15

CONFUSION

Sidney K. D'Mello and Arthur C. Graesser,
University of Notre Dame and University of Memphis

> A man may be absorbed in the deepest thought, and his brow will remain smooth until he encounters some obstacle in his train of reasoning, or is interrupted by some disturbance, and then a frown passes like a shadow over his brow.
>
> (Darwin, 1872, p. 220)

Almost a century and a half ago, Darwin published his seminal book *The Expression of Emotions in Man and Animals* that arguably launched the scientific study of emotion. In that book, he made a number of astute observations on frowns, the contexts that elicit them, their evolutionary value, their special status in the arsenal of human expressions, and their ubiquity as a form of emotional expression from infancy to mortality. Darwin observed that frowns often accompanied incongruence during deep thought and effortful deliberation, but not during simple reflection, orientation of attention, or meditation. He reasoned that frowns were an expressive correlate of the intention to focus attention on distant objects, which can be achieved by contracting the eye muscles so as to restrict incoming light to objects of immediate relevance. Though once associated with voluntary muscle control in the service of visual perception, Darwin hypothesized that through millions of years of evolution, the frown was involuntarily associated with information seeking, as is the case when one encounters a disruption in a train of thought.

Although not explicitly mentioned by Darwin, in some contexts, the furrowed brow is accompanied by feelings of *cognitive disequilibrium* (Piaget, 1952), or *cognitive*

Author Notes: The authors are grateful to members of the Emotive Computing Lab who have contributed to the research synthesized in this chapter (http://emotion.autotutor.org). This research was supported by the National Science Foundation (NSF) (ITR 0325428, HCC 0834847, DRL 1235958). Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF.

dissonance (Festinger, 1957), and the experience of *confusion.* Cognitive disequilibrium and confusion are triggered when individuals encounter incongruence in the form of impasses, anomalies, contradictions, disruptions of goals, extreme novelty that cannot be comprehended, and interruptions of organized sequences of actions. The importance of cognitive disequilibrium and cognitive dissonance in learning has a long history in psychology that spans the developmental, social, and cognitive sciences (Berlyne, 1960; Chinn & Brewer, 1993; Collins, 1974; Festinger, 1957; Graesser & Olde, 2003; Laird, Newell, & Rosenbloom, 1987; Mandler, 1976; Mugny & Doise, 1978; Piaget, 1952; Schank, 1999). The notion that cognitive disequilibrium extends beyond cognition and into emotions has also been acknowledged and investigated for decades. What is less clear, however, is the nature of the affective processes that are spawned by cognitive disequilibrium and how affect and cognition interact during learning. The focus on this chapter is on *confusion,* which is hypothesized to be the affective signature of cognitive disequilibrium and is expected to be highly relevant to both the processes and products of learning.

In our view, confusion is central to complex learning activities, such as comprehending difficult texts, generating cohesive arguments, solving challenging problems, and modeling complex systems. It is an inevitable consequence of effortful information processing, yet it has received considerably less attention in the mainstream scientific literature. Within the affective sciences, studies on confusion are essentially nonexistent when compared to emotions such as disgust and anger. Fortunately, there have been some recent efforts to investigate the phenomenon of confusion more carefully. This chapter synthesizes some of this literature with an emphasis on research on emotions and learning that we have conducted over the last decade. Our analysis of confusion is organized around seven fundamental questions: (a) What is confusion? (b) What are the appraisals that lead to confusion? (c) How is confusion expressed? (d) What are the temporal dynamics of confusion? (e) How is confusion regulated? (f) Why is confusion relevant to learning? and (g) When is confusion beneficial to learning? We conclude by discussing some of the implications of the findings, list open issues, and highlight opportunities in the scientific study of confusion.

WHAT IS CONFUSION?

The theoretical status of confusion in the affective sciences is quite mixed. Confusion has been considered to be a bona fide emotion (Rozin & Cohen, 2003a), a knowledge emotion (Silvia, 2010), an epistemic emotion (Pekrun & Stephens, 2011), an affective state but not an emotion (Hess, 2003; Keltner & Shiota, 2003), and a cognitive feeling state (Clore, 1992). It is beyond the scope of this chapter to go into these various conceptualizations of confusion, but the reader is referred to Rozin and Cohen (2003a, 2003b), Ellsworth (2003), Keltner and Shiota (2003), and Hess (2003) for an informative debate on the reasons for and against categorizing confusion as an emotion versus an affective or feeling state. In general, the confusion about the theoretical status of confusion as an emotion arises from (a) a lack of a clear definition of emotion, (b) multiple perspectives of emotion (Izard, 2010), and (c) a general paucity of basic research on confusion within the affective sciences (Rozin & Cohen, 2003b). This suggests that it might be useful to first ask a more basic question, "What is an emotion?" and then examine if, and to what extent, available data supports the classification of confusion as an emotion.

It is not surprising that the term emotion has stubbornly resisted any formal and widely accepted definition. To address this, Carroll Izard (2010), a noted emotion researcher, recently adopted a somewhat innovative approach to identify the defining characteristics of emotion. He asked 37 leading emotion researchers to provide written responses to six fundamental questions related to emotion. The question of interest to this chapter is "What is an emotion?" The results of a qualitative analysis of the written responses, published in a manuscript aptly titled "The Many Meanings/Aspects of Emotion: Definitions, Functions, Activation, and Regulation," yielded that emotions (a) involve neural circuits partially dedicated to "emotional processing" (8.92), (b) activate response systems in preparation for action (8.61), (c) have distinct feeling states (7.84), (d) play a role in expressive behavior and signaling systems (i.e., social functions) (6.56), (e) arise from results of appraisal processes (6.54), and (f) may involve cognitive interpretation of feelings (4.79). The numbers in parentheses beside each component reflect the extent to which a subset of the 37 respondents agreed on each of these six components in a subsequent survey on a scale ranging from 1 (not at all) to 10 (completely).

The six components identified by Izard (2010) reflect a number of different traditions, theories, and perspectives that have emerged over the last century (see Gross & Barrett, 2011, for a review). Hence, it is useful to consider whether confusion shares these six characteristics of an emotion. The case can easily be made for four of the components (items a, c, d, and e). First, it is clear that confusion arises from some form of neural interaction, as is the case when anomalies trigger EEG activities of the N400 (a negative event-related potential with a 400 ms post-stimulus onset) (Halgren et al., 2002; Kutas & Hillyard, 1980). Although the field of affective neuroscience is still in its infancy, it is unlikely that there is specific neural circuit or substrate dedicated solely to confusion. However, this should not weaken the status of confusion as an emotion because there is considerable debate as to whether specialized neural circuits exist for widely accepted emotions such as anger and disgust (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2011). Second, there is a distinct feeling state (subjective experience) that accompanies confusion (Rozin & Cohen, 2003a), although it is important to distinguish the form of short-term confusion that we are referring to here with long-term *mental confusion.* The latter is a pathological condition associated with mental disorientation and is symptomatic of dementia and other mental disorders (de Smet et al., 1982). Third, confusion has an expressive component consisting of the furrowed brow as initially noted by Darwin and subsequently confirmed in research studies (Craig, D'Mello, Witherspoon, & Graesser, 2008; Grafsgaard, Boyer, & Lester, 2011; McDaniel et al., 2007). Observers can also detect confusion from facial cues (Graesser et al., 2006). It is too early to say if there is a distinct facial expression for confusion, although failure to find one should not disqualify it as an emotion because despite decades of research, it is unclear if distinct facial expressions accompany emotions such as anger and disgust (Barrett, 2006; Russell, Bachorowski, & Fernandez-Dols, 2003). Fourth, as will be discussed in the next section, there is evidence to suggest that confusion arises from a cognitive appraisal of a mismatch between incoming information and existing knowledge (D'Mello, Lehman, Pekrun, & Graesser, 2014; Silvia, 2010).

