Catching Eureka on the Fly

Paulina Lindström and Agneta Gulz

Lund University Cognitive Science
Kungshuset Lundagård
SE-222 22 Lund
Paulina.Lindstrom@lucs.lu.se, Agneta.Gulz@lucs.lu.se

Abstract

An interesting class of intelligent tutoring systems are systems that are able to detect and recognize affective responses and states in students, and use such data to adapt its feedback, guidance and instruction. The article presents affective responses and states considered relevant for learning, and methods that have been proposed or used to detect such responses and states. We focus on the affective state of 'eureka', a relatively short-term state with corresponding difficulties of being detected. There are good reasons to attempt to catch this affective state, since it can mark when understanding is achieved and signal to an intelligent tutoring system that the student is ready to proceed in the learning process. A study is presented where a combination of the behavioural measure of eye tracking and two psycho-physiological measures – skin conductance and pupil size – are used in an attempt to identify this affective experience. The results showed significant differences on all measures between subjects who understood certain stimuli, and purportedly experienced eureka, and subjects who did not understand the stimuli. The psycho-physiological measure patterns were however not as dramatic as predicted. Finally, we discuss the potential benefits for students using ITS, that can detect eureka experiences.

Introduction

Successful tutoring is dependent on a tutor's ability to adapt her or his guidance and feedback, in content as well as in format and in style, to a student or a group of students. One source of information for a tutor in this adaptation process is the behavioural and verbal output from students: the moves they make in a problem-solving process, what they say in commenting an assignment or answering a question, etc. But expert tutors are also able to recognize students' *affective states*, such as whether a student is confused, engaged, curious, bored, etc., and respond towards this in ways that positively impact learning (Goleman 1995) – (even though there is data showing that this is not *always* the case (Graesser et al. 2006)). If our aim is to develop intelligent tutoring systems by means of artificial intelligence, it is thus of interest to develop systems capable of detecting, recognizing and interpreting such emotional responses and states in students, and to use such data to adapt teaching strategies and processes accordingly.

This article focuses on the affective state of eureka – or insight experienced as "Ah, now I get it" – and on the pedagogical potentials of an intelligent tutoring system that can identify this state. After a general discussion of learning and affect we go more specifically into eureka and learning. The following two sections focus on methods for the identification of affect and specifically the identification of eureka. A recent study is presented, and after discussing its results, we enter the discussion on the pedagogical potentials of an intelligent tutoring system that could identify the experience of eureka.

Learning and Affect

In recent years the role of emotions in learning has received increased attention (Kort, Reilly, and Picard 2001; Craig et al 2004; D'Mello et al. 2005; Graesser et al. 2006; Linnenbrink 2006; Meyer and Turner 2006; Schutz et al. 2006; Immordino-Yang and Damasio 2007). Note that this is a relatively novel sub-domain within the encompassing domain of cognition and emotion/affect, focusing on *learning* and on which, and how affective states are functionally significant in learning processes.

Kort, Reilly, and Picard (2001) describe a learning spiral where a learner moves through functionally different affect states in forming an understanding of a domain. For example, a learner interested in biology may be *curious* to know how a plant can grow. The learner is then presented with information about the photosynthesis and its role in the growing process. The learner might then be *confused* by some of the concepts. During the activity of making sense of the difficult concepts, a state of *flow* may occur. After a while, the photosynthesis is understood and the learner has incorporated the new knowledge coherently. The learner now has reached an understanding of how a plant can grow. In this way, affect states may provide cues to how the process of creating meaning proceeds in a learning situation.

Craig et al. (2004) present a study in which observations and emote-aloud protocols were exploited in order to identify affective states that are relevant from a learning point of view, such as frustration, boredom, flow, confusion and 'eureka'. The communication of affect states is useful since expressed affects provide cues to how the

tutor can adapt appropriately (D'Mello et al. 2006). For example, frustration communicates that help might be needed, whereas flow might indicate that a learner wants to be left alone to focus on whatever she or he is absorbed in. The Casey example (Burleson and Picard 2004) illustrates this well in the way the system offers just the needed amount of guidance and help, based on integrating various measures of affect states in the student.

