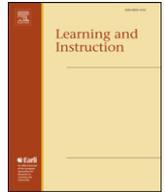


Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Dynamics of affective states during complex learning

Sidney D'Mello*, Art Graesser

202 Psychology Building, The University of Memphis, Memphis, TN 38152, USA

ARTICLE INFO

Article history:

Received 29 August 2010

Received in revised form

3 October 2011

Accepted 4 October 2011

Keywords:

Affective states

Affective dynamics

Emotions

Confusion

Cognitive disequilibrium

ABSTRACT

We propose a model to explain the dynamics of affective states that emerge during deep learning activities. The model predicts that learners in a state of engagement/flow will experience cognitive disequilibrium and confusion when they face contradictions, incongruities, anomalies, obstacles to goals, and other impasses. Learners revert into the engaged/flow state if equilibrium is restored through thought, reflection, and problem solving. However, failure to restore equilibrium as well as obstacles that block goals trigger frustration, which, if unresolved, will eventually lead to boredom. The major hypotheses of the model were supported in two studies in which participants completed a 32–35 min tutoring session with a computer tutor. Their affective states were tracked at approximately 110 points in their tutoring sessions via a retrospective affect judgment protocol. Time series analyses confirmed the presence of confusion–engagement/flow, boredom–frustration, and confusion–frustration oscillations. We discuss enhancements of the model to address individual differences and pedagogical and motivational strategies that are inspired by the model.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Efforts to learn difficult subject matter at deeper levels of comprehension involve a complex coordination of cognitive processes and affective states. Cognitive processes that underlie inference generation, causal reasoning, problem diagnosis, conceptual comparisons, and coherent explanation generation are accompanied by affective states such as irritation, frustration, anger, and sometimes rage when the learner makes mistakes, struggles with troublesome impasses, and experiences failure. On the other hand, positive affective states such as flow, delight, excitement, and eureka are experienced when tasks are completed, challenges are conquered, insights are unveiled, and major discoveries are made. Simply put, emotions are systematically affected by the knowledge and goals of the learner, as well as vice versa (Mandler, 1976, 1999; Stein & Levine, 1991). Cognitive activities such as causal reasoning, deliberation, goal appraisal, and planning processes operate continually throughout the experience of emotion. For example, flexibility, creative thinking, efficient decision-making, and conceptually-driven relational thinking have been linked to positive affect, while negative affect has been associated with narrower localized attention and stimulus-driven

referential processing (Clore & Huntsinger, 2007; Fielder, 2001; Fredrickson & Branigan, 2005; Isen, 2008; Schwarz, in press).

The inextricable link between affect and cognition is a fundamental assumption adopted by the major theories of emotion. Although the contemporary theories of emotion (Barrett, 2009; Frijda, 2009; Izard, 2007; Russell, 2003; Scherer, 2009) convey general links between cognitive processes and affective states, they do not directly explain and predict the sort of emotions that occur during complex learning (described below in *Scope of the Model* subsection), such as attempts to master physics, biology, or critical thinking skills. Fortunately, theoretical frameworks that predict systematic relationships between affective and cognitive states during complex learning are beginning to emerge in the fields of psychology (Deci & Ryan, 2002; Dweck, 2002; Immordino-Yang & Damasio, 2007), education (Buff, Reusser, Rakoczy, & Pauli, 2011; Csikszentmihalyi, 1990; Huk & Ludwigs, 2009; Linnenbrink, 2006; Meyer & Turner, 2006; Schutz & Pekrun, 2007), and artificial intelligence in education (Calvo & D'Mello, 2011; Conati & Maclaren, 2009; Forbes-Riley & Litman, 2010). Emotion research in education during the last century mainly focused on test anxiety or motivation-based traits, whereas researchers in the last decade have begun to investigate a much broader set of *academic emotions*.

Pekrun (2010) has classified academic emotions into four categories that include achievement emotions, topic emotions, social emotions, and epistemic emotions. *Achievement* emotions (e.g., contentment, anxiety, and frustration) are linked to learning

* Corresponding author. Tel.: +1 901 378 0531.

E-mail addresses: sdmello@memphis.edu, sidney.dmello@gmail.com (S. D'Mello).

activities (e.g., homework, taking a test) and outcomes (e.g., success, failure), whereas *topic* emotions are aligned with the learning topic (e.g., empathy for a protagonist while reading classic literature). On the other hand, *social* emotions such as pride, shame, and jealousy are not directly related to the topic but reflect the fact that educational activities are socially situated. Finally, *epistemic* emotions arise from cognitive information processing, such as surprise when novelty is encountered or confusion when the student experiences an impasse. According to the control–value theory, these academic emotions arise from cognitive appraisals of control over the learning task and value in the learning activity, with reciprocal connections between the emotions, their antecedents, and their consequents (Pekrun, 2006).

The emerging research on student emotions in classrooms focuses on a broad array of affective responses that are elicited in a number of learning contexts. Other research has focused on a more in-depth analysis of a smaller set of emotions that arise during deep learning in more restricted contexts and over shorter time spans, from 30-min to 1.5 h (Baker, D'Mello, Rodrigo, & Graesser, 2010; Conati & Maclaren, 2009; Forbes-Riley & Litman, 2010; Rodrigo & Baker, 2011; Woolf et al., 2009). These learning contexts include a multitude of computer environments, such as preparation for high-stakes test taking, problem solving, reading comprehension, and essay writing. The emotions that appear to be prominent in these learning sessions include boredom, engagement/flow, confusion, frustration, anxiety, curiosity, delight, and surprise (see Calvo & D'Mello, 2011).

The definitions of most of these learning-centered affect states are well known, but confusion and engagement/flow require some clarification. Although most would agree that confusion is not a 'basic' emotion (e.g., anger, disgust, sadness, fear) (Ekman, 1992), there is some debate about classifying confusion as either an emotion (Keltner & Shiota, 2003; Rozin & Cohen, 2003; Silvia, 2009), an epistemic affective state (Pekrun, 2010), or a mere cognitive state (Clare & Huntsinger, 2007). The present paper considers confusion to be an epistemic affective state and refers to the entire set of states under consideration in this article (boredom, confusion, etc.) as affective states or cognitive-affective mixtures instead of emotions *per se*.

Engagement/flow is a cognitive-affective state that sometimes has a short time span, but at other times forms part of Csikszentmihalyi's (1990) conception of flow (Baker et al., 2010). It is important to point out that a learner can be engaged without necessarily experiencing flow; for example, being engaged in order to avoid failure when one is anxious. Quite different from this form of engagement, we conceptualize engagement/flow as a state of engagement with a task such that concentration is intense, attention is focused, and involvement is complete. However, it need not involve some of the task-related aspects which Csikszentmihalyi (1975) associates with flow, such as clear goals, balanced challenge, or direct and immediate feedback. It also may not involve some of the aspects of Csikszentmihalyi's conceptualization that refer to extreme intensity to the extent that there is time distortion or loss of self-consciousness. Such intense flow experiences would be welcome, but are rare in the learning environments under investigation in the present study.

The identification of the affective states that occur during learning is critical, but it could be argued that merely knowing *what* states occur has limited utility. What is missing is a specification of *how* these states evolve, morph, interact, and influence learning and engagement. An analysis of mood states during a learning session will not suffice, because states such as confusion, frustration, and delight arise and decay at much faster timescales (a few seconds) compared to moods (several minutes or a few hours) (Ekman, 1984; Rosenberg, 1998). Furthermore, an analysis that

focuses on whether a learner is generally in either a positive or negative mood during an *entire* learning session is also unsatisfactory, because learners oscillate between positive and negative states throughout the session. What is required is a fine-grained analysis of the rapid dynamics of both positive and negative affective states that naturally occur during effortful learning activities. The present paper presents such a model and reports two studies that empirically test its major predictions.

2. A model of affective dynamics during learning

The proposed model highlights the critical role of cognitive disequilibrium in driving deep learning and inquiry. Cognitive disequilibrium is a state of uncertainty that occurs when an individual is confronted with obstacles to goals, interruptions of organized action sequences, impasses, contradictions, anomalous events, dissonance, incongruities, unexpected feedback, uncertainty, deviations from norms, and novelty. The importance of cognitive disequilibrium in learning and problem solving has a long history in psychology that spans the developmental, social, and cognitive sciences (Berlyne, 1978; Chinn & Brewer, 1993; Collins, Warnock, Aiello, & Miller, 1975; Festinger, 1957; Graesser & Olde, 2003; Laird, Newell, & Rosenbloom, 1987; Miyake & Norman, 1979; Piaget, 1952; Schank, 1999). The notion that cognitive disequilibrium extends beyond cognition and into emotions has also been acknowledged and investigated for decades (Festinger, 1957; Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005; Lazarus, 1991; Mandler, 1976; Piaget, 1952). What is less clear, however, is the trajectory of cognitive-affective states that are spawned by cognitive disequilibrium and how these trajectories impact learning and problem solving. It is this trajectory of affective states that is the focus of the present paper.

The proposed model posits that the complex interplay between external events that trigger impasses and obstacles that block goals, coupled with goal appraisal (Mandler, 1999; Stein & Levine, 1991), cognitive disequilibrium (Chinn & Brewer, 1993; Graesser, Lu, et al., 2005; Graesser & Olde, 2003; Piaget, 1952), and impasse resolution (Siegler & Jenkins, 1989; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003), is the key to understanding the affective states that underlie complex learning. The model is presented in Fig. 1 in the form of a state transition network. The nodes (circles) in the figure represent the affective states (in parentheses) and their presumed causes (in bold). Links represent situations that trigger transitions between the different states. Solid links represent primary components of the model, whereas other (dashed) links are secondary.