It is unclear to what extent confusion involves bodily response systems via changes in physiology and priming of actions (item b). The lack of available data is not due to a failure to associate confusion with specific bodily changes, but rather stems from the lack of systematic research on how confusion is manifested in the body and how it recruits action systems (Rozin & Cohen, 2003a). As will be subsequently discussed in the section on expression, physiological-based machine learning models have recently achieved some success in distinguishing confusion from the neutral state and discriminating confusion from other emotions (AlZoubi, D'Mello, & Calvo, 2012). Though far from conclusive, this suggests that there is a link between confusion and the underlying physiology. Finally, it is unclear if confusion is a cognitive interpretation of a feeling (item f), an idea that originated with James (1884) and has been in and out of fashion for over a century. This was the sixth criteria listed by Izard (2010), and it obtained the lowest ratings (4.79 out of 10), so we will not let it influence confusion's fate as an emotion.

In summary, it is possible to make an initial case for confusion as an emotion because it arises out of neural interactions, involves bodily response systems, has a distinct feeling state, has an expressive component, and is an antecedent of cognitive appraisal. Confusion might also be considered to be an epistemic emotion or a knowledge emotion (Pekrun & Stephens, 2011; Silvia, 2010) since it arises out of information-oriented appraisals of external or internal knowledge (see next section). Some evidence indicates that confusion is likely perceived as a negative activating emotion (i.e., negative valence + moderate arousal) (Sazzad, AlZoubi, Calvo, & D'Mello, 2011) and can be positioned in the upper left quadrant of the Circumplex (see Russell, 1980 for details on the Circumplex model of affect). This categorization of confusion as an emotion should be taken to be tentative until there is more data to support or refute this position.

WHAT ARE THE APPRAISALS THAT LEAD TO CONFUSION?

The categorization of confusion as a knowledge emotion or an epistemic emotion implies that it has something to do with the state of an individual's knowledge. Indeed, confusion is hypothesized to occur when there is a mismatch of information, a violation of expectations, and other clashes of cognition during the processing of information. According to Mandler's interruption (discrepancy) theory (Mandler, 1990), individuals are constantly assimilating new information into existing knowledge schemas (e.g., an existing mental model). When new or discrepant information is detected (e.g., a conflict with prior knowledge or expectations), attention shifts to discrepant information, arousal increases in the autonomic nervous system, and the individual experiences a variety of possible emotions, depending on the context, the amount of change, and other relevant appraisals. Surprise is expected to occur when the degree of unexpectedness is high. Confusion is hypothesized to occur when there is a mismatch between incoming information and prior knowledge, or when new information cannot be integrated into existing mental models, thereby initiating cognitive disequilibrium. Confusion and surprise need not be mutually exclusive since surprise can precede confusion when an unexpected stimulus is appraised as being incomprehensible (Silvia, 2010).

Kagan (2009) provides a useful framework to discriminate among different states of uncertainty that are induced when low probability events are encountered. He identifies eight distinct states that emerge from appraising stimuli with respect to *familiarity* (familiar vs. unfamiliar), *expectation* (expected vs. unexpected), and *outcome* (desired vs. aversive). For example, a sudden clash of thunder while taking a stroll on a sunny day can be categorized as familiar (because one has heard thunder before), unexpected (because it is a sunny day), and aversive (because one has no umbrella). Uncertainty and confusion are expected to be maximized when the situation is unfamiliar, unexpected, and somewhat aversive, but this is entirely an empirical question.

Kagan also distinguishes *stimulus novelty,* which occurs when the unexpected events pertain to sensory information, from *conceptual novelty,* which is related to a mismatch of expectations in terms of an individual's knowledge structures and existing schemas. For example, hearing an unexpected high-pitched tone while learning Newtonian physics would be an example of stimulus novelty. On the other hand, watching a simulation of an elephant and a pebble being dropped from a skyscraper and noting that they both hit the ground at the same time would be conceptually novel if the individual has a fundamental misconception of Newton's second law of motion (i.e., the individual believes that heavier objects accelerate faster during free fall). The confusion that stems from conceptually novel events is of relevance to learning.

We have conducted a number of experiments to test the claim that confusion is elicited when there is conceptual novelty stemming from expectation violations and the presence of discrepancies in the information stream. In one set of experiments, confusion was induced while individuals performed a device comprehension task, such as trying to understand how toasters, doorbells, and other devices work from studying technical illustrated texts (D'Mello & Graesser, in review). The experimental trials consisted of presenting individuals with descriptions of device breakdowns (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 malfunction after they had studied a functioning device and had constructed a mental model of how it functions under normal operating conditions. The control trials simply involved comprehending the illustrated text without any breakdown descriptions. Confusion was measured via online self-reports after studying each device and was reported at significantly higher levels in the experimental trials than the control trials.

Contradictions are hypothesized to be another class of discrepant events that can induce confusion. We tested this hypothesis in three experiments that induced confusion by planting contradictory information during the learning of research methods. Specifically, learners discussed the scientific merits of sample research studies with two animated pedagogical agents: a tutor agent and a peer learner agent (D'Mello et al., 2014; Lehman et al., 2011). Contradictory trials involved the two animated agents expressing divergent opinions (one inaccurate and the other accurate or both inaccurate) and asking the (human) learners to decide which opinion had more scientific merit. Confusion was measured via a cued-recall procedure where participants made affect judgments by viewing videos of their faces and screens that were recorded during the learning task, with online self-reports, and by analyzing their response patterns (accuracy and consistency) immediately following contradictory trials. There was significantly higher confusion in the contradictory trials compared to the control trials that had no contradictions or inaccuracies.

Feedback plays an important role in learning because it is *directive* (i.e., tells learners what needs to be fixed), *facilitative* (i.e., helps learners conceptualize information), and has *motivational* functions (Shute, 2008). What are the consequences of false or inaccurate feedback? Will the novelty and violation of expectations caused by false feedback yield confusion? Indeed, a recent experiment indicated that learners self-reported more confusion and had longer response times when they received inaccurate feedback (i.e., correct responses received negative feedback from the computer tutor) (Lehman, D'Mello, & Graesser, 2012) compared to accurate feedback.

Earlier we categorized confusion as a knowledge emotion because it involves appraisals of information. It is useful to ascertain the extent to which the appraisal structure of confusion is aligned with other knowledge emotions, such as interest and surprise. Silvia (2010) posits that confusion and interest share an appraisal space consisting of novelty (familiar vs. unfamiliar) and comprehensibility (low vs. high). While both confusion and interest are expected to be triggered by highly novel events, Silvia hypothesized that confusion would be associated with appraisals of low comprehensibility, while interest would arise from high comprehensibility appraisals. In other words, a novel stimulus that could not be understood would be confusing, but a novel stimulus that could be understood would spark interest. This hypothesis was confirmed in an experiment involving comprehension of novel poems that were either comprehendible because background information required to understand the poem was provided (thereby triggering interest) or not comprehendible when participants had no background information (triggering confusion).

In summary, these experiments indicate that unexpected discrepant events induce confusion. This has been observed when the discrepancy is in the form of breakdowns, contradictions, false feedback, or the presentation of novel information that cannot be easily comprehended.

HOW IS CONFUSION EXPRESSED?

As Williams James (1884) put it so eloquently, "if we fancy some strong emotion, and then try to abstract from our consciousness of it all the feelings of its characteristic bodily symptoms, we find we have nothing left behind, no 'mind-stuff' out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains" (p. 193). Taking a cue from James that emotions and their expressions are inextricably coupled, we consider how confusion is expressed via the face, speech, posture, physiology, and language. Our emphasis is on studies that investigate naturalistic expressions of confusion instead of acted or posed expressions.

