explores the role that affect seems to play in human collaboration with an eye toward creating agents that can use this information and produce it as well. The term “affect” or “affective expression” will be used to discuss emotional displays because “emotion” connotes high intensity displays, whereas the full range of facial displays includes many that do not appear as intense. Indeed facial displays are often subtle, even when the underlying emotional state of the expressor is intense.

It has long been understood that “affect” plays an important role in social interaction. Matsumoto, Keltner, Shiota, O’Sullivan, and Frank (2010) argue that evolutionist approaches to emotions (beginning with Darwin) hold that facial emotions have an important social function, especially in solving social problems. Facial affect expression provides information about the expressor’s affect, as well as about his or her relationship with a person receiving the affective display. It is particularly important to gain insight into the ways in which facial expression as well as other affective devices of the body contribute to the collaboration process between two or more people.

Faces alone do not give all the critical information about affect. The tenor of the voice, including pitch, duration, and voice quality, also provide information about affect (Scherer, 1986a). Faces and voices (and bodies) work together to provide the information that humans use to discern a person’s affective state. Furthermore, people are better able to determine the affective state when the face and voice are congruent (Jaywant & Pell, 2012). Creating computer agents that can do this is an active area of research. Producing spoken utterance with affect has been an active area of research since Cahn’s (1990) work.

A great deal of scientific literature has been directed at understanding what affective expressions people can recognize (Ekman, 1993), what muscles of the face are used in which emotions (see Matsumoto et al., 2010, for an extensive list), and the dynamics of facial expression (Bassili, 1979; Scherer & Ellgring, 2007). These efforts have inspired computational scientists to develop intelligent agents that can use their faces to convey affect (using when possible, the same kind of simulated “muscles” that human faces do), and reproducing the dynamics.

Creating intelligent agents that can discern a human’s affective state and produce affect in the face, voice, and body is still a work in progress. Researchers have been creating virtual agents that can reliably produce affective expressions through the face, the body, and the voice (Gratch, Marsella, Wang, & Stankovic, 2009; Pelachaud, 2009; Wagner, Lingenfelser, Bee, & André, 2011) for some time. Expressive robot faces that may or may not bear
a close resemblance to a human face are also being explored. Figure 13.6 presents the faces for the commercial “iCat robot” (Breeman, Yan, & Meerbeek, 2005) and for the robot head Einstein created by Hanson Robotics (see: http://www.youtube.com/watch?v=vx35zMyFJ94).

Some researchers have attempted to program intelligent agents to have an affective internal state, notably the work of Gratch et al. (2009) using appraisal theory (Scherer, 2001). The ability of an intelligent agent to discern affect is more difficult, due to the challenges of visual processing to recognize

![Figure 13.6 The iCat and Einstein robots.](image-url)
affect, although research is also taking place in this area (see El-Kaliouby & Robinson, 2005; Malatesta et al., 2009; Schroder et al., 2011; Tscherepanow, Hillebrand, Hegel, Wrede, & Kummert, 2009).

Nonetheless, if such capabilities are available, will they be needed? Given an interest in collaboration, one must ask: what role does affect play for collaborators on shared goals and tasks? One way to answer this question is to ask and answer a related one: Do human collaborators express affect in their faces, with their bodies, and their voices? If they do, when and why?

**Affective Expression and Collaboration**

Our pre-scientific, everyday experiences make clear to us that we have all been in collaborations where affective displays occur. If we are building something together, and it falls apart, this unexpected event causes us and our partners to, at least, express surprise. We can imagine just how our face is: our eyes widen, our brows and eyebrows go up, and perhaps our jaw drops to open our mouths. We may voice a sound of surprise or words that express surprise. When some task we wish to finish fails, we can express our disappointment with some unhappy expression in our faces, as well as in our voices. We may frown or look exasperated. We may express our displeasure with sounds or an utterance. If we are called on to do something distasteful, we express our distaste, in mild or extreme form in our faces, voices, and bodies. There is not sufficient space in this chapter to give a catalog or calculus of just what affective displays accompany which circumstances, though many behavioral scientists and physiologists have pursued the problem of how a face shows emotion, and whether those expressions are universal across culture (see, e.g., Ekman & Friesen, 1986; Izard, 1971). Far more remains to be explored on how the mind generates and reflects outward affect, as well as the effects of individual variation. Whatever further study may reveal, it is clear that when collaborations have difficulties, we express an affective response to it.

What happens when collaborations are not failing, that is, the parts of the collaboration are proceeding smoothly? The collaboration has no distasteful chores or serious disappointments. That is, the collaboration goes on in just the way the collaborators expect. Do people need to take into account affective displays in the normal course of things in collaboration?

This question can be addressed using the common scientific approach for developing new theories and models: studying the data, in this case, human data. Three different datasets of collaborations between people were
carefully observed to determine which, if any, affective displays take place and the circumstances in which they do. These datasets were collected for other research projects, but were revisited for this chapter to determine if and how affect arises during collaborations.