The model assumes that learners are typically in a prolonged state of either (a) engagement/flow as they pursue the superordinate learning goal of mastering the material in the learning environment or (b) disengagement (boredom) when they abandon pursuit of the superordinate learning goal. When learners are deeply engaged, they attempt to assimilate new information into existing knowledge schemas. When new or discrepant information is detected, attention shifts to the discrepant information, the autonomic nervous system increases in arousal, and the learner experiences a variety of possible affective states, depending on the context, the amount of change, and whether important goals are blocked (Mandler, 1984, 1999; Stein & Levine, 1991).

Learners experience cognitive disequilibrium when they are confronted with a contradiction, anomaly, system breakdown, or error, and when they are uncertain about what to do next (Carroll & Kay, 1988; Graesser, Lu, et al., 2005; Graesser & Olde, 2003; Siegler & Jenkins, 1989; VanLehn et al., 2003). Confusion is a key signature of the cognitive disequilibrium that occurs when an impasse is detected (Link 1A). Learners must engage in effortful problem solving activities in order to resolve the impasse and restore

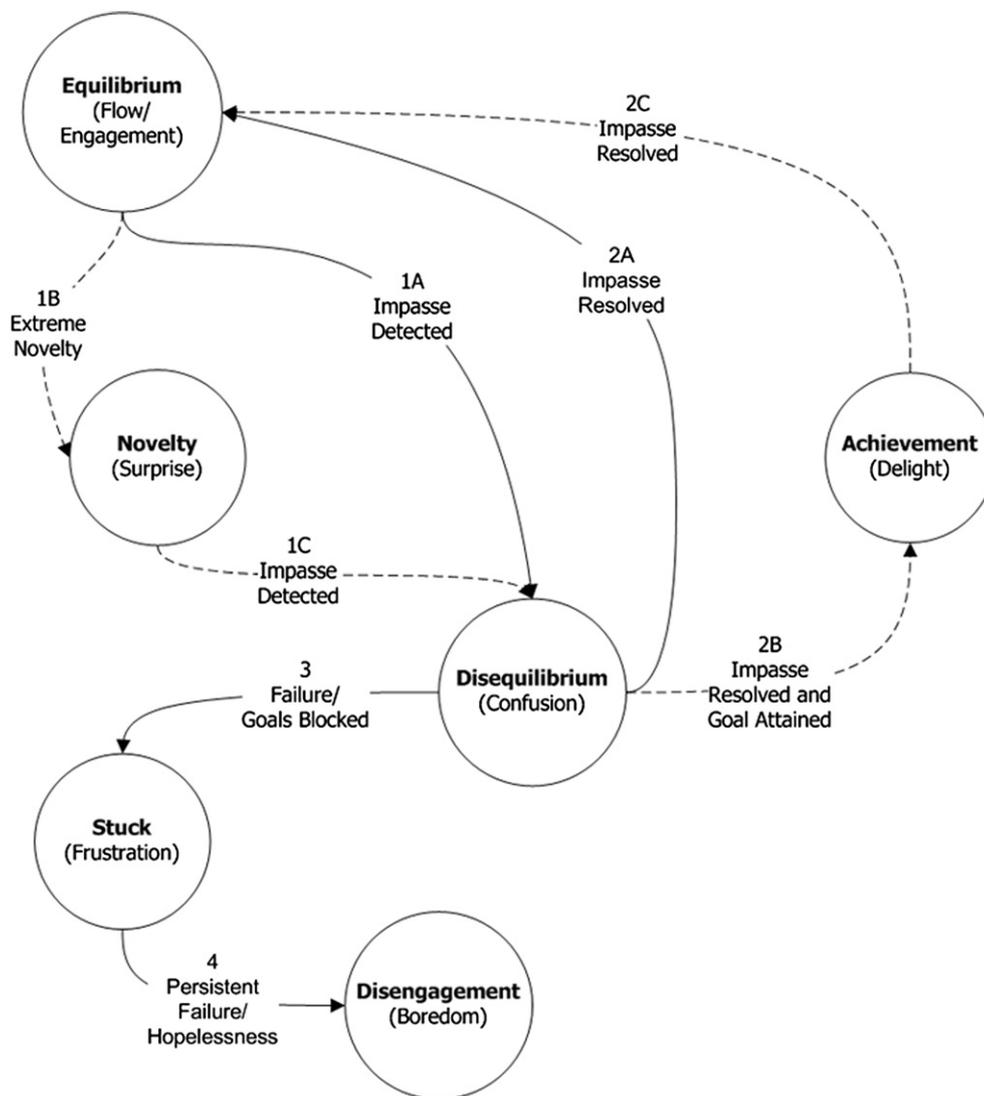


Fig. 1. Hypothesized model of affect dynamics.

equilibrium. The learner initiates a subgoal of resolving the impasse through effortful reasoning and problem solving. Equilibrium is restored when the source of the discrepant information is discovered and the impasse is resolved (Link 2A).

The above form of productive confusion associated with impasse resolution can be contrasted with hopeless confusion. Hopeless confusion occurs when the impasse cannot be resolved, the student gets stuck, there is no available plan, and important goals are blocked. The model hypothesizes that learners will experience frustration in these situations (Link 3). Furthermore, consistent with forced-effort theories of boredom (Larson & Richards, 1991; Robinson, 1975), persistent frustration may transition into boredom, a crucial point at which the learner disengages from the learning process (Link 4).

The model also hypothesizes some secondary links involving surprise and delight (dashed links in Fig. 1), in addition to the four major links (Links 1A, 2A, 3, and 4). Link 1B is activated when an event may be appraised as being unexpected, in which case the event evokes surprise (Link 1B). Confusion is triggered if the novelty resonates with an impasse (Link 1C).

In addition to the Confusion → Engagement/Flow link, which occurs when an impasse has been resolved, if the resolution triggers an achievement of an important goal or there is positive

feedback on an action, then there is a positive emotion such as contentment or delight (Link 2B). The positive emotion eventually dissipates to the equilibrium state (Link 2C) where the superordinate goal of learning is resumed.

2.1. Scope of the model

It is important to define the scope of the proposed model. First, the model applies to learning at deeper levels of comprehension (deep or complex learning). Complex learning requires learners to generate inferences, answer causal questions, diagnose and solve problems, make conceptual comparisons, generate coherent explanations, and demonstrate application and transfer of acquired knowledge (Graesser, Ozuru, & Sullins, 2010). This form of learning can be contrasted with shallow learning activities (memorizing key phrases and facts) and procedural learning. Complex learning is inevitably accompanied by failure, so the learner experiences a host of affective states. In comparison, they are presumably more affectively neutral in shallow learning sessions under similar task conditions.

Second, the model assumes that learners are in a base state of engagement and are actively processing the material for the impasse-driven dynamics to unfold. A radically different set of

affective trajectories are expected when the learner is bored or passively processing the material at shallower levels of comprehension. Such learners might experience persistent boredom or engage in non-productive activities such as zoning out, checking email, surfing the web, or talking to a friend. The model does not apply to these noncritical passive learners.

Third, the model focuses on a subset of the affective states that routinely arise during the comprehension of difficult subject matter in relatively short learning sessions (e.g., learning physics from a computer tutor or preparing for a standardized exam). From the standpoint of the four categories of emotions proposed by Pekrun (2010), this includes a subset of the achievement (e.g., boredom, frustration) and epistemic emotions (e.g., confusion). Social emotions and topic emotions, although critical to academic settings (Pekrun, 2010), are currently beyond the scope of the model.

2.2. Testing the model at multiple levels

The aforementioned model on affect dynamics can be tested at a number of levels of granularity. At a minimum, the hypothesized affective states (boredom, engagement/flow, confusion, frustration, delight, and surprise) should be observed at significant rates during complex learning activities such as studying for a test, reading difficult technical material, solving complex problems, or interacting with a computer or human tutor. This prediction has been supported in a number of studies that have tracked emotions in a number of contexts and with a variety of methodologies (see Calvo & D'Mello, 2011, for several such studies). One finding is that confusion, frustration, boredom, and engagement/flow are the major affective states that students experience across diverse learning contexts, student populations, and methods to track emotions. States such as surprise and delight are consistently observed, but with lower frequencies. Curiosity and anxiety are observed with relatively high frequencies, but only in some contexts (e.g., high level of intrinsic motivation for curiosity and high-stakes testing or being in the presence of a human tutor for anxiety).

At a finer level of granularity, the model predicts that events in the learning session serve as antecedents or elicitors of specific affective states. One example prediction is that impasses and anomalies trigger confusion and cognitive disequilibrium. Another is that events that block goals and the general unavailability of a plan (being stuck) yield frustration. These predictions have also been supported by existing data, a very small subset of which is summarized in this section. Consistent with the predictions of the model, it appears that contradictory statements in tutorial dialogs, hints (e.g., What about X?), and descriptions of system breakdowns all appear to be antecedents of confusion (D'Mello, Craig, Sullins, & Graesser, 2006; Graesser, Lu, et al., 2005). Similarly, negative feedback from a tutor as well as interruptions that block goals have been established as potential antecedents of frustration (D'Mello et al., 2006; Kapoor, Bursleson, & Picard, 2007).

Testing the model at yet another level of granularity leads to the issue that is at the heart of this paper: what are the natural dynamics of the affective states? Specifically, the model hypothesizes the presence of particular transitions between the affective states and the absence of other transitions. Affective dynamics is an important phenomenon that has not received much emphasis, so tests of the predictions pertaining to affective transitions is an important step in validating the model.

An analysis at an even finer level of granularity would involve simultaneously testing both the predicted transitions as well as the events that are hypothesized to trigger these transitions. That is, for any particular predicted transition, one could assess the probability that the transition was triggered by event E_1 versus E_2 , and so on.

Although the coupling of transitions with events would ostensibly yield a more stringent test of the model than focusing on transitions alone, complications arise in some of the more naturalistic learning sessions. By naturalistic we simply mean that the events and affective states naturally evolve over the course of the interaction as opposed to being experimentally manipulated.