The experience of eureka is not so often approached in the context of learning research, but it is the focus of the present article.

Learning and Eureka

An Affective Receipt

The experience of eureka or of *insight* is subjectively experienced as: *"Aha, now I get it!"* It is characterized by an initial impasse where the person at first does not know how to solve a complex situation or a difficult problem. Then suddenly, she or he 'sees' the answer and the problem is solved. In particular when the context involves *difficult material,* a *challenging task or a problem*, the process of finally grasping the inner or hidden nature of the complexity can be accompanied by an eureka or an aha-experience.

In a sense such an experience is an affective receipt to the learner as to whether she has understood or not. It can be a scaffold in the process of creating meaning, clearly telling her: *"You did get it now!", "You do understand this now!"* Such affective information functions as *external* and *internal communication* of success feedback to the student's learning or problem solving efforts. (Even though sometimes, an affective sense of understanding can be the result of hindsight and overconfidence (c.f. Trout 2002)).

There is also the case of *external communication* to the teacher or tutor that realizes or identifies the 'aha'-experience in a student. Questions that may then be asked by the tutor are the following: What were the factors behind the student's reaching insight here? Was there something in the explanation that provided the student with the right scaffold? Is there something in the present learning setting or context that is useful to consider for future learning situations?

This knowledge can certainly be educationally fruitful, especially when one is dealing with domains, problems or issues that are well-known to be difficult to understand and are frequently struggled with by students (Feltovich, Coulson, & Spiro. 2001).

Zooming in on the Moment of Understanding

If one is interested in zooming in on the moment of *understanding*, eureka is a good phenomenon to approach. Overall, it is well established that emotions are important in learning (Pons et al. 2005; Schutz and Pekrun 2007), but less is known about the moments of understanding. Gestalt theory claims that insight occurs because something is restructured in the mind (Wertheimer 1959; Köhler 1969). We predict that this important moment of insight is accompanied by detectable emotional responses, given that cognition and emotions are tightly connected.

The subjective experience of understanding in eureka seems relatively distinct and it is delimited in time. We expect to find out more about the emotional responses occurring when understanding takes place by using instruments and measures that can give other and more fine-grained information to supplement subjective reports of understanding. We wish to know more about the processes in time and about *when* understanding occurs – a challenge identified in the field of emotion and learning (Schutz el al. 2006).

Methods for Identification of Affect in Learning

There are several different methods for detecting and identifying affect in learning, involving various forms of instruments and technology. Burleson and Picard (2004) proposed a number of measures for detecting affect state when discussing the learning companion Casey, such as skin conductivity, facial analysis, mouse pressure and seating posture. Kapoor and Picard (2005) used a multi-sensor affect recognition system for discerning interest and disinterest in children, which included facial expression, seating posture combined with the learner's activity on the computer. D'Mello et al. (2006) predicted affective states by using emote-aloud protocols when learners interact with AutoTutor. In their study they found that AutoTutor can detect affect states such as confusion, eureka and frustration by using dialogue features from the conversation between AutoTutor and the learner.

People can also be used. Graesser et al. (2006) involved trained judges, peers and the learners themselves for detecting and identifying some affect states in learning. When comparing, they found that peers were not as good as trained judges and the learners themselves in detecting emotion states.

Methods for discerning affect states in a learner thus ranges from observations performed by trained raters, think aloud protocols and questionnaires to more technically influenced methods such as measuring skin conductance, eye movements and seating posture. Other methods involve detection of facial expressions, use of dialogue features, voice features, etc.

Each method has benefits and drawbacks. A selection of methods is presented in table 1.

Table 1: A selection of methods with suggestions of some benefits and drawbacks.