There are two primary methods of investigating the expressive components of emotion. The *theory-guided* approach focuses on a small set of expressions or actions (e.g., puckered lips, rises in pitch, forward leans) that have some theoretical-grounding as an expressive component of an emotion (see Russell et al., 2003, for a review). The advantage of this approach is that it affords the systematic testing of theory and yields highly interpretable expressive models of emotion. The disadvantage of this approach is that a large number of potential cues are ignored because they have no adequate grounding in theory. For example, it might be difficult to advance a theory as to why the kurtosis of the third formant of a speech signal is diagnostic of confusion. Should this feature simply be ignored in our quest for the vocal correlates of confusion?

The second is more of a *data-driven approach* that consists of computing large feature sets (potentially in the thousands) and applying automated data mining techniques (specifically machine learning) to narrow the feature space by identifying features that correlate with human-provided judgments of confusion, such as self-reports, online observations by researchers, or coding of video (see Calvo & D'Mello, 2010, for a review of these studies). The advantage is that this method has the potential to identify complex features that no theoretician would conjure a priori. The disadvantages are the potential lack of alignment with theory, the increased risk of Type I errors (although this risk can be eliminated with appropriate cross-validation methods), and problems interpreting some of the predictive models (as is the case when a neural network is used for prediction). The subsequent review includes both approaches. Aside from philosophical differences that are unlikely to ever be resolved, both systematic decoding studies (theory-guided approach) and data mining (data-driven approach) offer useful insights into how confusion is expressed.

Several of our findings pertaining to the expressive components of confusion were obtained in a study involving 28 learners who completed a 32-minute tutorial session with AutoTutor (D'Mello, Craig, Witherspoon, McDaniel, & Graesser, 2008; D'Mello & Graesser, 2009, 2010b), an intelligent tutoring system with conversational dialogue (Graesser, Chipman, Haynes, & Olney, 2005). This study is henceforth referred to as the *AutoTutor Multiple Judge Study.* Videos of the learners' faces, their computer screens, posture patterns, and logs of the interaction were recorded during the tutorial session. Approximately 100 judgments of each learner's emotions (boredom, flow/engagement, confusion, frustration, delight, surprise, and neutral) were provided by the learners themselves (self-report), untrained peers, and two trained judges via a cued-recall protocol (Graesser et al., 2006). The primary analysis consisted of extracting features from each of the informational streams and linking them to specific emotions using traditional statistical techniques as well as more advanced machine learning methods. The findings specific to different modalities are discussed below.

Facial Expressions

Darwin's (1872) observations about the emergence of frowns during disruptions of thought has been systematically confirmed in the few studies that have investigated the facial correlates of confusion. Using a theory-guided approach, Craig and colleagues (2008) performed an emote-aloud study where seven learners verbally expressed their emotions (as they occurred) during interactions with AutoTutor. Video recordings of learners' faces were manually coded for facial movements using the Facial Action Coding System (FACS; Ekman & Friesen, 1978). The Action Units (AUs) were correlated with online verbal reports of confusion. They found that a lowered brow (AU4), tightened lids (AU7), and combinations of these two facial movements (AU4 + AU7) were associated with confused expressions (see Figure 15.1). The lip corner puller (AU12) yielded a weaker but notable association with confusion. In a subsequent study (McDaniel et al., 2007),

Figure 15.1 Facial expressions of confusion.

FACS coding was performed on the videos collected in the *AutoTutor Multiple Judge Study,* and the observed AUs were correlated with affect judgments provided by two trained judges (as stated previously). Once again, the furrowed brow with tightened lids (AU4 + AU7) was predictive of confusion, although there was a notable lack of a link between AU12 and the expression of confusion. There is some additional converging evidence that is suggestive of the link between brow movements and confusion (Grafsgaard et al., 2011; Rozin & Cohen, 2003a), but more work is needed to identify additional facial indicators of confusion if they exist.

Speech Contours

Speech transmits affective information though the explicit linguistic message (what is said) and the implicit paralinguistic features of the expression (how it is said). Although it is clear that affective information is encoded and decoded through speech, there is also some ambiguity with respect to how different acoustic features communicate different emotions. One reliable finding is that pitch (*fo* or fundamental frequency) appears to be a reliable index into arousal (Johnstone & Scherer, 2000). Pitch has also been identified as a positive predictor of uncertainty (Forbes-Riley & Litman, 2011). This finding was obtained via a data-driven approach that involved regressing human-provided judgments of uncertainty on several acoustic and lexical features extracted from learner responses during one-on-one human-computer tutorial dialogues with the ITSpoke speech-enabled intelligent tutoring system. Future research is needed to replicate this finding and to identify additional vocal correlates of confusion.

Body Movements

Bodily movements are a much neglected but excellent channel to study the expression of emotion because the body is large and has multiple degrees of freedom. Bodily movements are presumably unconscious so they are less susceptible to social editing, at least when compared to the face and speech. We have analyzed how specific postures, as well as subtle changes in bodily fluctuations, are indicative of confusion and other emotions. For example, in the *AutoTutor Multiple Judge Study,* the pressure exerted on the back and seat of a pressure-sensitive chair was recorded during a tutorial session with AutoTutor (D'Mello & Graesser, 2009). When compared to the neutral state, confusion was accompanied by a decrease in the pressure exerted on the back of the chair without any accompanying increase on the seat. This is suggestive of an upright or alert posture (D'Mello & Graesser, 2010a).

In addition to specific postures, we have also investigated how subtle, presumably unconscious, bodily fluctuations covary with the experience of confusion and other emotions. We recently (D'Mello, Dale, & Graesser, 2012) tracked these movement dynamics using 1/f noise, pink noise, or fractal scaling during naturalistic experiences of affect in two studies involving deep learning and effortful problem solving. The results indicated that body movement fluctuations of individuals experiencing cognitive equilibrium was characteristic of correlated pink noise (i.e., an expected balance between determinism and randomness), but there was a whitening (i.e., more disorder or randomness) of the signal when individuals experienced states that are diagnostic of cognitive distress such as confusion.

Physiology

One of the key evolutionary functions of emotion is to prepare for rapid action in response to relevant environmental events. This call to action is accompanied by higher activation of the sympathetic nervous system. A large body of research has attempted to identify how different emotions are manifested in a number of physiological channels and devices such as electrocardiogram (ECG), electromyogram (EMG), galvanic skin response (GSR), respiration (RESP), skin temperature (ST), blood volume pressure (BVP), photoplethysmograph (PPG), impedance cardiogram (ICG), and electroencephalographs (EEG) (see Larsen, Berntson, Poehlmann, Ito, & Cacioppo, 2008, for a review). Although there has been some difficulty associated with identifying specific physiological responses for each emotion, physiological changes have been reliably linked to variations in arousal and sometimes valence (Barrett, 2006). The classification of confusion as a knowledge emotion raises the question of whether it has a specific physiological correlate, at least when compared to the more visceral emotions like disgust and fear.

This question was recently addressed by AlZoubi et al. (2012) who attempted to discriminate among several nonbasic emotions (e.g., confusion, curiosity) using ECG (electrical activity of the heart), EMG from the corrugator (brow) muscle, and GSR from finger tips. The physiological signals were collected while 27 learners completed a 45-minute tutorial session with AutoTutor. A total of 117 features were extracted from these three physiological channels and were used to predict self-reports of emotion obtained at 15-second intervals using a cued-recall procedure. They were able to obtain moderate accuracy in discriminating confusion from neutral and from other emotions. This suggests that confusion is to some extent manifested in physiology, although the exact nature of this manifestation is still unclear because the internals of machine learning models used in this research are not readily interpretable.