Some of the complications associated with naturally evolving affect include: (a) the relatively low frequencies and skewed distributions of some affective states, (b) difficulties in the temporal alignment of states and events (i.e., when no obvious cause can be associated to a transition), (c) the inability to make causal claims when multiple events occur in close temporal proximity, and (d) interpretation difficulties that arise when the valence of multiple events is not complimentary (e.g., providing negative feedback followed by a motivational statement). Unfortunately, these complications are the rules, rather than the exceptions, in naturalistic learning sessions, as in the case of the present research where students' affective states are tracked during interactions with a dialog-based computer tutor.

One possible way to alleviate these complications is to systematically manipulate the learning session in order to test the causal influences of critical events on particular transitions. For example, one can induce an affective state (the predicted source), generate an event (the predicted trigger), and track the student's subsequent state (the predicted sink). Alternate source states and triggers can also be manipulated to serve as controls. This methodology would ostensibly solve a large number of the issues described above. Unfortunately, its significant limitation is that it does not provide any insights on the *naturalistic dynamics* of student affect, and the resulting manipulated events are sometimes awkward for bonafide learning environments. A methodology with manipulated events that run the risk of clashing with ecologically valid complex learning is incompatible with the goals of the present research.

In our view, the most defensible position is to test the model at all four levels of granularity. One can also introduce a fifth level where the model can be tested for its generalizability to different student populations, content areas, and learning environments. Testing the model at these levels will require a combination of correlational studies, randomized controlled experiments, field studies, and many others. This is an ambitious research program which far exceeds the scope of any single study. The goal of the present study is more modest, namely to test some claims about the dynamics of naturally evolving affective experiences. Our focus on affective transitions and not their causes is our attempt to test the model at a single level, knowing full well that additional experiments with more systematic manipulations and a simpler learning environment will be needed. We chose to focus on this level of analysis because failure to identify systematic transitions among affective states precludes testing at higher levels. Specifically, it is pointless to test causes of affect transitions if the transitions themselves do not spontaneously emerge. It should also be noted that previous studies (cited above) have verified some of the events that trigger particular states (e.g., negative feedback for frustration), so we have some idea as to the causes of some of the transitions.

2.3. Hypotheses and present studies

This study empirically tests four hypotheses of the model pertaining to transitions among affective states. Transitions involving delight and surprise were not addressed because these states are not at the crux of the model and do not occur with sufficient frequency in the learning environment we tested.

The first hypothesis, called the *disequilibrium* hypothesis (Hypothesis 1), states that learners in an engaged/flow state will experience cognitive disequilibrium and confusion when they detect

impasse (Link 1A). This hypothesis predicts the occurrence of transitions from engagement/flow to confusion and the *absence* of transitions from engagement/flow into frustration and boredom. By absence, we mean that a transition either, does not occur, is extremely rare, or its rate of occurrence is statistically indistinguishable from what could be expected by a random process (chance).

The second hypothesis is the *productive confusion* hypothesis (Hypothesis 2). According to this hypothesis, cognitive disequilibrium, impasses, and confusion provide learners with an opportunity to think, deliberate, and problem solve (Graesser & Olde, 2003; VanLehn et al., 2003). Confusion is alleviated and learners revert back to cognitive equilibrium if the problem solving process results in a resolution of the impasse. In contrast to productive confusion, the model hypothesizes that there should be a transition that resonates with *hopeless confusion* when learners cannot resolve an impasse (Hypothesis 3). Hence, the model supports links between confusion and engagement/flow (Hypothesis 2) and confusion and frustration (Hypothesis 3), but not confusion and boredom.

The fourth hypothesis, or the *disengagement hypothesis* (Hypothesis 4), states that persistent failure, which is related to frustration, eventually transitions into disengagement and boredom. According to this fourth hypothesis there should be a transition from frustration to boredom (Link 4), but frustration should not transition into engagement/flow.

The model does not make any predictions for transitions that either emanate from the neutral state (Neutral → Affect) or terminate into the neutral state (Affect → Neutral). It also does not make explicit predictions for four affective transitions. These involve a transition from frustration to confusion and three transitions stemming from boredom. Transitions from frustration to confusion might occur if an additional impasse is reached, whereas a learner is more likely to stay frustrated when they are stuck. Transitions that stem from boredom are equally probable. A bored learner might be re-engaged if presented with an engaging stimulus or might become frustrated from having to endure the learning session. We explored these four possibilities even though they were not guided by any explicit predictions from the model. However, we did not consider transitions involving neutral because the present focus is on Affect → Affect transitions.

The four major hypotheses were tested in two studies in which college students interacted with a validated intelligent tutoring system called AutoTutor. AutoTutor helps learners construct explanations by interacting with them in natural language with adaptive dialog moves similar to human tutors, as will be described below (Graesser, Chipman, Haynes, & Olney, 2005). Both studies involved a 30–35 min learning session after which participants provided self-reported affect judgments at approximately 100 points in their tutoring sessions via a retrospective affect judgment protocol (Graesser et al., 2006). The likelihoods of the hypothesized transitions were assessed from these judgments. The purpose of the first study was to systematically test the predictions of the model, whereas the second study assessed the extent to which the findings replicated with a different version of AutoTutor and with some important methodological changes.

3. Brief description of the learning environment (AutoTutor)

AutoTutor is a fully automated tutor for Newtonian physics, computer literacy, and critical thinking (Graesser, Chipman, et al., 2005; Graesser et al., 2004). The impact of AutoTutor in facilitating the learning of deep conceptual knowledge has been validated in over a dozen experiments on college students (Graesser et al., 2004; VanLehn et al., 2007). Tests of AutoTutor have produced learning gains of .4–1.5 sigma (a mean of .8), depending on the learning measure, the comparison condition, the subject

matter, and the version of AutoTutor. It should be pointed out that the amount of training time and the number of questions covered in the current studies were much less than previous studies that systematically assessed learning gains with AutoTutor. The goals of the present study were to analyze the affective states during learning from AutoTutor rather than assessing learning gains. Given the short amount of training time and the small number of questions covered, impressive learning gains were not expected, and systematic analyses on learning gains were not performed.

AutoTutor's dialogs are organized around difficult questions and problems (called main questions) that require reasoning and explanations in the answers. The following is an example of a challenging question in computer literacy: "How can John's computer have a virus but still boot to the point where the operating system starts?" In order to correctly answer a main question, students need to articulate 3–7 sentential expressions, which normally involves between 25 and 100 turns in a conversation. When presented with these questions, students typically respond with answers that are only one word to two sentences in length. In order to guide students in their construction of an improved answer, AutoTutor actively monitors learners' knowledge states and engages them in a turn-based dialog. AutoTutor adaptively manages the tutorial dialog by providing feedback, pumping the learner for more information, giving hints, correcting misconceptions, answering questions, and summarizing answers.

The dialog for a difficult question is implemented over a number of conversational turns. Each turn of AutoTutor in the conversational dialog has three information slots (i.e. constituents). The first slot of most turns is short feedback on the quality of the student's last turn. This feedback is either positive (e.g., "very good", "bravo"), negative (e.g., "not quite", "almost"), or neutral (e.g., "uh huh", "okay"). The second slot advances the coverage of the ideal answer with either prompts for specific words (e.g., "X is a type of what?"), hints (e.g., "What can you say about X?"), assertions with correct information (e.g., "X is required for"), corrections of misconceptions, or answers to students' questions (via information retrieval from a glossary or textbook). The third slot is a cue to the student for the floor to shift from AutoTutor as the speaker to the student. For example, AutoTutor ends each turn with a question or a gesture (rendered by the animated conversational agent) to cue the learner to do the talking. Discourse markers (e.g., "and also", "okay", "well") connect the utterances of these three slots of information within a turn.

AutoTutor can keep the dialog on track because it is always comparing what the student says to anticipated input (i.e., the expectations and misconceptions in the curriculum script). AutoTutor assesses the conceptual quality of the learner's response using state of the art natural language processing algorithms (Jurafsky & Martin, 2008). As the learner expresses information over many turns, the list of expectations is eventually covered and the main question is scored as answered. AutoTutor then presents students with the next main question in the series.

The AutoTutor interface along with a sample dialog between a college student and AutoTutor is presented in Fig. 2.

4. Study 1

4.1. Method

4.1.1. Participants

The participants were 28 undergraduate Psychology students from a mid-south university in the U.S. There were 5 males and 23 females. 37% of the students were Caucasian, 56% African-American, and 7% were classified as "Other". The participants

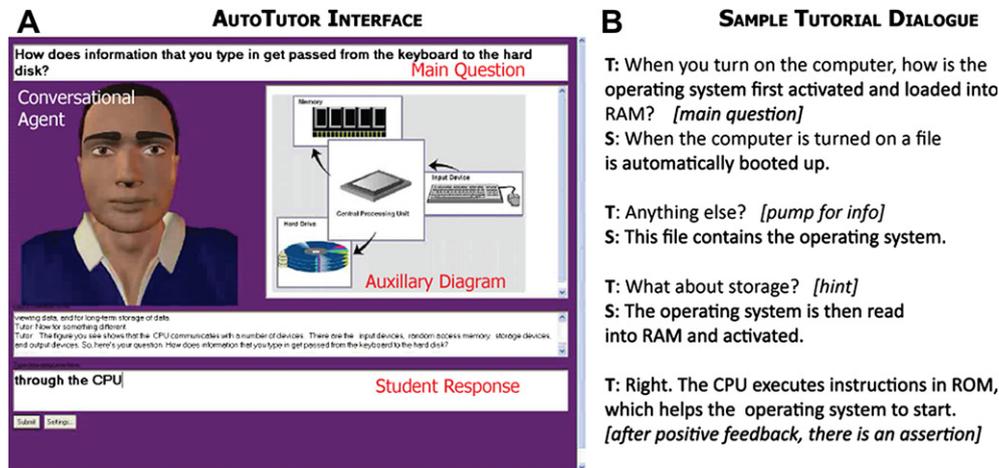


Fig. 2. The AutoTutor interface and sample dialog from an actual tutorial session with a knowledgeable student.

received course credit for their participation. Prior coursework in computer literacy was not required.