The first of these videos was collected in a quasi-experimental situation, in which two people were putting together a porch swing. The participants had their own reasons for doing this task, and their interaction was videotaped, with their permission, in one of their homes. The session lasted about 45 min. The second is a series of observations of pairs of people making canapés in a laboratory setting. One person was asked to teach a second person how to make canapés, and then the “student” taught a third person. Each teaching episode took about 8 min, and the pairs knew each other. The last video dataset was also collected in a laboratory setting by videotaping three different groups, each consisting of three people who are in a sales encounter. One person is buying a cellphone, the second is the buyer’s friend who is along to give advice, and the third person is the sales person who provides information and advice about which phone to buy. Each session lasts 5-10 min. Data from the canapés dataset was used to develop the model of connection events discussed earlier in this chapter.

Affective expressions, as one might guess, occur in all these encounters. Every single person, pair or triple, has occasions on which they show affect. Some of their affective displays result from surprise, disappointment, or embarrassment, as discussed above. For example, in the porch swing task, the following exchange occurs as the partners, C and D, finish reading the instructions about how to put the swing together. When D reads the last instructions, he is expecting them to indicate that the nuts are to be tightened at that time. What he finds is just the opposite.

Porch Swing Dialog

D reads: Do not tighten nuts at this time.
D says: No more directions.
D laughs and then C laughs.
D says: I think it means tighten any loose nuts at this time. D laughs.
C says: Entirely possible.

D and C’s initial laughter is an expression of surprise at the final part of the instructions. D’s subsequent laugh seems to be amusement. C’s comment “Entirely possible” does not respond to amusement affectively, but to the
content of D’s utterance and conveys agreement on how the last part should be done. Thus, affect and task are, in this example, intertwined.

In one cellphone sales encounter presented below, the salesperson displays embarrassment when the friend (F) asks the salesperson (S) about data plans for cellphones. Because the salesperson has not carefully studied his notes about the cellphones, he does not know how to answer this question, and he stalls by saying, “That’s a good question.” His expressions just after seem to convey embarrassment, and his partners smile, but work hard not to look at him. They are as embarrassed as he is. At the end of this excerpt, he discovers the information that F was looking for.

Cellphone Sales Dialog 1

F says: Um what are the data plans offered for um …
S says: Um.
S is smiling.
S says: That’s a good question.
S then looks at the cellphone for information and finds none.
F and B (the buyer) begin smiling but look at their notes and then the cellphone counter.
B puts her hand over her smiling mouth.
S makes a wry face.
S says: Darn good question.
S smiles then looks at his information sheet.
F says: Okay.
S says: Oh, right then I just… Um the Geophone data plan first year is …

In the reminder of the instances of affect in the three datasets, participants’ expressions involve smiling, and their voices often indicate laughter and amusement. In the sales interactions, all three participants, which consist of members who did not know each other before the encounter, begin their interaction with greetings and smiles. Their affective expressions are the familiar ones that accompany greetings and express welcome in the socially conventional way.

In many occurrences of laughter and smiling, amusement provides a social function to create some sense of solidarity (Spencer-Oatey, 1996; Wheeless, 1978), that is, the sense that the participants are in this situation together. For example, in the canapés dialog below, the participants, who are engineering students, make jokes about the term canapés, about what canapés are, and about the use of lace doilies (“fancy orientation”) for the plate for serving canapés. In each case, one participant reaches out to the other with a joke about canapés. Their jokes also indirectly serve the
collaboration because they provide relevant information for the collaboration. For example, when one participant P1 responds to the jokes about canapés, he explains what they are and how one dresses up the serving platter (which is referred to by the odd term “fancification”). However, the jokes do more, by making it clear that the participants both know they are involved in an odd activity, for in general engineering students do not make canapés.

Canapé Dialog

P1: Um, Today we’re making canapés. P1 smiles.
P2: (smiling) What is a canapé?
P1: Um a canapé is (P1’s face changes to a serious look) it’s a sort of, what you see at sort of an upscale function that goes on plates and they are sort of like hors d’oeuvres type thing.
P2: (serious face) right.
P1: Um, and so, so they start off with ah, crackers, followed by cream cheese or hummus, um and then they put on a smaller topping, such as pimentos, olives or raisins or something like that. And then they are sort of put onto a plate, in a sort of fancy orientation and they are served for dinner.
P2: What’s the other stuff?
P1: Those come in, in the fancification of the plate.

Affective expressions accompany what are mildly embarrassing events. In one of the canapés encounters, one participant smiles at his partner when the partner drops an olive while trying to put it on top of the canapé. Then the partner smiles back. On another occasion, the partner cannot open a jar; the instructor participant then grins at her while noticing this. These encounters are not just to show that one participant looks ridiculous and the other is amused at the first’s expense. The fact that they look at each other and smile conveys that the observer is expressing empathy and a sense of solidarity at the situation. Empathy also conveys that the observer sees the situation as understandable. That determination is important in maintaining and building trust between the collaborators.

A related purpose of smiling and amusement occurs in one of the sales encounters. After deciding which phone to purchase, the buyer (B) asks the salesperson (S):

Cellphone Sales Dialog 2

B says: well, can I have my phone now? B laughs.
S says: Sure, I’ll ring you up. S laughs.
B and S laugh and walk away.
Their shared joke centers on the fact that they are not in a real store and that no real purchasing is going to happen, but that instead, they are all play acting. In effect, they deny that the next task is going to actually occur, and make clear what they are doing together. Their whole exchange again expresses their solidarity in dealing with the situation at hand.