Methods for detecting affect states		
Method	*Benefits from a*	*Drawbacks from*

	research perspective	a research perspective
Trained judges	High ecological validity	Time consuming, differences in judgement of same affect state between judges
Think aloud	Subjective reports on experiences	Interrupts ongoing activity; subjective interpretation
Emote aloud	Subjective reports on experienced affect states	Interrupts ongoing activity; subjective interpretation
Questionnaires	Structured	Performed after the affect state
Eye tracking	Objective	Does not define a specific emotion, but attention allocation can be inferred
Skin conductance	Objective	Intrusive, other factors may influence such as anxiety etc
Seating posture	Non-intrusive	Not yet a commercial system
Calculating algorithms	Consider multiple parameters	Artificial classification of face features/ words etc.
Voice	Quality of voice; intonation, speed	Time consuming transcriptions for analysis
Language	High ecological validity	Time consuming transcriptions for analysis

A current challenge is to refine measures and improve detection of learners' affect states. Since different methods come with different strengths and weaknesses, combinations of methods and measures are in many cases favourable. Also, different methods may highlight different features of an affective state or process.

Furthermore there are advantages with combinations of methods not only for research that aims to catch and identify affective states, but also in a practical situation when one needs to be as certain as possible of the identification of an affective state.

Current Directions

When we are dealing with advanced interactive learning environments there are further possibilities to combine methods and data from different sources in order to identify affective states. Besides *learner or participant data* (from voice, posture, eye movements, skin conductance, language, etc.) we can use *system data* or *learning environment data*. As an example, McQuaggan, Lee, and Lester (2007), with the goal of predicting upcoming frustration in students, combined physiological data with data from the system itself. They measured blood volume, pulse and galvanic skin responses from the learner (blood volume is used to calculate student heart rate and changes in heart rate) and combined this with data from the interactive learning environment, such as: the amount of time that has elapsed since the student had achieved a goal, which goals were being attempted by the student, which goals were achieved, the rate of goal achievement, and the effort expanded to achieve a goal. Using this combined data, the authors have built a frustration model predicting student frustration both relatively early and accurately. This kind of methodology, we believe, can be fruitful both for research purposes and in the end for practical purposes.

Another current research and development direction within the field of learning and affect is the one suggested by Craig et al. (2004), namely to aim at gaining more precise identifications of *when* emotions occur during the learning process: in a sense, to try to catch emotions on the fly. The study to be presented (for a more detailed presentation, see Lindström and Holmqvist in preparation) was set up in line with this suggestion.

Detecting Eureka

Catching Eureka on the Fly

An overall goal of the present study (for a full report see Lindström and Holmqvist in preparation) was to investigate possible emotional correlates of understanding by focusing on the eureka experience. More specifically, we wanted to investigate whether or not there are significant differences in emotional responses when a person understands something compared to when a person does not.

The study should be seen as a contribution to the unfolding of *the moments of understanding* and of characteristics of possible corresponding emotional responses.

Furthermore, the aim is to find methods by which the subjective feeling of the eureka experience can be complemented with other criteria, and together used as indicators of what precedes understanding and what happens when understanding occurs.

As an object of study, the eureka experience has two desirable features. First, the experience is very well marked in time. Second, it is relatively easy to determine whether an insight actually occurred or not. A potential problem is that eureka is rare. For example Craig et al. (2004) found

that eureka was relatively rare in relation to frustration or curiosity. However, in a later study it was a frequently reported affective state (D´Mello et al. 2006). The reasons for this contradiction are discussed at the end of the article.

Method

Forty participants, 15 females and 25 males, with an average age of 24 years, were asked to identify 26 stimuli that consisted of 9 incomplete pictures (a stimulus drawn in an unconventional manner, middle picture in figure 1), 12 ambiguous pictures (two interpretations of the stimulus are possible, where one is obvious and the other is subtle such as the scene of nature vs. a baby in figure 1.) and 5 mathematical-logical pictures (figure 1), and that were presented in random order on a computer screen. All pictures were pre-tested to have the potential of evoking an eureka experience when identified or solved. Also, during the study subjects reported statements such as "Aha!" or "Ok, now I got it!".

Figure 1: Examples of stimuli used which were predicted to yield emotional responses when identified.