4.1.2. Materials and equipment

Two computers were used for presentation and data capture. One of the computers handled the tutorial interaction, whereas the other computer was used to record a video of the participant's face. Three channels of information were recorded and collected while participants interacted with AutoTutor on topics of computer literacy. First, an IBM® Blue Eyes Camera (Morimoto, Koons, Amir, & Flickner, 1998) was used for collecting data on facial movements at a fine-grained resolution. Second, body posture was tracked by the Body Pressure Measurement System (BPMS, version 5.23) developed by Tekscan® Inc (Tekscan, 1997). Third, Camtasia Studio™, a screen capturing software package, recorded a video of the participant's entire tutoring session with AutoTutor. The captured video also included an audio stream of the synthesized speech generated by the AutoTutor animated conversational agent.

4.1.3. Procedure

Participants interacted with AutoTutor version 2.0 for 32 min on one of three randomly assigned topics in computer literacy: hardware ($N = 9$), Internet ($N = 10$), or operating systems ($N = 9$). The three different topics were used to maximize content coverage rather than to systematically test for topic-related differences in affective profiles. The sessions were 32 min long because we were interested in obtaining 30-min tutorial sessions in addition to a 2-min introduction provided by AutoTutor. Sessions were terminated after 32-min to control for time on task across participants.

During the interaction the BPMS and the Blue Eyes camera recorded a video of the participants' posture and face respectively. The computer's display was recorded as well using the Camtasia screen capture software. After interacting with AutoTutor, the video streams from the AutoTutor screen (see Fig. 2) and the participant's face were synchronized and displayed to the participants. The screen capture included the tutor's synthesized speech, printed text, students' responses, dialog history, and images, thereby providing the context of the tutorial interaction.

Participants completed some questionnaires (not relevant to present study) for 10–15 min after the tutorial session. Next, they provided judgments on their affective states by viewing the face and screen videos that were recorded during the tutorial session. The videos were synchronized and at the end of each 20-s interval, the two video streams were paused (freeze framed) and the participant was asked to provide judgments of the affective states they experienced at that instant (fixed judgments). If they were

unsure of their state, they could rewind the window to the previous fixed point and replay the 20-s segment. Participants also designated any affective judgment they had experienced during the 20 s in between the previous pause and the current pause (spontaneous judgments) by manually pausing the videos and providing a judgment. Both fixed and spontaneous judgments occurred simultaneously over a single pass of the face and screen videos.

Participants were provided with a checklist of seven states for them to mark along with definitions of the states. The definitions were present on a piece of paper that participants retained throughout the affect judgment procedure. The experimenter answered any questions pertaining to these definitions prior to commencing the procedure.

Boredom was defined as “being weary or restless through lack of interest”. Confusion was defined as “a noticeable lack of understanding”, whereas engagement/flow was a “state of interest that results from involvement in an activity”. Frustration was defined as “dissatisfaction or annoyance”. Delight was “a high degree of satisfaction”. Surprise was “wonder or amazement, especially from the unexpected”. Neutral was defined as “no apparent emotion or feeling”. The affect judgments were not based on these definitions alone but on the combination of videos of their faces, contextual cues via the screen capture, the definitions of the emotions, and their recent memories of the interaction.

It is important to mention three important points pertaining to the present affect judgment methodology. This procedure was adopted because it affords monitoring participants' affective states at multiple points, with minimal task interference, and without participants knowing that these states were being monitored while completing the learning task. Second, this retrospective affect judgment method has been previously validated (Rosenberg & Ekman, 1994), and analyses comparing these offline affect judgments with online measures encompassing self-reports and observations by judges have produced similar distributions of emotions (Craig, D'Mello, Witherspoon, & Graesser, 2008; Craig, Graesser, Sullins, & Gholson, 2004). Third, the offline affect annotations obtained via this retrospective protocol correlate with online recordings of facial activity and gross body movements in expected directions (D'Mello & Graesser, 2010). Although no method is without its limitations, the present method appears to be a viable approach to track emotions at a relatively fine-grained temporal resolution.

4.1.4. Data treatment

4.1.4.1. *Preparing time series of affective responses.* The retrospective affect judgment procedure yielded 2547 fixed judgments and

430 spontaneous judgments, collectively yielding 2977 self-reported affect judgments for the 28 participants' AutoTutor sessions. On average there were 91 (SD = 10) fixed judgments and 17 (SD = 14) spontaneous judgments per participant. Since the goal of the present analyses is to analyze affect transitions across each participant's entire AutoTutor session, both fixed and spontaneous judgments were analyzed together rather than in isolation. On average there were 106 total judgments from each participant (SD = 9).

There were 483 instances of boredom, 593 engagement/flow, 533 confusion, 335 frustration, 849 neutral, 94 delight, and 80 surprise. We analyzed the proportional occurrence of the affective states experienced by the 28 learners. A repeated measures ANOVA indicated that there was a statistically significant difference in the distribution of states, $F(6, 162) = 10.81$, $Mse = .023$, $p < .001$, partial eta-square = .286. Bonferroni post-hoc tests revealed that the proportional occurrence of boredom ($M = .160$, $SD = .140$), confusion ($M = .180$, $SD = .127$), engagement/flow ($M = .199$, $SD = .161$), frustration ($M = .114$, $SD = .107$), and neutral ($M = .288$, $SD = .248$) were on par and significantly greater than delight ($M = .032$, $SD = .038$) and surprise ($M = .027$, $SD = .028$), which were equivalent to each other.

A time series that preserved the temporal ordering of the affective states was constructed for each participant. On average, there were 106 elements per time series (SD = 9). Since the goal of this paper is to investigate transitions between different states, and not persistence in the same state, the data was recoded to eliminate repetitions between states. For example, the sequence $X \rightarrow Y \rightarrow Y \rightarrow Z$ was converted to $X \rightarrow Y \rightarrow Z$. This process reduced the length of the time series to 1796 states with a mean of 64 states per participant (SD = 19). On average, there was a state transition every 32.4 s (SD = 11.2). The recoding process did not alter the distribution of the affective states.

4.1.4.2. Computing transition likelihoods. We developed a likelihood metric to compute the likelihood of a transition between any two affective states in order to test the major hypotheses of the model. The metric can be represented as $L(M_t \rightarrow M_{t+1})$, where M_t is the state at time t (the current state), M_{t+1} is the next state at $t+1$. We motivate the likelihood metric by contrasting it with the conditional probability of a transition to M_{t+1} given that the immediate state is M_t (see Eq. (1)). Although conditional probabilities are suitable for exploring the strength of association between M_t and M_{t+1} , they do so relative to M_t . However, if a particular M_{t+1} , say boredom (B), is very frequent, the conditional probability ($B|M_t$) will be high because $B|M_t$ is proportional to the frequency of boredom (also known as base rate).

$$\Pr(M_{t+1}|M_t) = \frac{\Pr(M_{t+1} \cap M_t)}{\Pr(M_t)} \quad (1)$$

The likelihood metric addresses the influence of base rate by penalizing associations that are not greater than an expected amount of association. Formally, the likelihood metric is defined as:

$$L(M_t \rightarrow M_{t+1}) = \frac{\Pr(M_{t+1}|M_t) - \Pr(M_{t+1})}{1 - \Pr(M_{t+1})} \quad (2)$$

The reader may note significant similarity to Cohen's κ kappa agreement between raters (Cohen, 1960) and indeed the likelihood metric can be justified in a similar fashion. The definition of Cohen's κ is listed in Eq. (3).

$$\kappa = \frac{\Pr(A) - \Pr(E)}{1 - \Pr(E)} \quad (3)$$

In Eq. (3), $\Pr(A)$ is the agreement between two raters, and $\Pr(E)$ is the expected agreement. So κ removes the agreement expected by chance and then normalizes by the total possible agreement (1: perfect agreement) minus expected agreement. The likelihood metric proceeds in the same way, however, rather than agreement, we use conditional probability as a measure of association (see Eq. (2)). The expected degree of association is $\Pr(M_{t+1})$, because if M_{t+1} and M_t are independent, then $\Pr(M_{t+1}|M_t) = \Pr(M_{t+1})$. Therefore, the numerator of Eq. (2) equals the degree of association observed minus the degree of association expected under independence. If the observed degree of association is higher than that expected under independence, then the numerator will be positive (a positive association). If the observed association is equal to that expected under independence, then the numerator will be zero (no association). Finally, if the observed association is less than that expected under independence, the numerator will be negative (a negative association). Thus the sign and the magnitude of L is intuitively understandable as the direction and size of the association between M_{t+1} and M_t , accounting for the base rate of M_{t+1} .

4.2. Results and discussion

Eq. (2) was used to compute the likelihoods of all possible state transitions excluding repetitions between states. This yielded 7×6 or 42 possible transitions. Raw frequencies of these states transitions after summing over all participants are presented in the Appendix. Instead of focusing on all 42 transitions, significance testing was performed on the eight theoretically specified state transitions. These included four transitions that were hypothesized to occur above chance (Engagement/Flow \rightarrow Confusion, Confusion \rightarrow Engagement/Flow, Confusion \rightarrow Frustration, and Frustration \rightarrow Boredom) and four transitions whose occurrence was expected at chance levels (Engagement/Flow \rightarrow Boredom, Engagement/Flow \rightarrow Frustration, Confusion \rightarrow Boredom, and Frustration \rightarrow Engagement/Flow). We performed one-sample, t -tests to test whether likelihoods were significantly greater than or equivalent to zero (no relationship between immediate and next state).