Language

Communication is one of the functions that is shared by language and emotions. It is perfectly clear that emotional content is routinely encoded in language as is the case when individuals write movie reviews, product reviews, blogs, and e-mail messages (see Pang & Lee, 2008, for an extensive review of sentiment analysis). But to what extent do individuals express emotions (particularly confusion) during learning? This question was investigated by analyzing 1,167 learner responses collected over the course of 28 tutorial interactions collected in the *AutoTutor Multiple Judge Study* (D'Mello & Graesser, 2012). We were only able to identify one occurrence of an explicit emotional expression ("I'm confused") in this corpus of learner utterances. Therefore, individuals experiencing confusion very rarely overtly label this emotion to a computer tutor. A similar finding was obtained in an analysis of transcripts from 50 tutorial sessions between learners and human tutors (D'Mello & Graesser, 2012). The lack of verbal emotion expressions in these learner utterances is somewhat surprising because an in-depth analysis of videos of both the human-computer and human-human sessions indicated that there were numerous emotional episodes (D'Mello & Graesser, 2012; Lehman, Matthews, D'Mello, & Person, 2008). This suggests that a more systematic textual analysis of tutorial dialogues might be necessary to uncover cues that might be diagnostic of learner emotions.

We explored this possibility by investigating the extent to which particular emotions are reflected in learner responses by considering a broad profile of language characteristics measured by the Linguistic Inquiry and Word Count (LIWC) (Pennebaker, Francis, & Booth, 2001) and Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004). LIWC is a validated computer tool that analyzes bodies of text using a large lexicon of words that have been rated on approximately 80 psychological and linguistic features. Coh-Metrix automatically analyzes text with respect to hundreds of measures of different types of cohesion (e.g., co-reference, referential, causal), genre, syntactic complexity, characteristics of words, and readability. Confusion was predicted by learner responses that were lacking in connectives (e.g., "hence," "because"), which is indicative of fragmented and less-cohesive responses. Confusion was also predicted by an increased use of inhibitory terms akin to "block," "constrain," and "stop" as measured by LIWC. This analysis revealed that although learners do not directly express their confusion, their responses inevitably convey their confusion by the words they use and by the connectives that hold their responses together.

Discourse Features and Contextual Cues

One advantage of investigating emotions with a dialogue-based intelligent tutoring system like AutoTutor is that the dialogue history provides a rich trace into the contextual underpinnings of learners' emotional experiences. To what extent is confusion manifested in these features of discourse and other conversational cues? To address this question, we analyzed the interaction logs collected in the *AutoTutor Multiple Judge Study.* Specifically, we examined the tutorial dialogue (i.e., the context) over 15-second intervals that culminated in episodes of confusion (D'Mello, Craig, Witherspoon, McDaniel, & Graesser, 2008). An event triggering confusion could either be tutor generated (e.g., the tutor provided a vague hint), learner generated (e.g., the learner has a misconception), or session related (e.g., early vs. late in the session). The results indicated that confusion occurred earlier in the session, within the first few attempts to answer a question, with slower and less verbose responses, with responses that had low conceptual quality, with frozen expressions (e.g., "I don't care" or "Please repeat" instead of domain-related contributions), when the tutor was less direct (i.e., more vague hints rather than explanations), and when the tutor provided negative feedback. These relationships between the various discourse features and confusion are generally in the expected directions.

WHAT ARE THE TEMPORAL DYNAMICS OF CONFUSION?

One aspect of confusion and of emotions in general that has not received sufficient attention is the chronometry or temporal dynamics of emotion. As an initial step to understanding temporal dynamics, we present a sketch of a model that predicts specific confusion trajectories on the basis of the severity of the discrepant event that triggers confusion and the results of confusion regulation processes. We also present some preliminary data that supports parts of the model.

The model assumes that individuals encounter discrepancies at multiple levels as they attempt to assimilate incoming information into existing mental models. There is some threshold T_a that needs to be exceeded before the individual is confused. Discrepancies that are not severe enough to exceed T_a are not detected by the individual, and there is no confusion. Sometimes the severity of the discrepancy greatly exceeds T_a, and the individual is bewildered or flustered. Let us denote this threshold as T_b.

A moderate level of confusion is experienced when the severity of the discrepancy meets or exceeds T_a but is less than T_b. The individual may not elect to attend to the confusion and shift attentional resources elsewhere. When this occurs, confusion is alleviated very quickly, and the length of confusion is less than duration D_a. If the length of the confusion episode exceeds D_a, then the individual has begun to attempt to identify the source of the discrepancy in order to resolve the confusion. When confusion resolution fails and the individual is confused for a long enough duration D_b, then there is the risk of frustration. With a longer duration D_c, there is a persistent frustration and the risk of disengagement and boredom (i.e., the learner gives up). There is potentially a *zone of optimal confusion*, which occurs when: T_a > *discrepancy* < T_b and D_a > *duration* < D_b.

Some evidence in support of this model can be obtained from some recent research that identified *confusion-engagement, confusion-frustration,* and *frustration-boredom* oscillations during interactions with AutoTutor (D'Mello & Graesser, 2012). These oscillations are depicted in Figure 15.2. The confusion-engagement transition is presumably linked to experiencing discrepancies (engagement to confusion) and successfully resolving the confusion (confusion to engagement). The confusion-frustration transition likely occurs when a learner experiences failure when attempting to resolve an impasse (confusion to frustration) and experiences additional impasse(s) when frustrated (frustration to confusion). Transitions involving boredom and frustration are ostensibly related to a state of being stuck due to persistent failure to the point of disengaging (frustration to boredom) and annoyance from being forced to persist in the task despite having mentally disengaged (boredom to frustration).

In addition to these transitions across states, we have also made some progress towards fitting exponential decay curves to study the decay characteristics of confusion

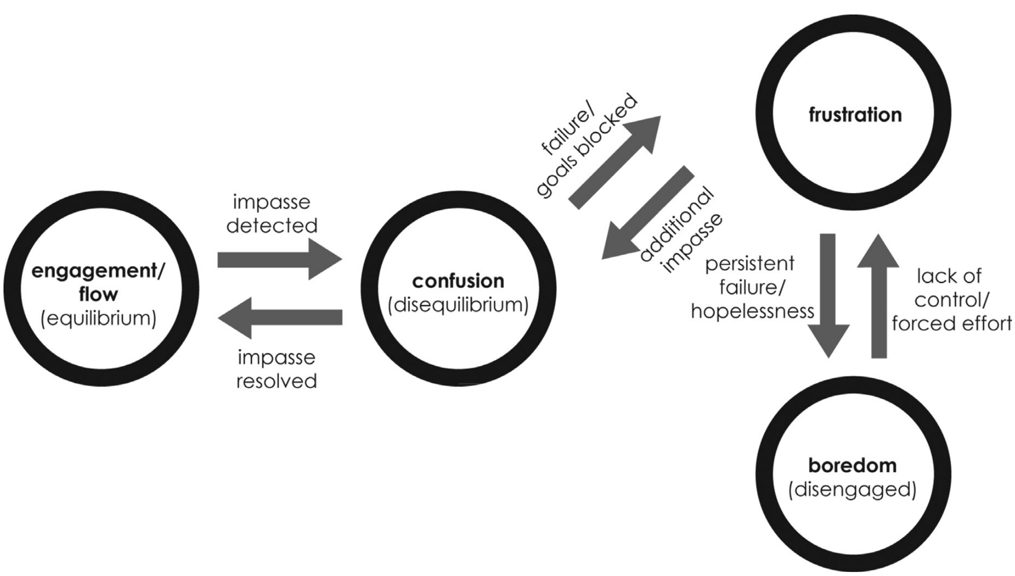

Figure 15.2 Affect transitions.

trajectories of individual learners (D'Mello & Graesser, 2011). What is missing, however, is the task of specifying and testing the various durations and thresholds of the model, which likely depend on some interaction between individual differences and the complexity of the materials and task. Systematically fitting these parameters in a manner that is sensitive to constraints of the individual, the environment, and their interaction is an important item for future work.

HOW IS CONFUSION REGULATED?