Each picture was accompanied by a written sentence to provide subjects with a context to interpret the stimulus within.

Participants were instructed to click the mouse button as soon as they identified the stimuli. When they clicked the mouse, a blank screen was displayed, and they were instructed to verbally report their answer as soon as the blank screen appeared. In turn, this verbal report was used to validate what subjects had seen. (A verbal record was collected to serve as a control to compare subject's answer with the presented stimulus.)

This method for evaluating whether a subject understood or not is arguably non-intrusive to the experience itself.

To encourage participants' efforts in trying to identify as many pictures as possible, subjects were told that those participants identifying many pictures would receive an additional reward besides the value card they knew they would get after completing the experiment.

The measurements consisted of a combination of a behavioural measure – eye movements – with two physiological measures – skin conductance and pupil dilation. The use of eye tracking for discerning what and how people attend, was particularly suitable in the present experiment with a very clear visual component in the problem solving process. In all stimuli we defined *a critical area of interest*, i.e. what one ought to look at in order to interpret, solve and understand the problem (Grant and Spivey, 2003). For example, in line with Yarbus

(1967) it is likely to attend to the eye region, when there is a face or a person in a picture. *Misleading areas of interest* are areas that are not helpful for identifying or understanding the stimulus or something distracting from the task.

It has been showed that increased attention allocation on critical areas is important for different kinds of problem–solving processes (Grant and Spivey (2003); Knoblich, Ohlsson, and Raney (2001)).

Using a combination of different kinds of measurements, as related above, is in line with current approaches in affect detection. Other examples of combinations of measurements are psycho-physiological or behavioural measures of affect in combination with questionnaires; (Picard and Daily 2005); judges of different skills and conversational features (Graesser et al. 2006); affect sensing and task performance (Burleson and Picard 2004); multi-sensors including facial expressions, posture and user activity on the computer (Kapoor and Picard 2005); dialogue features from emote aloud protocols (D´Mello et al. 2006); facial expressions, a body posture measurement system and conversational features from AutoTutor's log file (D´Mello et al. 2005).

We chose to measure skin conductance and pupil dilation because they are commonly used as measure of emotional reactions and we predicted that an eureka experience would be readily detected in these measures. Attention allocation was chosen as a measure because eye movements show how online processing of a stimulus is proceeding. This gave, furthermore, a desired combination of a behavioural measure with physiological measures known to relate to subjective experiences.

Predictions

In the study (Lindström and Holmqvist, in preparation) it was predicted that there would be significant differences in responses between subjects identifying a stimulus compared to subjects not identifying the stimulus, in that the former would (i) have a larger pupil dilation (ii) have a higher skin conductance, (iii) look more on critical areas of interest in the stimulus picture and (iv) look less on non-critical areas of interest.

When formulating the predictions, we also selected a time frame to use for the analysis and comparisons of the data.

The identification or understanding of the stimulus plus clicking the button takes place within a limited time frame, and our decision was to analyze the 20 seconds leading up to the button click. We considered this time frame to be long enough to capture the relevant behaviour and responses, since insight is a sudden experience (Metcalfe 1986). The button click would also provide a time marker to allow us to backtrack the behaviour and the responses predicted to show the unfolding of understanding.

Results

For the analysis 4 of the 40 subjects were excluded due to technical problems. On the average, subjects identified 7,81 out of the 26 stimuli; the identification frequency was highest for the ambiguous stimuli, middle for the mathematical/logical stimuli and lowest for the incomplete stimuli.

Analyzing the data from the 20 seconds preceeding the button click, we found significant differences in support of our hypotheses ($p < .05$) (table 2) for pupil dilation, skin conductance levels as well as for attention allocation on critical and misleading areas of interest, when comparing subjects who identified the stimulus compared to subjects who did not identify the stimulus.

Table 2: Results

Measure	Effect size (d)
Pupil dilation	0,17
Skin conductance level	0,08
Critical area of interest	0,41
Misleading area of interest	0,24

Subjects identifying a stimulus had significantly larger pupil dilation – on average 4,7% larger – than subjects who did not identify the stimulus.