Descriptive statistics on the likelihood of the transitions and the results of the t -tests are presented in Table 1. It should be noted that there is some variation in the degrees of freedom for each test reported in Table 1, because some of the learners did not report experiencing all the states. In particular, while all 28 learners experienced at least one episode of confusion, this was not the case for the other states. Four learners never experienced engagement/flow, three did not report any episodes of frustration, and two learners never reported being bored during the task. Hence, it was not possible to compute transitions emerging from these states for those learners.

4.2.1. Testing occurrence of predicted transitions

Hypothesis 1, or the disequilibrium hypothesis, predicts the existence of a transition from engagement/flow to confusion but not to boredom or frustration (i.e., this transition should not occur or should occur at chance rates). The results confirmed this hypothesis as the Engagement/Flow \rightarrow Confusion was significant with a medium to large effect size (Cohen, 1992), whereas the Engagement/Flow \rightarrow Boredom and Engagement/Flow \rightarrow Frustration transitions occurred at chance levels.

Hypotheses 2 and 3, the productive confusion and hopeless confusion hypotheses, respectively, were also confirmed by the fact that the Confusion \rightarrow Engagement/Flow and Confusion \rightarrow Frustration transitions significantly occurred. Hence, engagement/flow and frustration are the consequents of productive and hopeless confusion, respectively. The results also confirmed that the

Table 1
Transition likelihoods for Study 1.

Transition	Descriptives			One-sample <i>t</i> -test			
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
Excitatory							
Flow → Confusion**	24	.175	.263	3.27	23	.003	.667
Confusion → Flow**	28	.147	.224	3.46	27	.002	.654
Confusion → Frustration**	28	.060	.133	2.37	27	.025	.447
Frustration → Boredom	25	.098	.259	1.90	24	.070	.379
Inhibitory/baseline							
Flow → Boredom	24	.022	.212	.515	23	.612	.105
Flow → Frustration	24	.017	.102	.807	23	.428	.165
Confusion → Boredom	28	-.033	.147	-1.18	27	.249	-.223
Frustration → Flow	25	.022	.251	.435	24	.667	.087
No prediction							
Frustration → Confusion	25	.056	.293	.961	24	.346	.192
Boredom → Confusion	26	-.012	.175	-.354	25	.726	-.069
Boredom → Flow	26	.062	.253	1.25	25	.224	.245
Boredom → Frustration**	26	.088	.181	2.47	25	.021	.485

Notes. ** $p < .05$. Flow refers to Engagement/Flow.

Confusion → Boredom transition occurred at chance levels, which is in line with the predictions of the model.

The marginally significant transition between frustration and boredom partially supports the disengagement hypothesis (Hypothesis 4), which states that persistent failure and the associated frustration will eventually transition into boredom. Furthermore, as expected, the frustration to engagement/flow transition was not significant.

4.2.2. Additional patterns

Although the model does not explicitly predict a Frustration → Confusion transition and any transitions from boredom, we tested these transitions as well. As could be expected, the Boredom → Confusion transition occurred at chance levels. Hence, boredom is not a precursor to cognitive disequilibrium and confusion and some level of engaged concentration is needed for an impasse to be detected. But contrary to our expectations, the Boredom → Frustration transition was significant, suggesting that bored learners might experience frustration from having to endure the learning session despite their ennui; this pattern is consistent with the forced-effort model of boredom (Larson & Richards, 1991).

The results also revealed that the Frustration → Confusion and Boredom → Engagement/Flow transitions, although not significant, had likelihood scores that were consistent with small effects. Study 2 tested whether these unpredicted patterns were noteworthy or spurious, and whether the predicted patterns were also observed.

5. Study 2

The results from Study 1 supported the major hypotheses of the model. We attempted to replicate these findings in Study 2, which was similar to Study 1, but with two important differences. While participants in Study 1 interacted with the traditional typed-input version of AutoTutor, Study 2 participants spoke their responses to a new spoken-input AutoTutor. It should be noted that previous experiments comparing the typed- and spoken-input versions of AutoTutor showed gains in deep learning for both conditions and an equivalence in learning gains across conditions (D'Mello, King, & Graesser, 2010).

The second difference between the two studies pertains to the retrospective affect judgment protocol. While participants in Study 1 provided affect judgments every 20 s and in-between each 20-s block, participants in Study 2 provided judgments at three pre-

selected points plus some random points in the tutorial session (described below). This alternate aperiodic affect-polling schedule was adopted in order to avoid concerns that the periodic nature of the affect sampling in Study 1 (every 20 s) inadvertently influenced the results. For example, these fixed 20-s points might correspond to instances where there is not much activity and other sampling methods might yield a different distribution of affective states and alternate transitions between the states.

5.1. Method

5.1.1. Participants

The participants were 30 undergraduate students (13 male and 17 female) from a mid-south university in the U.S. who participated for course credit. Prior coursework in computer literacy was not required.

5.1.2. Materials and equipment

5.1.2.1. *Sensory data recording.* The setup in Study 2 was the same as Study 1 with the exception that participants used a set of headphones that included a microphone. The recorded video also captured participants' spoken responses to the tutor, so they could hear their speech, the tutor's speech, and view videos of their faces while making their affect judgments.

5.1.3. Procedure

Similar to Study 1, participants interacted with the spoken-input AutoTutor for 35 min on one of three randomly assigned topics in computer literacy (same topics as Study 1). Affect ratings were obtained at three specific types of judgment points: (1) a few seconds after AutoTutor completed a dialog move (tutor move), (2) immediately before the learner started expressing his or her spoken response to the tutor (student move), and (3) other randomly selected points in the dialog (random point). We focused on obtaining judgments corresponding to student and tutor moves because these represent key instances in the tutorial session. Note that there is a gap between the completion of the tutor move and the student's response because the student needs to deliberate in order to respond to the tutor's question. Random points were also selected to ensure an even distribution across the session and to capture those moments of student thought and deliberation (i.e., when neither the student nor the tutor are speaking or preparing to speak).

The observation sampling points were selected by first obtaining all available points corresponding to tutor and student moves and randomly selecting 33 judgment points each, thereby yielding 66 affect judgment points corresponding to student and tutor moves. An additional 33 points were randomly selected from the entire tutoring session; these points did not correspond to any particular tutor or student action. Participants provided affective ratings at these 99 judgment points. These constituted the fixed judgment points. Similar to Study 1, the participants could stop the video at any time and make a spontaneous judgment.

5.1.4. Data treatment

5.1.4.1. *Preparing time series of affective responses.* The retrospective affect judgment procedure yielded 2716 fixed judgments and 383 spontaneous judgments, collectively yielding 3099 self-reported affect judgments for the 30 participants' AutoTutor sessions. On average there were 91 ($SD = 2$) fixed judgments and 17 ($SD = 13$) spontaneous judgments per participant. The average number of total judgments per participant was 103 ($SD = 14$).

There were 574 instances of boredom, 708 engagement/flow, 596 confusion, 409 frustration, 583 neutral, 160 delight, and 69 surprise. A repeated measures ANOVA indicated that there was a significant difference in the proportional distribution of affective

states, $F(6, 174) = 10.94$, $Mse = .017$, $p < .001$, partial eta-square = .274. Bonferroni post-hoc tests revealed that boredom ($M = .190$, $SD = .158$), confusion ($M = .188$, $SD = .097$), engagement/flow ($M = .232$, $SD = .199$), frustration ($M = .129$, $SD = .086$), and neutral ($M = .189$, $SD = .132$) were the major states experienced; delight ($M = .051$, $SD = .068$) and surprise ($M = .021$, $SD = .023$) were comparatively rare. The distribution of states was similar to Study 1.

Time series constructed for each participant's data had a mean length of 103 states ($SD = 14$). Time series that were recoded to eliminate repetitions between states had a total of 1957 states with a mean of 64 states per participant ($SD = 15.2$). On average, there was a state transition every 32.8 s ($SD = 9.58$). This is equivalent to the 32.4 mean transition length obtained in Study 1. Recoding the time series did not alter the distribution of states. These results indicate that despite the methodological changes in Study 2, the distribution of affective states, the length of the time series (with and without repetitions), and the rate of state transitions were similar across studies.

5.2. Results and discussion

We repeated the likelihood analyses to determine whether the patterns detected in Study 1 were replicated in Study 2. Raw frequencies of the 42 transitions summed over all 30 participants are presented in the Appendix. The subsequent discussion focuses on testing likelihoods associated with the predicted transitions.

5.2.1. Testing occurrence of predicted transitions

The results indicated that, with one exception, the major predictions of the model were supported in Study 2 (see Table 2). In particular, the Engagement/Flow → Confusion transition significantly occurred, whereas the Engagement/Flow → Frustration and Engagement/Flow → Boredom transitions occurred at chance levels, thereby supporting the disequilibrium hypothesis (Hypothesis 1). The Confusion → Engagement/Flow and Confusion → Frustration transitions were also significant, while the Confusion → Boredom transition occurred at chance levels, thereby confirming the productive confusion and hopeless confusion hypotheses (Hypotheses 2 and 3, respectively).

The disengagement hypothesis (Hypothesis 4) was partially supported in Study 2. As predicted, the Frustration → Boredom transition was significant. However, although the model predicts that a Frustration → Engagement/Flow transition should not occur, this transition was significant with a small to medium effect size.

Table 2
Transition likelihoods for Study 2.