Another example of affective responses occurs in both the canapés tasks and the sales task. When the participants finish, they smile at one another. These behaviors are social, a means of saying “look we’re done” and also providing the social need to acknowledge success of the group. Affective expression works both for task purposes and social ones.

In all of these encounters, a large portion of the affective displays represent what can be called “a serious working face.” The participants have tasks to do, determining how to build a porch swing from written instructions and then actually doing so, learning how to make canapés and doing so, or gathering information about various cellphones and deciding which best fits the buyer’s needs. During these efforts, when not making jokes, looking embarrassed, or the other affective expressions discussed above, the participants do not wear blank expressions on their faces. Their facial expressions reflect their attendance to the tasks at hand and signal that they are working. It seems likely that closer inspection of these periods of expression will reveal not just one kind of expression, but several expressions that change as they move from learning to making something or making a decision. These faces are not the ones normally associated with high intensity affective moments. They do, however, convey to other participants the mental state of the expressive participant, and they convey his or her involvement in the collaboration.

The Necessity for the Expression of Affect and Social Behavior in Collaboration

As the previous discussion of simple and commonplace collaborations illustrates, affective expressions reflect the function of solidarity and trust among collaborators. Affective expressions in response to problems in performing a task are useful in conveying one collaborator’s assessment of circumstances of the collaboration. One can ask whether virtual agents and robots should produce and recognize affective expression during collaborations. Further, one can ponder whether social behaviors are necessary to collaboration or just a side benefit to interaction.
Consider the following observations based on human behavior. First, recognizing and expressing affect in response to problems in collaborations contributes to the smoothness of the collaboration and getting the work done efficiently. The use of facial expressions and voice tone (as well as body expressions, which are not discussed here) provide additional ways to communicate about tasks and how they are going. People do not ignore this information or fail to recognize it in their encounters. It tells them how their partners view the current situation and gives them further insights for their own evaluations. A person who fails to provide such information requires his or her collaborator to either ask directly or to fail to recognize something about the circumstances that the partner has already ascertained. A person who fails to use affective expressions to convey his or her assessment of the situation must instead express that assessment by the semantic content of an utterance. However, early assessments of problems often do not come with a clear sense of what is wrong, just that something is, or that things are not going as expected. A collaborator who does not use the affective mechanisms available must invent language that will do the same. Using language that way is harder, and takes more time than the affective response.

Second, social interaction in collaboration may be necessary for two reasons. One rests on the undeniable fact that people are social beings who have always done their work together in social networks and social groups. The other reason concerns the ongoing requirement for assessing trust and solidarity between collaborators. Concerning humans as social beings, our Western emphasis on thought as our raison d’être would suggest that humans are first and foremost problem-solvers, tool-makers, and manipulators of all that surrounds them. In fact, virtually all human activity takes place in the context of families, tribes, clubs, companies, neighborhoods, and nations, simply said, in the context of social organizations, large, and small. That context is easy to overlook because it is so commonplace.

With all those social institutions, come conventions, rules, and knowledge for how to pursue work, how to get it done together, and how to come to value the group that has formed for the work. It also provides the knowledge of how to work together. One aspect of that knowledge is the assessment of the other person as a collaborator. Are they trustworthy? Can they be relied upon to do their part? Are they “on the same page with me,” even if they are not known to me? As the collaboration happens, collaborators assess their partners, both in what work they do and how they accomplish it, but also in the sense of whether they are going to continue to be valuable...
as partners, thus trust and solidarity are not fixed for all time, but evolve over time. As they evolve, they make it possible for the collaborators to see each other as part of a group, and to make use of whatever social knowledge that group has to support their effectiveness.

Should interactive intelligent agents supply affective responses, and if they do, for what purpose?

Work by Bickmore et al. (2005) has demonstrated that computer agents who express affective faces and who talk with human users about social matters during their collaboration and discussions of exercise tasks create a sense of alliance with their human users. Using standard psychological measures of alliance, Bickmore showed that people who interact with their “social” agents report a significantly higher sense of alliance than those who work with agents that show no affect and offer no social talk, and they report even less alliance with computer interfaces that have no agent at all.

That people feel a sense of alliance to computer agents that act socially strengthens the case for social agents. However, one may ask about the cost of completing the collaboration successfully if agents do not express affect and do not produce and assess social relationship features, such as trust and solidarity. It is essentially the same problem as having a human who cannot produce and assess affect and what it means socially. That is, if agents ignore these responses, they are missing out on factors that can influence what the person decides to do next. They miss signals about how the human is challenged or disturbed by collaboration failures, and they miss unspoken responses to their own difficulties. If agents cannot produce such responses themselves when such events occur, they slow down the collaboration in the same ways that humans would if they failed to produce them. People would not get the signals of impending problems or the indications that the agent is going to be a good collaborator or not. In sum, failing to recognize affective responses, failing to express affect about problems in the collaboration, and failing to convey social information make the agent less able to participate in the full range of factors that humans rely on in making decisions. Furthermore, as intelligent agents clearly are not people, their human partners will see their flawed abilities as evidence of their being stupid and unreliable.