Individuals who are confused ideally pursue effective ways to regulate their confusion in order to restore equilibrium. Emotions theorists have identified a number of strategies that individuals enact to regulate their emotions. These include situation selection, situation modification, attentional deployment, cognitive change, and response modulation (Gross, 2008) (also see Jacobs & Gross, 2014). The first two strategies, situation selection and situation modification, are regulatory strategies aimed at selecting and modifying contexts (situations) that minimize or maximize the likelihood of experiencing certain emotions. Attentional deployment involves either attending to (e.g., ruminating) or avoiding (e.g., distraction) an object or event that can trigger an emotion. Cognitive reappraisal (Dandoy & Goldstein, 1990) involves changing the perceived meaning of a situation in order to alter its emotional content. Finally, response modulation involves a sustained effort to either overemphasize or minimize (e.g., suppression) the expression of an emotion.

There undoubtedly are individual differences in how learners experience and regulate confusion. Some learners might attempt to avoid confusion (and other negative emotions) by seeking out tasks with minimal intellectual challenges (situation selection), immediately seeking help when challenged (situation modification), avoiding attending to events within a situation that might be challenging (distraction/attentional deployment), intentionally ignoring or misattributing the cause of discrepant events to avoid confusion (reappraisal), and even withholding bodily expressions by adopting a poker face when confused (response modification). In contrast to these cautious learners, academic risk takers (Clifford, 1988) might engage in tasks that are intellectually stimulating (situation selection and modification), persevere on difficult problems, and consider challenges and failure to be necessary conditions to develop proficiency (reappraisal).

At this point in science, it is unclear if and to what extent learners utilize these strategies to regulate confusion. It is likely that confusion is perceived to be an aversive state, so learners who experience cognitive disequilibrium must resolve their confusion in order to restore equilibrium. Hence, one way to regulate confusion is to engage in cognitive activities to resolve the confusion, but this is only an assumption at this time. Four possible (but nonexclusive) trajectories of confusion dynamics as a function of the outcome of effortful resolution processes are shown in Figure 15.3.

One possibility is that confusion quickly rises, but it rapidly dissipates soon after (*quick rise and rapid dissipation* trajectory, see Figure 15.3a). It is possible that a learner might never fully resolve his or her confusion, and it might even increase as time progresses, thereby producing the *slow rise but never peak* trajectory depicted in Figure 15.3b. Alternately, confusion might adopt a *rise, peak, hold, decay* model (see Figure 15.3c) (Davidson, 1998; Rosenberg, 1998). According to this model, confusion gradually rises

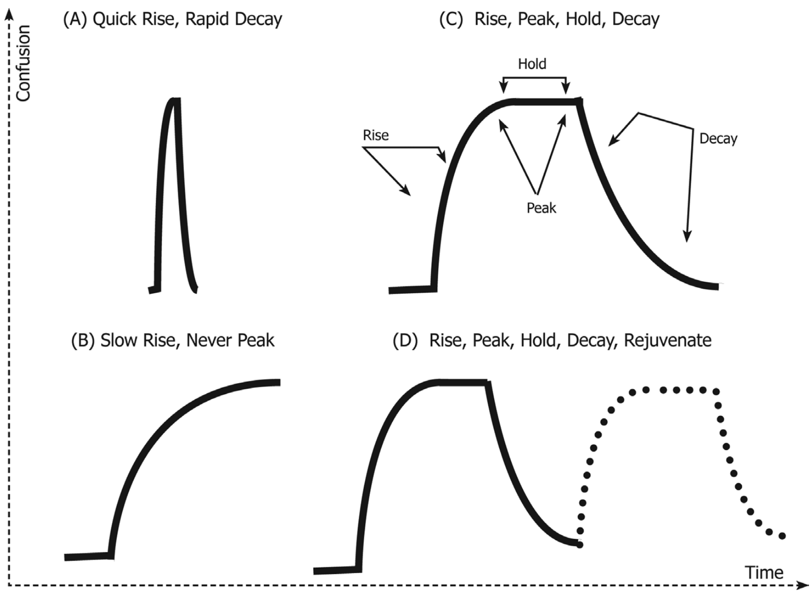

Figure 15.3 Confusion growth and decay dynamics.

until it peaks, presumably when an impasse is fully detected. Confusion is then held at its peak as the learner tries to resolve his or her confusion. Confusion begins to decay if and when the impasse is resolved or the source of a discrepancy is discovered. There is also the possibility that the learner might have not correctly resolved the impasse, and the rise, peak, hold, decay cycle is rejuvenated if a discrepancy is discovered (Figure 15.3d). Of critical importance is the observation that confusion is never fully resolved in the slow rise but never peak trajectory. This form of unresolved (hopeless) confusion is expected to accompany poor performance when compared to situations where confusion is immediately or eventually resolved.

The data from the device comprehension study (D'Mello & Graesser, in review) described earlier (see section on appraisals) was used to assess whether learners adhered to any or some of these trajectories and on the relationship between confusion resolution and learning. Specifically, learners participated in a cued-recall task in which they provided continuous confusion judgments by viewing videos of their faces that were recorded while they were attempting to diagnose the cause of device malfunctions. A second-by-second analysis of these confusion time series yielded two characteristic trajectories that successfully distinguished those learners who partially resolved their confusion (*rise, peak, hold, decay*) from those who remained confused (*slow rise but never peak*). As predicted, learners who partially resolved their confusion performed significantly better on a subsequent comprehension test than learners who remained confused. In addition to this study, Rodrigo and colleagues have reported some converging evidence to support this distinction between resolved and unresolved confusion and the differential impact of these processes on learning in more authentic contexts, such as learning computer programming in computer labs in schools (Lee, Rodrigo, Baker, Sugay, & Coronel, 2011; Rodrigo, Baker, & Nabos, 2010).

WHY IS CONFUSION RELEVANT TO LEARNING?

We have described a number of studies that have investigated different but related aspects of confusion. Although the context of these studies has been learning and problem-solving tasks, we now turn to the fundamental question of why confusion is relevant to learning. In our view, confusion plays a prominent role in learning activities that are pitched at deeper levels of comprehension and especially when the learner needs to bridge the gap between an existing (and usually faulty) mental model and an ideal conceptual model (Chi, 2008; Chinn & Brewer, 1993; Nersessian, 2008) (also see Sinatra, Broughton, & Lombardi, 2014). For example, a learner who has the faulty mental model that heavier objects accelerate faster than lighter objects during free-fall must confront this misconception in order to arrive at a mental model that is consistent with Newton's second law. The learner will be in a state of cognitive disequilibrium and experience confusion when they detect the misconception. The next section discusses some of the conditions where confusion might be beneficial to learning. Here, we focus on the incidence of confusion across multiple learning contexts.

In general, confusion is expected to be more the norm than the exception for complex learning tasks, such as learning the principles of ecological succession, comprehending a legal document, fixing a broken piece of equipment, and debugging errors in a computer program. Some compelling evidence to support this claim can be found in a recent meta-analysis that analyzed 24 studies that used a mixture of methodologies to systematically monitor the emotions (15 emotions plus neutral) of 1,740 middle school, high school, college, and adult learners in five countries over the course of more than 1,000 hours of continuous interactions with a range of learning technologies including intelligent tutoring systems, serious games, and simulation environments (D'Mello, 2013). The incidence of confusion was consistent with small or larger effects (i.e., Cohen's $d > 0.2$) compared to the other emotions in approximately half of the studies, which is reasonable given that the different learning environments varied with respect to the complexity of the learning task (e.g., writing an essay vs. learning about computer architecture). Confusion was found to be less frequent than engagement/flow, as frequent as boredom and happiness, somewhat more frequent than curiosity and frustration, and substantially more frequent than anxiety, contempt, delight, disgust, fear, sadness, and surprise. In addition to its prevalence during human-computer interactions, confusion has also been found to be quite frequent in human-human tutoring sessions. For example, Lehman and colleagues (2008) coded videos collected over the course of 50 hours of interactions between students and expert human tutors. They found that confusion was the most frequent emotion, comprising one third of all recorded emotion instances.