Transition	Descriptives			One-sample <i>t</i> -test			
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>Df</i>	<i>p</i>	<i>d</i>
Excitatory							
Flow → Confusion**	28	.179	.277	3.41	27	.002	.645
Confusion → Flow**	30	.103	.226	2.49	29	.019	.454
Confusion → Frustration**	30	.076	.152	2.74	29	.010	.500
Frustration → Boredom**	29	.078	.205	2.06	28	.049	.382
Inhibitory/baseline							
Flow → Boredom	28	-.030	.158	-1.00	27	.324	-.190
Flow → Frustration	28	-.006	.158	-.194	27	.848	-.037
Confusion → Boredom	30	.030	.221	.734	29	.469	.134
Frustration → Flow**	29	.111	.263	2.27	28	.031	.422
No prediction							
Frustration → Confusion	29	.045	.218	1.10	28	.280	.205
Boredom → Confusion	30	.095	.296	1.75	29	.090	.320
Boredom → Flow	30	.034	.283	.654	29	.518	.119
Boredom → Frustration**	30	.084	.222	2.06	29	.049	.376

Notes. ** $p < .05$. Flow refers to Engagement/Flow.

This transition was not significant in Study 1, hence, its presence in Study 2 could be attributed to some of the methodological changes. We suspect that transitions from Frustration → Boredom are a result of being unable to resolve an impasse after considerable effort has been invested (consistent with the model). On the other hand, Frustration → Engagement/Flow transitions might occur when frustration is caused by temporary, yet irritating, speech recognition problems; the learner reverts back into the engaged state once the system is able to comprehend his responses. However, these speculations need to be confirmed in future analyses.

5.2.2. Additional patterns

Out of the four transitions for which the model does not make any explicit predictions, the Boredom → Frustration transition was statistically significant. This transition was also significant in Study 1, indicating that the link between boredom and frustration is quite reliable (average $d = .43$). Hence, persistent boredom begets frustration.

One surprising finding was that the Boredom → Confusion transition, although not significant, occurred with some regularity in Study 2 but not Study 1. In contrast, the Boredom → Engagement/Flow transition had a non-significant but non-zero likelihood in Study 1, but was substantially less prevalent in Study 2. This pattern, which was not consistently observed in both studies, might be attributed to methodological differences across studies or some other spurious factor. Hence, we do not put too much stock in these transitions.

It is interesting that the Frustration → Confusion transitions were not significant in either study, but did show non-zero likelihoods ($M = .045$ for both studies). This pattern is intuitively plausible because a frustrated learner might transition into confusion if an additional impasse is detected. It might be the case that the small sample sizes in these studies did not yield the requisite power to detect this small effect, so replication with a larger sample is warranted.

6. General discussion

We have presented a model of affective dynamics that is theoretically grounded in perspectives highlighting the importance of goal appraisal, cognitive disequilibrium, and impasse resolution during learning and problem solving. The model provides an analysis of the dynamics of affective states that learners experience when they encounter novelty (surprise), experience impasses (confusion), diagnose what went wrong (confusion), make mistakes, get stuck (frustration), discover a relevant insight (with delight and possibly eureka), get re-engaged (possibly in an engaged and flow-like state), or simply give up and disengage (boredom). The major predictions of the model were verified in two studies that assessed the incidence and transition likelihoods of naturally occurring affective states.

With the exception of delight and surprise, which were comparatively rare, an analysis of the distribution of affective states across both studies indicated that complex learning was rife with experiences of engagement/flow, confusion, boredom, and frustration. Specifically, although the neutral state was also reported with non-zero frequencies, the six affective states comprised the majority of the observations (71.2% and 81.1% in Study 1 and 2, respectively). Complex learning is indeed an emotionally charged experience.

In addition to the incidence of the affective states, our results indicate that learners do not randomly shift between emotions. In stark contrast, there appears to be some systematicity in how learners transition between different affective states, and we

proposed and evaluated a model aimed at capturing some of these affective dynamics. Although the major predictions of the model were confirmed in two studies, the analyses also suggested that the model needs to be expanded to include a Boredom → Frustration link and perhaps even a Frustration → Confusion link. The revised model is presented in Fig. 3.

The revised model supports the assertion that students in the state of engagement/flow are continuously being challenged within their zones of optimal learning (Brown, Ellery, & Campione, 1998; Vygotsky, 1978) and are experiencing two-step episodes alternating between confusion and insight (Hypotheses 1 and 2). In contrast to these presumably *beneficial* engagement/flow–confusion–engagement/flow cycles, there are the *harmful* oscillations between boredom and frustration (Hypothesis 4 augmented to include a Boredom → Frustration transition). There are also *recurrent* oscillations between confusion and frustration when a confused learner gets stuck and experiences frustration, which can trigger confusion if an additional impasse is detected (Hypothesis 3 augmented to include a Frustration → Confusion transition).

The subsequent discussion focuses on some of the broader issues stemming from this research. These include a discussion on the central role of confusion and cognitive disequilibrium during deep learning, implications of our results for pedagogical strategies, and a discussion of limitations and possible resolutions.

6.1. Central role of confusion and cognitive disequilibrium in deep learning

A major assumption of the model is that impasses and the resultant cognitive disequilibrium and confusion are precursors to deep learning (Graesser, Lu, et al., 2005; Graesser & Olde, 2003; VanLehn et al., 2003). Although this research did not explicitly address learning gains, there is some evidence to support this assumption. In particular, previous research has correlated confusion (and uncertainty) with positive learning outcomes in computer and human tutoring contexts (Craig et al., 2004; D'Mello & Graesser, *in press*; Forbes-Riley & Litman, 2010; Graesser, Chipman, King, McDaniel, & D'Mello, 2007; VanLehn et al., 2003). There is also some evidence to suggest that deep learning cannot be satisfactorily achieved without impasses and confusion. For example, VanLehn et al. (2003) reported that the learning of

physics concepts was rare when students did not reach an impasse, irrespective of quality of the explanations provided by tutors.

It is important to note that learning is presumably not directly caused by confusion, but rather by the cognitive activities (e.g., impasse resolution, causal reasoning) that accompany its experience. We recently tested this hypothesis by conducting two experiments in which confusion was induced while participants performed a device comprehension task (understanding how devices such as toasters and doorbells work from technical illustrated texts) (D'Mello & Graesser, *submitted for publication*). The manipulation consisted of presenting participants with descriptions of device malfunctions (e.g., “When a person rang the bell there was a short *ding* and then no sound was heard”) and asking them to diagnose the problem. A second-by-second analysis of the dynamics of confusion yielded two characteristic trajectories that successfully distinguished those individuals who partially resolved their confusion over the course of diagnosing the breakdowns from those who remained confused. Importantly, individuals who partially resolved their confusion demonstrated significantly better device comprehension from individuals who remained confused.

6.2. Pedagogical implications of results

A major assertion that emerges from the model is that learning environments need to substantially challenge students in order to elicit critical thought and deep inquiry. It is a somewhat counter-intuitive claim that pedagogical tactics that challenge, perplex, and productively confuse learners are stimulating alternatives to the typical shallow information delivery systems that are comfortable to passive noncritical learners. Nevertheless, confusion is an affective state that correlates with learning gains because it accompanies impasses and the resulting deep thinking. Students are in a low state of arousal and in cognitive equilibrium when they are in comfortable learning environments involving passive reading and accumulating shallow facts without challenges. However, this does not yield deep learning. In contrast, deep learning is much higher in conditions that present challenges to inspire deep inquiry.

A tutor (human or artificial) that assumes that Engagement/Flow ↔ Confusion oscillations are beneficial to learning can induce confusion in order to jolt students out of their base state of passively attending to the tutoring session. This will place learners in a state of cognitive disequilibrium in which they will have to

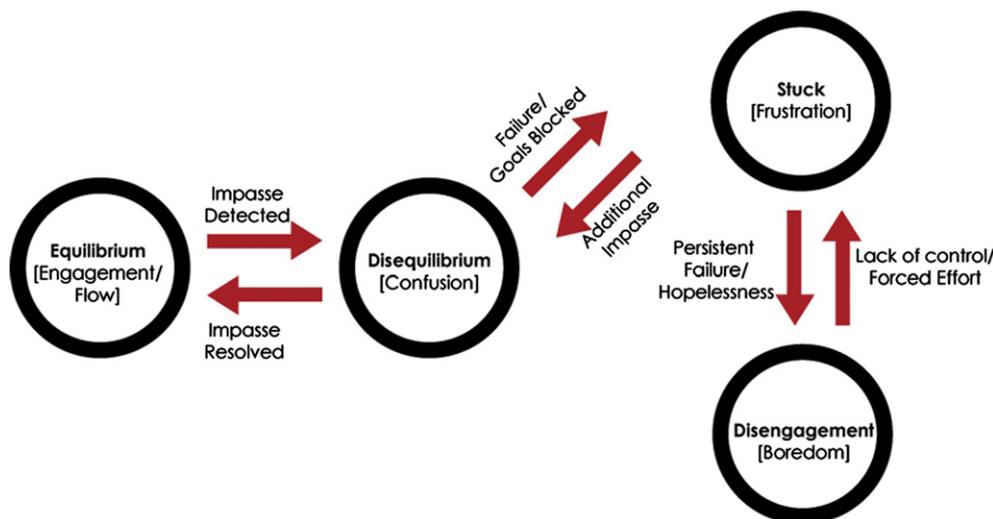


Fig. 3. Observed model of affect dynamics.

stop, think, reason, and be active problem solvers. One indirect way to induce confusion involves introducing new complex topics or concepts, which will hopefully cause the student to reengage with the material at a deeper level. Perhaps a more direct method to induce confusion would be to intentionally cause discrepancies that challenge students' existing mental models and prior knowledge. For example, we have successfully induced confusion by presenting participants with descriptions of device malfunctions (D'Mello & Graesser, submitted for publication) and by introducing contradictions in tutorial dialogs (Lehman et al., 2011).

If these manipulations succeed and the student is confused, then the tutor needs to help the learner manage his confusion. Perhaps the most promising strategy to scaffold confusion is one recommended by VanLehn in his research on impasses during learning (VanLehn et al., 2003). This strategy takes effect when confusion is detected and it entails: (a) prompting the student to reason and arrive at a solution, (b) prompting the student to explain their solution, and (c) providing the solution with an explanation only if the student fails to arrive at an answer. This strategy places the onus on the shoulders of the students and capitalizes on the benefits of impasses, confusion, and self-explanations.