RELATIONSHIPS WITH INTELLIGENT AGENTS

So far, the chapter has been hinting at an underlying but significant aspect of collaboration, namely the role of relationship among collaborators. The role of affective expression in collaboration serves not only the collaboration
itself, but also the formation and evolution of the collaborators as a group, that is, as having some relationship between the group members. Producing social behaviors from the most basic ones of being engaged in the interaction, to conveying affective expression, to using social dialog during work together are all capabilities that support the developing relationships between people. The relationships in turn support collaboration by offering each collaborator a sense that the other is going to be a team player, as the collaboration unfolds.

At present, few intelligent agents have any depth of knowledge about relationships with people. Should they have much at all? Why would it matter to a person if an intelligent agent knew about relationships? If intelligent agents collaborate with people, should they not then know enough about groups to understand what membership in a group entails? However, if a collaboration is short-lived, it may be that group relationship knowledge is not so vital.

The author is currently investigating intelligent agents that use a basic model of relationships and explicitly reason about their evolving relationship with a person. These agents will not have just short-term interactions with users. They will instead have long-term collaborations with people, over a period of months, and be present in people’s homes all the time. The agent is “always-on,” which means that whenever the agent senses the person’s presence, it makes itself available to the person (Sidner et al., 2013). In Figure 13.7, the older adult and the agent placement can be seen, as well as the actual virtual agent greeting her human partner in the morning.

The target population for interaction with the always-on agent is healthy, isolated older adults who live alone. The agent’s tasks for working with these adults are tasks that will provide a bit of day-to-day companionship as well as support the adult’s personal connections to other people, both in cyberspace and the physical neighborhood of the adult. During a typical day, the agent might talk about the day’s weather or the appointments in the adult’s daily calendar. It can play a hand of cards in a social way, with talk about how the game is going, as well as other topics, such as how a favorite sports team is doing in the current week. For connecting the adult to others, the agent acts as an exercise coach (Bickmore & Picard, 2005) to keep encouraging the adult to exercise. As a side benefit, suggesting walks, e.g., can get the older adult out and about in the local neighborhood. In cyberspace, the agent can provide the adult with the means to connect to family members and old friends via social networking tools, such as Google Hangout, a video-calling application.
Figure 13.8a provides a screen capture of the virtual agent in a conversation with a person as they play cards, and Figure 13.8b shows the equivalent robotic agent, when it has just provided the user with a list of possible activities to do. The virtual agent has just spoken the words displayed in the thought balloon. The person’s choices of what to say in response are shown below the agent in clickable boxes. The state of the card game is shown on the left of the figure.

The always-on agent reasons about its relationship with its human partner. It judges when the relationship moves from being one of strangers to being acquaintances, to being in a relationship that one might call...
companions (Coon, Rich, & Sidner, 2013). As strangers, the agent will undertake only the most basic social encounters, such as discussing weather or playing games. As the relationship changes, the agent will offer to participate in scheduling events in the user’s calendar, telling stories, and connecting the adult to friends and family in cyberspace. The agent waits until it determines it has become “companionship familiar” with the person before
undertaking to discuss more personal topics, such as exercise or asking the user to tell a personal story.

An always-on agent can be a virtual agent that appears on a computer screen, as shown in Figure 13.8a, but for a small number of homes, the agent will be the small robot with a head and face, as shown in Figure 13.8b. The agents will produce some affective expressions. It remains an unresolved challenge for this project to develop algorithms that will determine the types of affect to use and the timing of any affective expressions that always-on agents will display in their daily tasks with their human partner.

CLOSING THOUGHTS

Intelligent agents will soon be part of our everyday lives. While this claim sounds as if it is imaginary science fiction, this chapter has explored some of the research that will make it possible. Agents will be able to engage and collaborate with us, express and interpret affective expressions, and will be able to form relationships with us.

ACKNOWLEDGMENTS

My thanks to my collaborators on much of the work reported here: Charles Rich, Tim Bickmore, Brett Ponsler, Will Coon, Aaron Holroyd, Bahador Nooraei, Dan Schulman, Mohammad Shayganfar, Morteza Behrooz, Lazlo Ring, and Zessie Zhang.

This work is supported in part by the National Science Foundation under awards IIS-0811942 and IIS-1012083. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

REFERENCES


Note: Page numbers followed by f indicate figures and t indicate tables.

**A**

Activation-deactivation adjective check list (AD-ACL), 216

AERA. See American Educational Research Association (AERA)

Affect-aware educational technologies, 57

Affectiva, 98

Affective computing, 146–147

Affective tutoring systems (ATS), 75

Affect-sensitive computer technology automated classification, affect states

AutoTutor log file, 9

body movements, 9

computational linguistics, 9–10

dialog cues, 10

discourse interaction history, 9

emotional experience, 8–9

facial actions, 9

noninvasive sensing methods, 8

semantic match algorithms, 9–10

computer system

confusion, 10–11

ITS, 11

learning gains, 10–11

motivation, 11

learners

achievement emotions, 5–6

automated computer tracking, 7

boredom, 6–7

cognitive disequilibrium model, 6–7

Cohen’s kappa, 7–8

demot-aloud task, 7–8

engagement and possibly flow, 6–7

human judgments, 7

learning-centered emotions, 5–6

moment-to-moment emotions, 5–6

online judgment task, 7–8

trained judges, 7–8

American Educational Research Association (AERA), 91

Analysis of variance (ANOVA), 218, 220, 220t, 246–247

Animated pedagogical agents.