It should be noted that confusion is more than a mere incidental corollary of complex learning activities. Confusion is also related to learning outcomes. In a detailed analysis of human-human tutorial dialogues, VanLehn and colleagues (2003) reported that learning of conceptual physics concepts was rare if learners did not reach an impasse (which we assume to involve some level of confusion) irrespective of the explanations provided by the tutor. Craig, Graesser, Sullins, and Gholson (2004) conducted an online observational study in which the affective states (frustration, boredom, engagement/flow, confusion, eureka) of 34 learners were coded by observers every five minutes during interactions with AutoTutor. When learning gains were regressed on the incidence of the individual emotions, confusion was the only emotion that significantly predicted

6241-0277-PII-015.indd 302 1/15/2014 7:59:32 PM

learning. This finding of a positive correlation between confusion and learning has subsequently been replicated in follow-up studies with AutoTutor that used different methods to monitor emotions (D'Mello & Graesser, 2011; Graesser, Chipman, King, McDaniel, & D'Mello, 2007). Some recent data has also causally linked confusion and learning gains, but this depends on how confusion is attended to and the extent to which it is effectively regulated. This data is discussed in the next section.

WHEN IS CONFUSION BENEFICIAL TO LEARNING?

Confusion is expected to be beneficial to learning because it signals that there is something wrong with the current state of the world. This jolts the cognitive system out of equilibrium, focuses attention on the anomaly or discrepancy, and motivates learners to effortfully deliberate, problem solve, and restructure their cognitive system in order to resolve the confusion and return to a state of equilibrium. These activities inspire greater depth of processing, more durable memory representations, more successful retrieval, and consequently enhanced learning. It is not the confusion itself, but the cognitive activities that accompany its experience, that presumably influence learning. In this respect, confusion may not have a direct causal effect on learning, but rather serves as some form of a moderator on learning outcomes.

We have recently conducted three experiments to test for a moderation effect of confusion on learning (D'Mello et al., 2014; Lehman et al., 2011). These experiments were briefly introduced in the section on appraisals but are discussed in more detail in this section. The learning context for these experiments was the teaching of conceptual skills pertaining to scientific reasoning, such as stating hypotheses, identifying dependent and independent variables, isolating potential confounds in designs, interpreting trends in data, determining if data support predictions, and understanding effect sizes (Halpern, 2003; Millis et al., 2011). We developed a multimedia learning environment that attempted to teach these fundamental scientific inquiry skills by presenting example cases of studies (including the research design, participants, methods, results, and conclusions) that were frequently flawed because they violated principles of good research. Learners were instructed to evaluate the merits of the studies and point out flaws in the design.

The critiques of sample research studies were accomplished by holding multiturn trialogues with two embodied conversational agents and the human learner. One agent called the *tutor agent,* or *Dr. Williams,* led the tutorial lessons and served as an expert on scientific inquiry. The second agent, *Chris,* was the *peer-agent,* who simulated a peer of the human learner (i.e., the participant in the experiment). The human learners interacted with both agents by holding conversations in natural language that were designed to mimic human-human tutorial interactions.

Confusion was experimentally manipulated over the course of these multiturn trialogues by a manipulation of contradictory information. This occurred by having the animated agents occasionally disagreeing on ideas by voicing inaccurate information (experimental trials) and asking the human learner to intervene and decide which opinion had the most scientific merit. The source, timing, and content of the contradictions varied across conditions and experiments, details of which are beyond the scope of this chapter. What is important is that confusion was induced by providing misleading and sometimes incorrect information. However, all misleading information was corrected over the course of the trialogues, and learners were fully debriefed at the end of the experiment.

The results were illuminating in a number of respects. One finding was that the contradictions were quite effective in inducing confusion. Interestingly, the learners were somewhat reticent to admit that they were confused, but their underlying confusion was revealed through more objective measures consisting of their responses to probe questions immediately following the contradictions. As predicted, confusion moderated the effect of the contradictions on learning gains. Learning gains for contradictory trials were statistically equivalent to no-contradiction control trials when learners were not confused by the contradictions. However, learners who were confused by the contradictions had substantially higher learning gains in the contradictory trials than in the control trials. This effect was observed for simple multiple choice tests of knowledge and on subsequent transfer tests, some of which consisted of identifying flaws in case studies that were radically different than the case studies discussed during the trialogues.

Some of these effects have also been observed in a recently completed study where confusion was induced via a false feedback manipulation in lieu of contradictions (Lehman et al., in 2012). Learners who initially provided correct answers but received negative feedback reported more confusion, had longer response times immediately following the false feedback (processing incongruities), and spent more time studying an explanatory text (greater depth of processing) than controls. Importantly, learners demonstrated enhanced learning gains compared to those who received accurate feedback (positive feedback for correct responses), but only when they reported being confused by the feedback.

In summary, although systematic research on the potential facilitative effects of confusion on learning is in its infancy, there appear to be some measurable benefits to productively confusing learners in order to promote deeper inquiry. These findings, which highlight the beneficial role of confusion to learning, are consistent with Piaget's (1952) notion of *accommodation* because learners must, to some extent, alter their mental models in order to resolve their confusion. These findings also contribute to an impressive body of evidence on the facilitative effects of negative mood states on the process of accommodation; this literature is surveyed in considerable detail by Fiedler and Beier (2014). Although it is tempting to merely attribute the facilitative effects of confusion to the fact that it is a negatively valenced emotion, it is important to note that all negative affective states are not created alike. Indeed, there is a world of difference between a background negative mood state that subtly biases cognition and an intense experience of a negative emotion that overtakes cognition (Rosenberg, 1998). Frustration, for example, is a negative activating emotion (similar to confusion), but it is unlikely to yield any of the learning benefits associated with confusion. For that matter, neither are disgust, fear, or contempt.

SUMMARY, IMPLICATIONS, FUTURE WORK, AND CONCLUSIONS

The last decade has ushered in considerable excitement for research on emotions in the affective, learning, and computer sciences. Some landmarks include the launch of the APA journal *Emotion* in 2001, the launch of *Emotion Review* in 2009, Schutz and Pekrun's (2007) edited volume *Emotions in Education*, and numerous special issues on affect and its relationship with learning (e.g., Linnenbrink-Garcia & Pekrun, 2011).

Computer scientists and engineers are also fascinated by emotion, a movement that can be traced to Picard's (1997) book *Affective Computing.* The 2010 launch of *Transactions in Affective Computing,* a scholarly journal published by the Institute for Electrical and Electronic Engineers (IEEE), offers further evidence that we now live in a world of *computational emotions* (systems that sense, induce, respond to, and synthesize emotions).

We are also living in an era of interdisciplinary research as emotion, education, and computing researchers forgo traditional disciplinary boundaries in a collaborative effort to do basic research on emotions during learning and to leverage these insights towards the development of technologies that help students learn by coordinating emotion and cognition. Some of this emerging interdisciplinary research has been compiled in Calvo and D'Mello's (2011) edited volume *New Perspectives on Affect and Learning Technologies.* As with any burgeoning research area, there are currently more open questions than answers, but this only fuels interest and enthusiasm for more research.

In keeping with this interdisciplinary spirit, much of the research described in this chapter has adopted an interdisciplinary approach that has encompassed multiple theoretical frameworks, methodologies, and instruments to shed light on one ubiquitous but inconspicuous emotion—confusion. We made an effort to argue in favor of categorizing confusion as an emotion, discussed the appraisals that lead to confusion, examined how confusion is expressed across multiple modalities that encompass the mind and body, explored the temporal dynamics of confusion, and described how confusion might be regulated. After examining these interrelated aspects of confusion, we discussed why confusion is very relevant to learning and explored circumstances in which confusion moderates learning outcomes.