It is important, however, that the VanLehn strategy should be tailored to the individual learner because these interventions might backfire and the learner might get frustrated and eventually disengage. Successful learners might be allowed to work out their own confusion in a discovery learning environment (Bruner, 1961; Vavik, 1993) that requires self-regulated cognitive activities. An alternative method would systematically scaffold the student out of the confused state. This method would presumably work better for learners with lower domain knowledge and lower ability to self-regulate their learning activities.

Pedagogical interventions would also be required to help learners regulate negative states such as frustration and boredom. Frustration regulation is particularly important because there is evidence for oscillations between frustration and boredom. If the affective state of frustration is detected, then the tutor could respond by changing its dialog strategies to include more direct feedback, assertions, and corrections of detected misconceptions. The tutor might also convey a degree of empathy to alleviate frustration (Lepper & Woolverton, 2002; Woolf et al., 2009). If the learner is bored, then the tutor should engage the learner in a task that increases interest and cognitive arousal, provides options of choice to increase perceived control, or provides a challenge (Csikszentmihalyi, 1975; Lepper & Woolverton, 2002; Pekrun, Goetz, Daniels, Stupnisky, & Raymond, 2010).

6.3. Limitations and future work

It is important to acknowledge five limitations of this research. One limitation is that the current model and reported data do not explicitly address individual differences in the manner in which learners experience and regulate their affective states. One important individual difference pertains to the causes that students attribute their emotions to (Jarvenoja & Jarvela, 2005). Attribution theory predicts that an impasse could be attributed to either an external cause (e.g., a problem in the learning material) or an internal cause (e.g., lack of student knowledge and effort) (Heider, 1958; Weiner, 1986). Such attributions should have an impact on the manner in which learners attempt to regulate confusion and other emotions. Causal attributions of failure also play a role on how frustration is regulated. Some learners blame themselves and go down a negative spiral. Others regard the failure as temporary and want to conquer the challenge.

Individual differences in academic risk taking, self-efficacy, and motivation (Bandura, 1997; Clifford, 1988; Meyer & Turner, 2006;

Zimmerman, 2000) are expected to play a role in how learners regulate their confusion. Mastery orientated learners, academic risk takers, and individuals with high self-efficacy will presumably regulate their confusion through effortful problem solving and cognitive deliberation. In contrast, cautious learners and learners with low self-efficacy might detect the impasses but be hesitant to step up and confront their confusion because of self-doubt and the threat of failure. These learners are also more likely to experience feelings of hopelessness and disengage when they get confused. In terms of motivation, performance-oriented students who lack intrinsic motivation might show a tendency to ignore experiences of confusion and simply forge ahead without directly confronting the underlying source of their perplexity (Pekrun, Elliot, & Maier, 2006).

The second limitation of this research is that the analyses did not consider the causes of some of the critical state transitions (i.e. links in Fig. 1) by analyzing the tutor's log files. There were two complications with the present data that precluded a more detailed evaluation of the model. One problem is that there are substantially more affective state reports (approximately 3000) than dialog events. Hence, a majority of the transitions do not have an intervening dialog move which can be examined to identify a causal link between affective transitions. A different challenge occurs for the few transitions that have a dialog event that could serve as a potential trigger for the transition. The problem arises when two or more events occur in close succession, which is more the rule than the exception with AutoTutor. This is a problem because it is difficult or impossible to determine which event influenced the state transition. For example, consider a situation where a student in some affective state submits a response, the tutor assesses this response, deems it to be of high quality, and provides positive feedback, followed by a hint to advance the dialog. The student then reports being in a different affective state. This entire chain of events (original affective state → student answer → positive feedback → hint → new affective state) transpires over a few seconds, so it is difficult to attribute the transition to any one of the three intervening events. It is possible to consider combinations of events as potential causes of state transitions. However, the large number of possible combinations is incompatible with the available data in small datasets that capture naturalistic tutorial interactions.

The third limitation is that we did not systematically analyze transitions involving the neutral state. To be clear, the neutral state was included in the time series that we analyzed (see Appendix), but we did not perform any additional statistical analyses on Affect → Neutral and Neutral → Affect transitions. Transitions involving neutral, particularly Affect → Neutral transitions, were not considered because the presumably numerous causes for these transitions (e.g., the learner suppresses the emotion, the emotion is reappraised, the emotion naturally dissipates, the emotion dissipates because the student gets distracted by some event) are beyond the scope of the present model. Refining the model to include some of these causes is an important item for future work.

The fourth limitation pertains to potential concerns with the present methodology. Although this method has several advantages (see Method of Study 1), it is limited by the learner's ability and sensitivity to his or her emotions, as well as by the learner's desire to be honest. Social biases may also affect how learners report their affective states. For example, a learner might be unwilling to report confusion or frustration, due to the obvious negative connotations associated with these states. Another limitation of this method is that it is restricted to consciously accessible affective states, and overlooks unconscious feelings and background moods.

The fifth limitation pertains to the generalizability of our findings. The proposed model was tested on the sample

population of students (undergraduate psychology majors from a southern University in the US), with similar methods, and with one particular tutoring system, namely AutoTutor. This raises the question of whether the model would be supported when different populations of students engage in alternate learning contexts (e.g., reading textbook or interacting with other computer-based learning systems), and when alternate methods are used to track their emotions such as online emote-alouds (Craig et al., 2008) or observations by external judges (Rodrigo & Baker, 2011).

In summary, the five limitations of this research are that we have not explicitly addressed individual differences, we have not linked the affect transitions to their causes (dialog moves) and effects (learning), transitions involving neutral were not considered, there were some methodological concern, and we have not assessed the generalizability of the findings. These issues will of course never be addressed in a single study and call for a larger program of research, which we are currently pursuing. Ongoing research plans along this front include: (a) developing an enhanced model that specifies how individual difference influences state transitions via testable hypotheses, (b) conducting experiments to systematically induce affect transitions (e.g., repeatedly providing negative feedback or intentionally inducing confusion), (c) expanding the model to include transitions involving the neutral state, (d) empirically testing the enhanced model with longer tutorial sessions so that deep learning gains can be accurately measured, and (e) testing the model on new populations in different learning contexts and with alternate methodologies to track affect.

Acknowledgments

We thank our research colleagues in the Emotive Computing Group and the Tutoring Research Group (TRG) at the University of Memphis (<http://emotion.autotutor.org>). Special thanks to Patrick Chipman, Scotty Craig, Brandon King, Bethany McDaniel, Jeremiah Sullins, Kristy Tapp, and Amy Witherspoon for data collection.

This research was supported by the National Science Foundation (NSF) (ITR 0325428, HCC 0834847), Institute of Education Sciences (IES), and U.S. Department of Education (DoE), through Grant R305A080594. Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF, IES, or DoE.

Appendix

Frequency of transitions between states for Studies 1 and 2. Observed frequencies of state transitions. The significant transitions (after correcting for base rates) and which are discussed in the paper are bolded (see Tables 1 and 2).

Current State	N	Next State						
		Flow	Con	Fru	Bor	Del	Sur	Neu
Study 1								
Flow	373	–	120	61	72	22	21	72
Confusion	354	116	–	70	44	13	14	95
Frustration	245	62	62	–	51	7	9	49
Boredom	288	71	50	49	–	5	3	103
Delight	85	20	18	6	5	–	5	30
Surprise	78	22	18	9	7	6	–	15
Neutral	373	75	83	49	107	31	21	–

Appendix (continued)

Current State	N	Next State						
		Flow	Con	Fru	Bor	Del	Sur	Neu
Study 2								
Flow	353	–	102	56	59	21	14	94
Confusion	415	95	–	93	70	20	16	118
Frustration	296	60	80	–	82	9	8	55
Boredom	327	59	73	78	–	22	6	82
Delight	110	26	20	11	17	–	6	23
Surprise	64	12	19	7	10	8	–	8
Neutral	392	97	111	51	87	30	12	–

Note. Flow refers to Engagement/Flow.

References

- Baker, R., D'Mello, S., Rodrigo, M., & Graesser, A. (2010). Better to be frustrated than bored: the incidence and persistence of affect during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, 68(4), 223–241.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Barrett, L. F. (2009). Variety is the spice of life: a psychological construction approach to understanding variability in emotion. *Cognition & Emotion*, 23(7), 1284–1306.
- Berlyne, D. (1978). Curiosity in learning. *Motivation and Emotion*, 2, 97–175.
- Brown, A., Ellery, S., & Campione, J. (1998). Creating zones of proximal development electronically in thinking practices in mathematics and science learning. In J. Greeno, & S. Goldman (Eds.) (pp. 341–368). Mahawah, NJ: Lawrence Erlbaum.
- Bruner, J. (1961). The act of discovery. *Harvard Educational Review*, 31(1), 21–32.
- Buff, A., Reusser, K., Rakoczy, K., & Pauli, C. (2011). Activating positive affective experiences in the classroom: "Nice to have" or something more? *Learning and Instruction*, 21(3), 452–466.
- Calvo, R., & D'Mello, S. (Eds.). (2011). *New perspectives on affect and learning technologies*. New York: Springer.
- Carroll, J., & Kay, D. (1988). Prompting, feedback and error correction in the design of a scenario machine. *International Journal of Man-Machine Studies*, 28(1), 11–27.
- Chinn, C., & Brewer, W. (1993). The role of anomalous data in knowledge acquisition – a theoretical framework and implications for science instruction. *Review of Educational Research*, 63(1), 1–49.
- Clifford, M. (1988). Failure tolerance and academic risk-taking in ten- to twelve-year-old students. *British Journal of Educational Psychology*, 58, 15–27.
- Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, 11(9), 393–399.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Collins, A., Warnock, E. H., Aiello, N., & Miller, M. L. (1975). Reasoning from incomplete knowledge. In D. Bobrow, & A. Collins (Eds.), *Representation and understanding* (pp. 383–415). New York: Academic Press.
- Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-adapted Interaction*, 19(3), 267–303.
- Craig, S., D'Mello, S., Witherspoon, A., & Graesser, A. (2008). Emote aloud during learning with AutoTutor: applying the facial action coding system to cognitive-affective states during learning. *Cognition & Emotion*, 22(5), 777–788.
- Craig, S., Graesser, A., Sullins, J., & Gholson, J. (2004). Affect and learning: an exploratory look into the role of affect in learning. *Journal of Educational Media*, 29, 241–250.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco, CA: Jossey-Bass.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
- Deci, E., & Ryan, R. (2002). The paradox of achievement: the harder you push, the worse it gets. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 61–87). Orlando, FL: Academic Press.
- D'Mello, S., Craig, S., Sullins, J., & Graesser, A. (2006). Predicting affective states expressed through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education*, 16(1), 3–28.
- D'Mello, S., & Graesser, A. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-adapted Interaction*, 20(2), 147–187.
- D'Mello, S., Graesser, A. The half-life of cognitive-affective states during complex learning. *Cognition & Emotion*, in press.
- D'Mello, S., King, B., & Graesser, A. (2010). Towards spoken human-computer tutorial dialogues. *Human-Computer Interaction*, 25(4), 289–323.
- D'Mello, S. K., Graesser, A. C. Inducing and tracking confusion and cognitive disequilibrium with breakdown scenarios, submitted for publication.
- Dweck, C. (2002). Messages that motivate: how praise molds students' beliefs, motivation, and performance (in surprising ways). In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 61–87). Orlando, FL: Academic Press.