See also Virtual avatars (VAs)

agent behaviors and learner emotions, 57

agent design and emotion, 231–232

benefits of

context-specific learning strategies, 227

gaze and gestures, 228

interactive demonstration, 227

as learners’ attentional guides, 228

learners’ motivation, positive impact on, 227

learning materials, elements of, 227

as navigational guides, 227–228

non-verbal feedback and conversational signals, 228

physical tasks, 227

cracter building qualities, 230–231

computer-based instruction

(see Computer-based instruction)

textual relevance, 230

definition, 226

design and implementation of, 49–50

tions, tutorial interactions

computer tutors, 50–51

emotion and empathy, 51–52

human tutors, 50–51

politeness, 51

guided activity principle, 141–142

history, 48–49

human-computer interaction

as co-learner, 229–230

tutoring systems architecture, 229

virtual world, assistance in, 229

impacts of

affective outcomes, 54–55

cognitive outcomes, 53–54

motivation, interest and self-efficacy, 55–56

self-perceptions and feelings, learning, 55

learning gains, 48

learning-related conversations, 47–48
Animated pedagogical agents (Continued)
motivational agents, 230–231
nonverbal fidelity, 57
real relationships, 58
research evidence
impact of, 234
learners’ physical activity levels, 232
negative emotion, 233
neutral condition, 233
positive emotion, 233
psychological theories of emotion, 232–233
relational agent and non-relational agent, 233
tutorial lesson, 233–234
role of
advisory-style agent, 226–227
assistant-style agent, 226–227
dialogs and interactions, 227
use of, 226–227
social interactions, 47–48
verbal and non-verbal communication, 230–231
Artificial intelligence (AI) techniques, 89–90
Ashametrics, 98
ATS. See Affective tutoring systems (ATS)

B
Betty’s Brain, 5
Blogs, 242
external opportunities, 259
external threats, 259
internal strengths, 253
internal weaknesses, 253
participants’ comfort levels, 249t
participants’ frequency of, 250–251t
personal-impersonal dimension, 248
social presence, students’ sense of, 248, 252t, 253f
SWOT analysis, 254t
Blue Eyes camera system, 95

C
CBLEs. See Computer-based learning environments (CBLEs)
Center for the Study of Adult Literacy (CSAL), 3–5, 18–20
CERT. See Computer Expression
Recognition Toolbox (CERT)
Circumplex model, 64–65
Cognitive theory of multimedia learning (CTML), 145–146
Coh-Metrix, 104
Collaboration, 274f
computational models of, 273
definition of, 273
human–computer collaborations, 275
intelligent computer agents (see Intelligent agents, collaboration)
joint and individual behaviors, 274
shared goal and shared beliefs, 274
Collagen Collaboration project, 273
Computer-based instruction
as co-learners, 229–230
context-specific learning strategies, 227
emotion-like qualities, 232
learners’ motivation, positive impact
on, 227
learning materials, elements of, 227
tutoring system architecture, 228–229
verbal and non-verbal communication, 230–231
virtual world, assistance in, 229
Computer-based learning environments (CBLEs)
artificial intelligence (AI) techniques, 89–90
emotions
affective grid, 92
basic emotions, 92–93
cognitive-affective hybrids, 92–93
components, 92
measurement techniques, 89–90
interdisciplinary methods, 89
learners’ emotional states
body posture, 96–97
facial expressions, 93–96
log-files, 102–105
physiological patterns, 97–99
self-report measures, 99–102
multi-method emotion classification
automated learning companion, 106–107
AutoTutor, 106
facial expression coding, 105
JavaTutor, 106–107
meta-analysis, 107–108
self-report measures, 105
skin conductance bracelets, 106–107
user-dependent and user-independent classifiers, 108
Wayang Outpost, 106
theoretical and analytical considerations, 108–109
Computer Expression Recognition Toolbox (CERT), 95
Computer literacy classes, 213
Computer-mediated communication (CMC), 163–164, 241
Confusion matrix
offline testing, 38t
real-time experiment, 39t
Control-value theory, 201–202
Control value theory of achievement emotions, 143–145
CSAL. See Center for the Study of Adult Literacy (CSAL)
Cyberspace, 289–292

D
Delicious bookmarking tool, 243, 245, 257t, 261–262
Design-A-Plant, 229
Dialogs
active learning prospects, 16
anticipated misconceptions, 12
assigning texts, student’s ability, 15
AutoTutor, 12
categories
answers, 13
assertions, 13
corrections, 13
hints, 12–13
mini-lectures, 13
off-topic comments, 13
prompt completions, 13
prompts, 13
pumps, 12–13
short feedback, 12–13
summaries, 13
discourse patterns, students, 14
floor management, 13–14
human and computer tutoring, 12–13
human-computer collaborations, 275
social dialog, 288–289
students choices, 14–15
summarization, 16
supportive short feedback, 15–16
Digital media, emotional design.
See Emotional design, digital media
Discrete Fourier transform (DFT), 35