Skin conductance level was higher for subjects who identified a stimulus compared to subjects who did not identify a stimulus.

Subjects who identified a stimulus attended the critical areas in question significantly more, whereas reversely, subjects who did not identify a stimulus on average looked significantly more at the misleading areas of interest.

The increased level in responses for subjects identifying stimuli may reflect a *predisposition* for understanding to occur. This would be in line with previous research on interest where it has been shown that students who are interested in a topic pay more attention, persist longer and gain a deeper understanding (Hidi 1990).

The continuously larger pupil dilation of subjects while identifying the stimulus suggests a difference in level of cognitive processing, as supported by our data and data from studies where pupil size increased as a result of increased difficulty (Kahneman and Beatty 1966; Hess and Polt 1964) and more complex cognitive processing (Beatty and Wagoner 1978).

In summing up, our results were in line with the predictions. They were clearly so for the eye movement data (with a distinct peak occurring around 6 seconds before reporting identification of the stimulus by clicking the button.) Yet, the overall pattern of the physiological measures was *not as dramatic* as we expected from "a *sudden* aha-experience".

Interpretation

A possible reason for not identifying a *dramatic* change is that the units of analysis, data per second, was too wide. However, pupil dilation and skin conductance are rather slow reacting systems if compared with recording brain activity such as ERP (event related potentials), and our analysis frame is within the comparable range with similar studies such as Grant and Spivey (2003) and Knoblich, Raney and Ohlsson (2001) who used 30 and 15 seconds intervals respectively.

Another possibility is that *it is* precisely *the increased level* in the physiological data that is essential in the process of identifying stimulus and not a dramatic pattern as we had foreseen.

Such a line of reasoning can also be related to the conflicting results between, on one hand, a study by Craig et al. (2004) and, on the other hand, a study by D'Mello et al. (2006). These studies report very different frequencies of eureka. In the first study, eureka was discerned by trained judges using a structured protocol filled in every 5 minutes to determine the present affective state of the learner during a 30 second interval. Here eureka was rarely reported. In contrast, in the D´Mello et al. (2006) study exploiting an emote-aloud protocol, many eureka experiences were reported. The authors suggest that subjects in the second study might be using the category label "eureka" as a functional equivalence to happiness or delight rather than as indicating 'a true eureka experience' (D´Mello, et al. 2006). But it is also the case that the methods used in the two studies are quite different, which might account for the differences in reported eureka experiences. In the Craig et al. (2006) study a limited time frame for judging emotional states was set, and an *observer* should report on experiences of someone else. With emote-aloud protocols as used in the D'Mello et al. (2006) study, an individual reports his or her *own* affective state *whenever* experienced.

As to the concept of 'true eureka experience', it is indeed possible that subjects in the D'Mello et al. 2006 study used the labels incoherently or not in the intended way, but we also believe that there is a wider *span* of the eureka experience than suggested by the concept 'true eureka experience'. An eureka experience in the sense of *Newtonian insight* may be prototypical in our conceptions of eureka. Such intense moments or states are probably quite rare. However, *less intense*, but still similar experiences seem to occur much more frequently in learning contexts when something falls into place and is finally or suddenly clear or solved.

Indeed, the tasks and situation in the actual study are more likely to produce such affective states than Newtonian eurekas.

And in line with this, the discrepancy in the results from the two studies described above may reflect how different methods capture different features of the experience, and how we probably would need more fine grained criteria for delimiting the phenomenon of eureka in its entire manifestations. But for such a project, the quest for more precise measures for detecting the experience is necessary.

Discussion and Future Work

Catching Eureka - Possibilities for the Future

On one hand, the combination of methods used in the presented study worked functionally well. On the other hand, eureka, at least in some of its manifestations, appears to be a relatively distinct affective state. Therefore it is of interest to continue a quest to further pinpoint what *are* the defining features of eureka. In order to do so, i.e. to attempt to fully capture this somewhat illusive experience, refinements of stimuli as well as of the learning contexts and the measuring methods will be required. Not least, we may need more fine-grained criteria in order to delimit the feature span of the experience.