- Ekman, P. (1984). Expression and the nature of emotion. In K. Scherer, & P. Ekman (Eds.), *Approaches to emotion* (pp. 319–344). Hillsdale, NJ: Erlbaum.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3–4), 169–200.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Fielder, K. (2001). Affective states trigger processes of assimilation and accommodation. In K. Martin, & G. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 85–98). Mahwah, NJ: Erlbaum.
- Forbes-Riley, K., & Litman, D. (2010). Designing and evaluating a wizarded uncertainty-adaptive spoken dialogue tutoring system. *Computer Speech and Language*, 25(1), 105–126.
- Fredrickson, B., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought–action repertoires. *Cognition & Emotion*, 19(3), 313–332.
- Frijda, N. H. (2009). Emotions, individual differences and time course: reflections. *Cognition & Emotion*, 23(7), 1444–1461.
- Graesser, A., Chipman, P., Haynes, B., & Olney, A. (2005). AutoTutor: an intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612–618.
- Graesser, A., Chipman, P., King, B., McDaniel, B., & D'Mello, S. (2007). Emotions and learning with AutoTutor. In R. Luckin, K. Koedinger, & J. Greer (Eds.), *13th International conference on artificial intelligence in education* (pp. 569–571). Amsterdam: IOS Press.
- Graesser, A., Lu, S., Olde, B., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: comprehending illustrated texts on devices when the devices break down. *Memory and Cognition*, 33, 1235–1247.
- Graesser, A., Lu, S. L., Jackson, G., Mitchell, H., Ventura, M., Olney, A., et al. (2004). AutoTutor: a tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers*, 36, 180–193.
- Graesser, A., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. (2006). Detection of emotions during learning with AutoTutor. In R. Sun, & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the cognitive science society* (pp. 285–290). Austin, TX: Cognitive Science Society.
- Graesser, A., & Olde, B. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, 95(3), 524–536.
- Graesser, A., Ozuru, Y., & Sullins, J. (2010). What is a good question? In M. McKeown, & G. Kucan (Eds.), *Bringing reading research to life* (pp. 112–141) New York: Guilford.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley & Sons.
- Huk, T., & Ludwigs, S. (2009). Combining cognitive and affective support in order to promote learning. *Learning and Instruction*, 19(6), 495–505.
- Immordino-Yang, M. H., & Damasio, A. R. (2007). We feel, therefore we learn: the relevance of affective and social neuroscience to education. *Mind, Brain and Education*, 1(1), 3–10.
- Isen, A. (2008). Some ways in which positive affect influences decision making and problem solving. In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of emotions* (3rd ed.). (pp. 548–573) New York, NY: Guilford.
- Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, 2(3), 260–280.
- Jarvenoja, H., & Jarvela, S. (2005). How students describe the sources of their emotional and motivational experiences during the learning process: a qualitative approach. *Learning and Instruction*, 15(5), 465–480.
- Jurafsky, D., & Martin, J. (2008). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.
- Kapoor, A., Burleson, B., & Picard, R. (2007). Automatic prediction of frustration. *International Journal of Human–Computer Studies*, 65(8), 724–736.
- Keltner, D., & Shiota, M. (2003). New displays and new emotions: a commentary on Rozin and Cohen (2003). *Emotion*, 3, 86–91.
- Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar – an architecture for general intelligence. *Artificial Intelligence*, 33(1), 1–64.
- Larson, R. W., & Richards, M. H. (1991). Boredom in the middle school years – blaming schools versus blaming students. *American Journal of Education*, 99(4), 418–443.
- Lazarus, R. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Lehman, B., D'Mello, S., Chauncey, A., Gross, M., Dobbins, A., Wallace, P., et al. (2011). Inducing and tracking confusion with contradictions during critical thinking and scientific reasoning. In S. Bull, & G. Biswas (Eds.), *Proceedings of the 15th international conference on artificial intelligence in education* (pp. 171–178). New York/Heidelberg: Springer.
- Lepper, M., & Woolverton, M. (2002). The wisdom of practice: lessons learned from the study of highly effective tutors. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 135–158). Orlando, FL: Academic Press.
- Linnenbrink, E. A. (2006). Emotion research in education: theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review*, 18(4), 307–314.
- Mandler, G. (1976). *Mind and emotion*. New York: Wiley.
- Mandler, G. (1984). *Mind and body: Psychology of emotion and stress*. New York: W.W. Norton & Company.
- Mandler, G. (1999). Emotion. In B. M. Bly, & D. E. Rumelhart (Eds.), *Cognitive science. Handbook of perception and cognition* (2nd ed.). (pp. 367–382) San Diego, CA: Academic Press.
- Meyer, D., & Turner, J. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychology Review*, 18(4), 377–390.
- Miyake, N., & Norman, D. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, 18(3), 357–364.
- Morimoto, C., Koons, D., Amir, A., & Flickner, M. (1998). *Pupil detection and tracking using multiple light sources*. IBM: Almaden Research Center.
- Pekrun, R. (2006). The control–value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341.
- Pekrun, R. (2010). Academic emotions. In T. Urdan (Ed.), *APA educational psychology handbook, Vol. 2*. Washington, DC: American Psychological Association.
- Pekrun, R., Elliot, A., & Maier, M. (2006). Achievement goals and discrete achievement emotions: a theoretical model and prospective test. *Journal of Educational Psychology*, 98(3), 583–597.
- Pekrun, R., Goetz, T., Daniels, L., Stupnisky, R. H., & Raymond, P. (2010). Boredom in achievement settings: exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3), 531–549.
- Piaget, J. (1952). *The origins of intelligence*. New York: International University Press.
- Robinson, W. P. (1975). Boredom at school. *British Journal of Educational Psychology*, 45, 141–152.
- Rodrigo, M., & Baker, R. (2011). Comparing the incidence and persistence of learners' affect during interactions with different educational software packages. In R. Calvo, & S. D'Mello (Eds.), *New perspective on affect and learning technologies* (pp. 183–202). New York: Springer.
- Rosenberg, E. (1998). Levels of analysis and the organization of affect. *Review of General Psychology*, 2(3), 247–270.
- Rosenberg, E., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition & Emotion*, 8(3), 201–229.
- Rozin, P., & Cohen, A. (2003). High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of Americans. *Emotion*, 3, 68–75.
- Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145–172.
- Schank, R. (1999). *Dynamic memory revisited*. Cambridge University Press.
- Scherer, K. R. (2009). The dynamic architecture of emotion: evidence for the component process model. *Cognition & Emotion*, 23(7), 1307–1351.
- Schutz, P., & Pekrun, R. (Eds.). (2007). *Emotion in education*. San Diego, CA: Academic Press.
- Schwarz, N. Feelings-as-information theory. In P. Van Lange, A. Kruglanski, T. Higgins (Eds.), *Handbook of theories of social psychology*. Sage, in press.
- Siegler, R., & Jenkins, E. (Eds.). (1989). *Strategy discovery and strategy generalization*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Silvia, P. J. (2009). Looking past pleasure: anger, confusion, disgust, pride, surprise, and other unusual aesthetic emotions. *Psychology of Aesthetics Creativity and the Arts*, 3(1), 48–51.
- Stein, N., & Levine, L. (1991). Making sense out of emotion. In W. Kessen, A. Ortony, & F. Kraik (Eds.), *Memories, thoughts, and emotions: Essays in honor of George Mandler* (pp. 295–322). Hillsdale, NJ: Erlbaum.
- Tekscan. (1997). *Body pressure measurement system user's manual*. South Boston, MA: Tekscan Inc..
- VanLehn, K., Graesser, A., Jackson, G., Jordan, P., Olney, A., & Rose, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, 31(1), 3–62.
- VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, 21(3), 209–249.
- Vavik, L. (1993). Facilitating discovery learning in computer-based simulation learning environments. In R. Tennyson, & A. Baron (Eds.), *Automating instructional design: Computer-based development and delivery tools* (pp. 403–449). Berlin, Germany: Springer-Verlag.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.
- Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: recognizing and responding to student affect. *International Journal of Learning Technology*, 4(3/4), 129–163.
- Zimmerman, B. (2000). Attaining self-regulation: a social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation*. Academic Press.