E
Ekman’s program of research, 92–93
Electrocardiogram (ECG), 97, 106
Electroencephalography (EEG), 8, 97–98, 107–108, 110
Electromyogram (EMG), 97, 106
Electronically mediated communication (EMC), 163–164
asynchronous EMC, 163–164, 173
emoticons (see Emoticons)
synchronous EMC, 163–164, 173
EMC. See Electronically mediated communication (EMC)
Emoticons, 164
bounded learning communities, 171–173
communication, improvement of facial expressions, 168–169
illocutionary force, 168–169
miscommunication and misunderstanding, 169–170
nonemotional sentiments, 168–169
punctuation effect, 169
second language learners, 169
text-based messages, 169
“community of inquiry” model, 172
critical feedback, 177
expectations for, 177
and group differences, 166–167
keystroke-based, symbolic emoticons, 165–166
limitations, 174–175
paralanguage, 172
personalized feedback, 177
and social contexts, 167–168
social presence, 170–171, 175–176
Emoticons (Continued)
student engagement and achievement, online courses, 175
teaching presence, 176

Emotion
action readiness, 131–132
affective quality, 132–133
and cognition, 134–135
as coping response, 131–132
differential emotions theory, 134
emotional interpretations, 131–132, 134
emotion schemas, 131–132, 134, 149–150
mood, 133
response types, 132
role of, 184–186

Emotional communication, 52
Emotional design, digital media
information representation
baby-face bias, 138
follow-up study, 138–139
game characters, 139, 140f
visual design, learning materials, 138–139
interaction design
animated pedagogical agents, 141–142
INTERACT model, 139, 140f
situational interest, 140–141
research agenda for
consumer grade tools, 154–155
HRV, 154–155
learners’ emotion, impact of, 152–153
learning outcomes, 153–154
physiological measures, 154
substantial intrapersonal and interpersonal variability, 156
Emotional interpretations (EIs), 131–132, 149
Emotional Multi-Agents System for Peer-to-peer E-Learning (EMASPEL), 71
Emotional scaffolding
constructive emotions, 203
definition of, 203
explicit emotional scaffolding, 204t, 208t
implicit emotional scaffolding, 204t, 208t
sociocultural theory, 202–203

Standford Teacher Education Program, 203
unconstructive emotions, 203
and VAs (see Virtual avatars (VAs))

Emotion icons. See Emoticons

Emotion schemas, 131–132, 134, 149–150
Emotions in learning, 135–137
academic challenges, 136–137
affective computing, 146–147
cognition in education
caring attitude, 185–186
competence, 185
parenting and perceived care, 186
teacher support, 185
TIMSS exam, 184–185
control value theory of achievement emotions, 143–145
CTML, 145–146
digital media, emotional design in
(see Emotional design, digital media)
educational research, 136
impact of, 136
integrated cognitive–affective model of multimedia learning, 147–151
intrapersonal multiplicity, 137
self-regulation, 137
stereotype threat, 151
technology environment, students’ learning in (see Technology-mediated learning)

F
Facebook, 165–166, 242, 245, 256t, 260–261
FaceReader (5.0), 94
Face-saving method, 51, 55
Facial action coding system (FACS), 93–96
Facial expressions, CBLES
automatic facial expression programs, accuracy of, 94–95
AutoTutor, 93–94
Blue Eyes camera system, 95
CERT, 95
disadvantages of, 95–96
FaceReader (5.0), 94
FACS, 93
human coders, 93–94
learner-centered emotions, 93–94
MetaTutor, 94
FACS. See Facial action coding system (FACS)
Finite impulse response (FIR) filter, 35–36
Focus-group interviews, 246–253
Folksonomy, 243

G
Google Blogger, 242, 245, 253, 254f, 259
Graphic user interface (GUI), 29–30

H
Heart-rate variability (HRV), 154–155
Honestly significant difference (Tukey HSD) follow-up procedure, 218
Human-computer interactions (HCI), 27–28, 142, 144–148
Human-robot interaction (HRI), 183–184, 195

I
iDRIVE, 5
Instructional Material Motivation Survey (IMMS), 216
Integrated cognitive affective model of learning with multimedia (ICALM), 147–151, 150f
Intelligent agents, collaboration
collaborative agent interfaces, 275f
engagement role in
adjacency pair, 277
affective expression, 279–286
appraisal theory, 281–282
backchannel, 277
directed gaze, 277
initiation of, 276–277
laboratory invention, 278–279
mutual facial gaze, 277
normal looking gestures, 278–279
quasi-experimental situation, 283
social behavior, 286–288
tangram puzzles, 277–279
health change behaviors, counseling on, 276f
relationships with
always-on agent, 289–292, 290–291f
cyberspace, 289–292
robotic agent, 290, 291f
social dialog, 288–289
virtual agent, 290, 291f, 292
Intelligent tutoring systems (ITSs)
adult readers, 3
affect-sensitive computer technology, 5–11
challenges and limitations, 20–21
computer applications, 3–4
conversational agents, 3–4
dialogs and trialogs, 11–20
natural language interactions, 48
reading comprehension, 4
iSTART, 5
ITSs. See Intelligent tutoring systems (ITSs)