Many of the studies on confusion featured in this chapter have been laboratory studies with limited ecological validity. These studies have been instrumental in confirming some expected patterns (e.g., the link between a furrowed brow and expressions of confusion) and revealing some nonobvious patterns (e.g., positive correlation between confusion and learning). However, it is unclear whether these patterns will be observed in more authentic learning contexts where a large number of extraneous variables come into play. Replicating and extending these initial laboratory findings in classrooms and other learning situations would represent an important step forward. It is also highly likely that previously unforeseen patterns will be discovered when confusion is investigated in more authentic learning contexts.

We conclude this chapter by briefly describing some of the important implications, challenges, and opportunities for a research program centered on confusion. Although such a discussion can warrant a chapter in itself, we focus on three major points. First, the empirical status of confusion as an emotion currently suffers from a lack of positive evidence rather than a surplus of negative evidence. Hence, there is a pressing need for basic research to validate or disprove our tentative categorization of confusion as an emotion. The phenomenon of confusion itself is completely oblivious to its categorization as an emotion, a cognition, or a blend of the two, so one might question the utility of advancing a research program to test the *confusion as an emotion* hypothesis. Although we are sympathetic to this view, and have previously argued against the false cognition versus emotion dichotomy (Graesser & D'Mello, 2011), the reality is that the scientific study of confusion is likely to flourish if there is sufficient empirical evidence to elevate it to the privileged status of a bona fide emotion, on par with the basic emotions of happiness, sadness, fear, disgust, anger, and surprise. It is somewhat paradoxical that one must first conduct a large body of research on confusion to show that it is an emotion before researchers are encouraged to scientifically investigate confusion as an emotion.

Second, within the educational realm, there appear to be some learning benefits associated with confusion. A somewhat controversial implication of our research is that pedagogical practices that attempt to productively confuse learners might be attractive alternatives to the typical information delivery systems that are comfortable for passive learning but rarely promote deep insight. One can imagine a world where interventions that expose misconceptions might be cherished instead of chastised, complexity might be a valuable substitute or complement for clarity, and less cohesive texts and lectures might replace the polished information deliveries of textbooks and formal lectures. Learning of difficult conceptual material is chaotic, gritty, and confusing, so there might be advantages to interventions with embedded challenges and other desirable difficulties (Bjork & Bjork, 2011), especially if the goal is to promote learning at deeper levels of comprehension. We have not formally studied this issue, but we suspect that most students and teachers perceive confusion to be reflective of failure and negativity, so there is an initial challenge of changing this simplistic and somewhat inaccurate mindset.

To be clear, we are not advocating learning environments that intentionally confuse low-achieving learners, learners with minimal motivation, and learners who risk dropping out when there is hopeless and unproductive confusion. It is worth noting, however, that stemming from Piaget's (1952) theory of cognitive development, there have been several attempts at promoting conceptual change by inducing cognitive conflict in classrooms (see Limón, 2001, for a review of these studies), so our suggestions are not entirely radical. Nevertheless, there obviously is no one-size-fits-all approach to learning, so these somewhat unconventional interventions should be differentially and dynamically sensitive to individual learners. Adapting pedagogical strategies to individual learners is difficult to achieve in formal learning contexts, but this is precisely the niche in which advanced learning technologies excel. Intelligent tutoring systems have made significant advances in creating fine-grained models of learner knowledge and have leveraged these models to select learning trajectories that are optimized to individual learners (Corbett & Anderson, 1994; Koedinger & Corbett, 2006). These systems can be augmented with the ability to induce confusion at the appropriate time and with the appropriate level of discrepancies, track the induced confusion using state-of-the art affect detection systems (Calvo & D'Mello, 2010; D'Mello & Graesser, 2010b), and implement scaffolds that help learners regulate their confusion so that they correct problematic misconceptions, resolve impasses, and revise faulty mental models. This is exactly the sort of scientific and technological infrastructure that is needed to design interventions that keep learners balanced between the extremes of boredom and bewilderment by selecting materials and challenges within their *zones of optimal confusion.*

REFERENCES

AlZoubi, O., D'Mello, S. K., & Calvo, R. A. (2012). Detecting naturalistic expressions of nonbasic affect using physiological signals. *IEEE Transactions on Affective Computing, 3*(3), 298–310.Barrett, L. (2006). Are emotions natural kinds? *Perspectives on Psychological Science, 1,* 28–58.

Berlyne, D. (1960). *Conflict, arousal, and curiosity.* New York, NY: McGraw-Hill.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology* *and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York, NY: Worth Publishers.

Calvo, R. A., & D'Mello, S. K. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1*(1), 18–37. doi: 10.1109/T-AFFC.2010.1

Calvo, R. A., & D'Mello, S. K. (2011). *New perspectives on affect and learning technologies.* New York, NY: Springer.

Chi, M. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 61–82). New York, NY: Routledge.

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. doi: 10.2307/1170558

Clifford, M. (1988). Failure tolerance and academic risk-taking in ten- to twelve-year-old students. *British Journal of Educational Psychology, 58,*15–27. doi: 10.1111/j.2044–8279.1988.tb00875.x

Clore, G. L. (1992). Cognitive phenomenology: Feelings and the construction of judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 133–163). Hillsdale, NJ: Erlbaum.

Collins, A. (1974). Reasoning from incomplete knowledge. *Bulletin of the Psychonomic Society, 4,* 254–254.

Corbett, A., & Anderson, J. (1994). Knowledge tracing—Modeling the acquisition of procedural knowledge. *User Modeling And User-Adapted Interaction, 4*(4), 253–278.

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.

Dandoy, A. C., & Goldstein, A. G. (1990). The use of cognitive appraisal to reduce stress reactions—A replication. *Journal of Social Behavior and Personality, 5*(4), 275–285.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London, England: John Murray.

Davidson, R. J. (1998). Affective style and affective disorders: Perspectives from affective neuroscience. *Cognition & Emotion, 12,* 307–330.

de Smet, Y., Ruberg, M., Serdaru, M., Dubois, B., Lhermitte, G., & Agid, Y. (1982). Confusion, dementia and anticholinergics in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 45,* 1161–1164.

D'Mello, S. K. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology, *Journal of Educational Psychology, 105*(4), 1082–1099.

D'Mello, S., Craig, S., Witherspoon, A., McDaniel, B., & Graesser, A. (2008). Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction, 18*(1–2), 45–80.

D'Mello, S., Dale, R., & Graesser, A. (2012). Disequilibrium in the mind, disharmony in the body. *Cognition & Emotion, 26*(2), 362–374. doi: 10.1080/02699931.2011.613668

D'Mello, S., & Graesser, A. (2009). Automatic detection of learners' affect from gross body language. *Applied Artificial Intelligence, 23*(2), 123–150.

D'Mello, S., & Graesser, A. (2010a). Mining bodily patterns of affective experience during learning. In A. Merceron, P. Pavlik, & R. Baker (Eds.), *Proceedings of the third International Conference on Educational Data Mining* (pp. 31–40). International Educational Data Mining Society.

D'Mello, S., & Graesser, A. (2010b). 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. (2011). The half-life of cognitive-affective states during complex learning. *Cognition & Emotion, 25*(7), 1299–1308.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction, 22,* 145–157. doi: 10.1016/j.learninstruc.2011.10.001

D'Mello, S. K. & Graesser, A. C. (2012). Language and discourse are powerful signals of student emotions during tutoring. *IEEE Transactions on Learning Technologies, 5*(4), 304–317.D'Mello, S., & Graesser, A. (in review). Confusion and its dynamics during device comprehension with breakdown scenarios.

D'Mello, S. K., Lehman, B. Pekrun, R., & Graesser, A. C. (2014). Confusion can be beneficial for learning, *Learning & Instruction, 29*(1), 153–170.