K
Keepon robot, 121–122
Kinect depth cameras, 96

L
Learner’s positive affect
affect and cognition, 115–116
humanoid robots
designers and children, 122–123
educational robotics, 119–120
humanoid robot Atti, 121–123
LEGO Mindstorms, 119–120
robots and affect, 120
online videos
compensatory strategies, 124–126
instructor-learner interaction, online learning, 123–124
virtual peers
affective computing, 117–118
affective role models, 118–119
computer-assisted tutoring, 116–117
educational and commercial purposes, 116–117
learning tasks, 116–117
virtual peer affect, 117–118
Learning interest, VAs. See Virtual avatars (VAs)
Likert scale, 215–216
Linear prediction-based cepstral coefficients (LPCC), 32–33, 36–37
Linguistic Inquiry and Word Count Tool (LIWC), 104
Log-files
context feature mining, 103–104
language and discourse feature analyses, 104–105
LPCC. See Linear prediction-based cepstral coefficients (LPCC)

M
MathGirls, 119
Math test, 101–102
Max-wins voting strategy, 37
Mel-frequency cepstral coefficients (MFCC), 32–33, 36–37
MFCC. See Mel-frequency cepstral coefficients (MFCC)
Mutual facial gaze, 277, 279

O
OCC model, 66–67, 67f
Online learning
adaptive speaker-independent model, 44–45
affective and cognitive functions, 29
application
challenges, 42
emotion detection, use of, 40–41
experiment scenario, 43
negative emotions, 43
positive emotions, 43
results, 43
speech data, 43
course delivery modes, 28
and emoticons (see Emoticons)
emotional states, 27
experiments
offline experiment, 38–39
real-time recording and recognition, 39–40
HCI technology, 27–28
individual interest, 205, 206f
feeling-related valences, 206–207
value-related valences, 206–207
pre-processed database, 27–28
real-time speech emotion recognition system, 30–38
situational interest, 205, 207–209
emerging individual interest, 207, 208t
maintained situational interest, 207, 208t
triggered situational interest, 207, 208t
well-developed individual interest, 207, 208t
VAs (see Virtual avatars (VAs))
Operation ARA, 5

P
PBWorks, 242, 245, 255t, 259–260
Pedagogical agent. See also Virtual avatars (VAs)
artificial emotions
affective computing, 72–73
consciousness and artificial feelings, 74
flexible and rational decision-making, 72
reactions/momentary emotions, 73
cognitive tasks, 63
discrete approach, 64
educational software, 80–84
economic agent, architecture of circle of affect, 77, 77f
convolutional agents, 75
educational benefits, 76–77
e-learning expenditure, 75
embodied animated character, 77
human-like behavior, 75
intelligent/affective tutoring system, 75–76
moods, 78
partial grading, 75
teaching/learning process, 76
verbal encouragement/warnings, 74
facial expressions, basic emotions, 81t
intelligent tutoring systems, 64
learning process, student’s emotions
AutoTutor group, 68
cognitive theories, 68
cumulative knowledge, 70
EECA, 71
EMASPEL, 71
facial expressions, 71
four-quadrant learning spiral model, 68, 69f
Kort’s Learning Spiral, 71–72
OCC theory, 68
SECs system, 71
task emotion identification, 71
tutoring agent, 72
un-learning and positive affect, 70
moods, mathematical rules, 79
quiz educational software, 84
satisfaction mood, sanguine temperament, 83
three-dimensional approach
OCC model, 66–67, 67
Plutchik’s multidimensional approach, 65, 66
Turtur’s circumplex model, 64–65
Perceptual linear prediction (PLP) cepstral coefficients, 32–33
Persona effect, 54–55
Physiological sensors, 99
Picard’s MIT-based research group, 98
Plutchik’s multidimensional approach, 65, 66
Pre-emphasis filter, 35–36
Prime Climb, 68, 99–100
ProQuest, 165

R
Really Simple Syndication (RSS), 253
Real-time speech emotion recognition system
GUI, 30
major functions, 30–31
methods and procedure
acoustic features, 36–37
emotion frequency, statistics analysis of, 38
feature extraction and model learning, 32
flowchart, 33
signal pre-processing, 35–36
speech segmentation, 35
speech signals, 32
supervised classification, 32
SVMs-based learning model, 37
voice activity detection, 32–34
Robot-based learning
affection and gratitude, 192–193
caring emotions, 193
displacement of emotions, 191–192
education curricula, 188–189
negative emotions, mathematics learning, 193–194
Turtle geometry, 190
Robots
always-on agent, 290, 291
Einstein robot, 281
humanoid robots, 192–193
iCat robot, 281
in K-12 education, 188–189
moving robot, 278–279
non-moving robot, 278–279
tangram puzzles, 277–279
Russell’s circumplex model of affect, 64
S
Science, technology, engineering and mathematics (STEM), 4, 119–120
Second Life®, 243, 245, 258, 262–263
Self-determination theory (SDT), 143, 145
Self-regulated learning, 55
Similarity-attraction theory, 118
Social agency theory, 54–55, 57
Social behavior, 286–288
Social bookmarking, 243, 265
external opportunities, 262
external threats, 262
internal strengths, 261–262
internal weaknesses, 261–262
participants’ comfort levels, 249
participants’ frequency of, 250–251
social presence, students’ sense of, 248, 252
SWOT analysis, 257
Social contexts, learning, 115–116
Social interaction, 280, 287
Social media. See Web 2.0 technologies
Social modeling theory, 118
Social networking, 242, 265
external opportunities, 261
external threats, 261
immediate-non-immediate dimension, 248
interactive-non-interactive dimension, 248
Social networking (Continued)
internal strengths, 260–261
internal weaknesses, 260–261
participants’ comfort levels, 249
participants’ frequency of, 250–251
sociable-unsociable dimension, 248
social presence, students’ sense of, 248, 252, 253
SWOT analysis, 256
warm-cold dimension, 248

Social presence
blogs, 248, 252, 253
definition of, 240
emoticons, 170–171, 175–176
intimacy and immediacy, 266–267
learning interest, VA (see Virtual avatars (VAs))
significance of, 239–240
social bookmarking, 248, 252, 253
social networking, 248, 252, 253
virtual world, 248, 252, 253
wikis, 248, 252, 253
Standford Teacher Education Program, 203
STEM. See Science, technology, engineering and mathematics (STEM)
Stimulus evaluation checks (SECs), 71
Strengths, weaknesses, opportunities and threat (SWOT) analysis, 246–247, 267–268
blogs, 254
social bookmarks, 257
social network, 256
virtual world, 258
wikis, 255
Support vector machines (SVMs), 32, 37

T
Tactical Language and Culture System, 5
Technology-mediated learning
budgets, 186
decision-making process, 187
displacement of emotions, 191–192
educational technologies, 187
fear of learning, 189–191
learning experiences, 187–188
microcomputers, 187
personal computer, 187
robotics technology (see Robot-based learning)
The International Handbook of Emotions in Education, 91
Trends in International Mathematics and Science Study (TIMSS), 184–185

Trialogs
adult learners, 17
collaborative problem-solving mode, 18–20
CSAL’s AutoTutor, 18–20
designs, 17
example trialog, 19
facial expressions, 20
grill-test mode, 18–20
helping mode, 18–20
Jeopardy, 18–20
lecture mode, 18–20
motivation and learning, 20
pedagogical agent design, 16–17
self-efficacy and self-esteem, 18–20
short feedback, 18
vicarious learning designs, 17
Turtle geometry, 190, 194

V
VAD. See Voice activity detection (VAD)
Virtual avatars (VAs)
credible persona, 209–210
data analysis, 218–220, 219–220
definition of, 204
individual interest
feeling-related interest, 215–216
value-related interest, 215–216
instructional design, 208
instructor-like persona, 209
vs. instructor-provided scaffolding, 204–205, 205
interest-inducing messages, 210–211, 219–220, 220–221
companion role, 214
computer literacy class students, 213
instructor role, 214–215
no interest-inducing message, 215
research design, 213
text message, 213–214, 215

Index
learners’ interest and academic emotion, 204
media features, 209–210
person-like persona, 209
positive effects on, 204
situational interest
  arousal measurement, 216
  attention measurement, 216
  participants’ involvement, 216
virtual agent scaffolding model, interest
development, 212f
  contingency, 211
  fading, 211
  transfer of responsibility, 211
voice feature, 209–210
web-based course material, 217–218
Virtual world, 243
  external opportunities, 263
  external threats, 263
  internal strengths, 262–263
  internal weaknesses, 262–263
  participants’ comfort levels, 249t
  participants’ frequency of, 250–251t
  social presence, students’ sense of, 248, 252t, 253f
  SWOT analysis, 258t
Voice activity detection (VAD), 32–34

W
Web-logs. See Blogs
WebQuest project, 244–245
Web 2.0 technologies
  awareness of, 243–244
  blogs (Google Blogger) (see Blogs)
  data analysis,
    246–247
  data collection, 245–246
  feelings, assessment of, 246
  focus-group interviews, 246
  inquiry-based learning technique,
    244–245
  interactive–non-interactive dimension,
    248, 265
  levels of learning and student perspectives,
    243–244
  personal–impersonal dimension, 248, 264–265
  qualitative results, 248–263
  quantitative results, 247–248
  in learning process, 251t
  participants’ comfort levels, 249t
  participants’ frequency of, 250–251t
  research participants, 244
  sensitive–insensitive dimension, 248, 264–265
  sociable–unsociable dimension, 248, 264–265
  social bookmarking (Delicious) (see Social bookmarking)
  social networking (Facebook) (see Social networking)
  social presence, students’ sense of, 248, 252t, 253f
  technology integration, 246–247, 267–268
  virtual world (Second Life®) (see Virtual world)
  warm–cold dimension, 248, 265
WebQuest project, 244–245
wikis (PBWorks) (see Wikis)
Wiki, 242
  external opportunities, 259–260
  internal strengths, 259
  internal weaknesses, 259
  participants’ comfort levels, 249t
  participants’ frequency of, 250–251t
  perceived external threats, 260
  social presence, students’ sense of, 248, 252t, 253f
  SWOT analysis, 255t
Wizard of Oz method, 122
Working memory, 145, 150