Ekman, P., & Friesen, W. (1978). *The Facial Action Coding System: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Ellsworth, P. C. (2003). Confusion, concentration, and other emotions of interest: Commentary on Rozin and Cohen (2003). *Emotion, 3*(1), 81–85.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Fiedler, K., & Beier, S., (2014). Affect and cognitive processes in educational contexts. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *Handbook of emotions in education.* New York, NY: Taylor & Francis.

Forbes-Riley, K., & Litman, D. J. (2011). Benefits and challenges of real-time uncertainty detection and adaptation in a spoken dialogue computer tutor. *Speech Communication, 53*(9–10), 1115–1136. doi: 10.1016/j.specom.2011.02.006

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. doi: 10.1109/TE.2005.856149

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.), *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 569–571). Amsterdam, Netherlands: IOS Press.

Graesser, A., & D'Mello, S. (2011). Theoretical perspectives on affect and deep learning. In R. Calvo & S. D'Mello (Eds.), *New perspective on affect and learning technologies* (pp. 11–22). New York, NY: Springer.

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., McNamara, D., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36,* 193–202.

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. doi: 10.1037/0022–0663.95.3.524

Grafsgaard, J., Boyer, K., & Lester, J. (2011). Predicting facial indicators of confusion with hidden markov models. In S. D'Mello, A. Graesser, B. Schuller, & J. Martin (Eds.), *Proceedings of the 4th International Conference on Affective Computing and Intelligent Interaction (ACII 2011)* (pp. 97–106). Berlin Heidelberg, Germany: Springer.

Gross, J. (2008). Emotion regulation. In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 497–512). New York, NY: Guilford.

Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review, 3*(1), 8–16.

Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *NouroImage, 17*(3), 1101–1116.

Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking* (4th ed.). Mahwah, NJ: Erlbaum.

Hess, U. (2003). Now you see it, now you don't—the confusing case of confusion as an emotion: Commentary on Rozin and Cohen (2003). *Emotion, 3*(1), 76–80.

Izard, C. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review, 2*(4), 363–370. doi: 10.1177/1754073910374661

Jacobs, S. E., & Gross, J. J. (2014). Emotion regulation in education: Conceptual foundations, current applications, and future directions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *Handbook of emotions in education.* New York, NY: Taylor & Francis.

James, W. (1884). What is an emotion? *Mind, 9,* 188–205.

Johnstone, T., & Scherer, K. (2000). Vocal communication of emotion. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 220–235). New York, NY: Guilford Press.

Kagan, J. (2009). Categories of novelty and states of uncertainty. *Review of General Psychology, 13*(4), 290–301.

Keltner, D., & Shiota, M. (2003). New displays and new emotions: A commentary on Rozin and Cohen (2003). *Emotion, 3*(86–91). doi: 10.1037/1528–3542.3.1.86

Koedinger, K., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–78). New York, NY: Cambridge University Press.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar—an architecture for general intelligence. *Artificial Intelligence, 33*(1), 1–64. doi: 10.1016/0004–3702(87)90050–6

Larsen, J., Berntson, G., Poehlmann, K., Ito, T., & Cacioppo, J. (2008). The psychophysiology of emotion. In M. Lewis, J. Haviland-Jones, & L. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 180–195). New York, NY: Guilford.

Lee, D. M., Rodrigo, M. M., Baker, R. S., Sugay, J., & Coronel, A. (2011). Exploring the relationship between novice programmer confusion and achievement. In S. D'Mello, A. Graesser, B. Schuller, & J. Martin (Eds.), *Proceedings of the 4th bi-annual International Conference on Affective Computing and Intelligent Interaction* (pp. 175–184). Berlin, Germany: Springer.

Lehman, B., D'Mello, S., Chauncey, A., Gross, M., Dobbins, A., Wallace, P.… . Graesser, A. C. (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, NY: Springer.

Lehman, B., D'Mello, S. K., & Graesser, A. C. (2012). Confusion and complex learning during interactions with computer learning environments, *The Internet and Higher Education, 15*(3), 184–194.Lehman, B., Matthews, M., D'Mello, S., & Person, N. (2008). What are you feeling? Investigating student affective states during expert human tutoring sessions. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (pp. 50–59). Berlin, Germany: Springer.

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: a critical appraisal. *Learning and Instruction, 11*(4–5), 357–380. doi: 10.1016/s0959–4752(00)00037–2

Lindquist, K. A., Wager, T., D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2011). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences, 173*(4), 1–86.

Linnenbrink-Garcia, L., & Pekrun, R. (2011). Students' emotions and academic engagement: Introduction to the special issue. *Contemporary Educational Psychology, 36*(1), 1–3.

Mandler, G. (1976). *Mind and emotion.* New York, NY: Wiley.

Mandler, G. (1990). Interruption (discrepancy) theory: Review and extensions. In S. Fisher & C. L. Cooper (Eds.), *On the move: The psychology of change and Transition* (pp. 13–32). Chichester, United Kingdom: Wiley.

McDaniel, B., D'Mello, S., King, B., Chipman, P., Tapp, K., & Graesser, A. (2007). Facial features for affective state detection in learning environments. In D. McNamara & G. Trafton (Eds.), *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 467–472). Austin, TX: Cognitive Science Society.

Millis, K., Forsyth, C., Butler, H., Wallace, P., Graesser, A., & Halpern, D. (2011). Operation ARIES! A serious game for teaching scientific inquiry. In M. Ma, A. Oikonomou, & J. Lakhmi (Eds.), *Serious games and edutainment applications* (pp. 169–196). London, United Kingdom: Springer.

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology, 8*(2), 181–192.

Nersessian, N. (2008). Mental modeling in conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 391–416). New York, NY: Routledge.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1–135.

Pekrun, R., & Stephens, E. J. (2011). Academic emotions. In K. Harris, S. Graham, T. Urdan, S. Graham, J. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (pp. 3–31). Washington, DC: American Psychological Association.

Pennebaker, J., Francis, M., & Booth, R. (2001). *Linguistic inquiry and word count (LIWC): A computerized text analysis program.* Mahwah, NJ: Erlbaum.

Piaget, J. (1952). *The origins of intelligence.* New York, NY: International University Press.

Picard, R. (1997). *Affective computing.* Cambridge, MA: MIT Press.

Rodrigo, M., Baker, R., & Nabos, J. (2010). *The relationships between sequences of affective states and learner achievement.* Paper presented at the Proceedings of the 18th International Conference on Computers in Education, Putrajaya, Malaysia.

Rosenberg, E. (1998). Levels of analysis and the organization of affect. *Review of General Psychology, 2*(3), 247–270. doi: 10.1037//1089–2680.2.3.247

Rozin, P., & Cohen, A. (2003a). 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.

Rozin, P., & Cohen, A. B. (2003b). Reply to commentaries: Confusion infusions, suggestives, correctives, and other medicines. *Emotion, 3*(1), 92–96.
Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178.
Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54,* 329–349.
Sazzad, M. S., AlZoubi, O., Calvo, R. A., & D'Mello, S. K. (2011). Affect detection from multichannel physiology during learning. In S. Bull & G. Biswas (Eds.), *Proceedings of the 15th International Conference on Artificial Intelligence in Education* (pp. 131–138). New York, NY: Springer.
Schank, R. (1999). *Dynamic memory revisited.* Cambridge, England: Cambridge University Press.
Schutz, P., & Pekrun, R. (Eds.). (2007). *Emotion in education.* San Diego, CA: Academic Press.
Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.
Silvia, P. J. (2010). Confusion and interest: The role of knowledge emotions in aesthetic experience. *Psychology of Aesthetics Creativity and the Arts, 4,* 75–80. doi: 10.1037/a0017081
Sinatra, G. M., Broughton, S. H., & Lombardi, D. (2014). Emotions in science education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *Handbook of emotions in education.* New York, NY: Taylor & Francis.
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. doi: 10.1207/S1532690XCI2103_01