Our aim is to continue exploring the eureka phenomenon in its entire manifestation by focusing on more typical learning contexts and tasks and to continue to work on refinements of the measuring methods.

ITS and Adaptation

It can easily be seen how the recognition of *confusion* or *frustration* can be useful in guiding pedagogical interventions. But how could identification of the relatively brief and clearly positive emotion of eureka be useful? Specifically, how could an adaptive intelligent tutoring system make use of such information?

At first sight, this may seem questionable. There is no obvious reason for *intervening* when recognizing eureka in a student. Recognizing confusion or frustration, indeed, functions (or ought to function) as a signal for a teacher – and for an intelligent tutoring system – to intervene in some suitable way. But in the case of a student experiencing the relatively brief and clearly positive emotion of eureka, what is the need for intervention or adaptation on the part of a tutor?

There are two things that should not be forgotten here. Firstly, pedagogical adaptivity in a powerful ITS (as well as in a human teacher) involves *learnability* on the part of the teaching system. In order to *improve* ones teaching, an ability to analyze and identify the kinds of learning situations or activities that *precede* students' eureka experiences can be highly relevant. Not the least so with respect to material that is well known to be hard for students really to understand. Here an identification of eureka experiences can help to pinpoint the teaching strategies, the set of examples and tasks etc. that seem to work well in guiding students towards an understanding of such material that is often difficult to get a grip of.

Secondly, the communicative function of eureka, as touched upon earlier in the text, should be emphasized. An experience of understanding in a student both signals that the strategy used by the teacher is successful, but it also informs the teacher that the student is now ready to proceed in the learning process. Such communicative function might seem obvious or trivial, but can indeed be helpful in future interactive learning environments – perhaps targeted at difficult topics and problems – to attain the delicate balance between providing well-timed support and well-timed challenges to a student.

Finally, we suggest more elaborative measures and methods for trying to capture eureka in its entire manifestation, because we think that this often positive affect state, can be valuable as signals for an intelligent tutoring system as well as for the student to learn from. The learning system can use them to adopt and refine its teaching strategies. The student can gain more fine-grained feedback and access to useful strategies that were exploited in order to solve the problem or task at hand. For example, a student dealing with an especially difficult task may see a diagnostic review of the actions taken to solve the problem and be supported in developing meta-strategies for solving forthcoming problems.

Acknowledgments

References

Beatty, J., and Wagoner, B. (1978). Pupilometric signs of brain activation vary with level of cognitive processing. *Science* 199(4334): 1216-1218.

Burleson, W., and Picard, R. W. (2004). Affective agents: Sustaining motivation to learn through failure and a state of stuck. Retrieved 15/09/06 from http://web.media.mit.edu

Craig, S. D., Graesser, A. C., Sullins, J., and Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media,* 29(3): 241-250.

D'Mello, S., Craig, S.D., Gholson, B., Franklin, S., Picard, R. W., and Graesser, A. C. (2005). Integrating affect sensors in an intelligent tutoring system. Workshop at 2005 International Conference on Intelligent User Interfaces, 7-13.

D´Mello, S., Craig, S. D., Sullivan, J., and Graesser, A. C. (2006). Predicting affective states through an emote-aloud procedure from AutoTutors mixed-initiative dialogue. *International Journal of Artificial Intelligence in Education* 16: 3-28.

Feltovich, P., Coulson, R., and Spiro, R. (2001). Learners' (Mis)Understanding of important and difficult concepts. In Forbus. K. and Feltovich, P. eds. Smart Machines in Education, 349-375. AAAI Press / The MIT Press.

T. Okamoto, R. Hartley, Kinshuk and J. P. Klus .eds. *Proceedings IEEE International conference on Advanced Learning Technology: Issues, Achievements and Challenges*, 43-48.

Goleman, D. (1995). *Emotional intelligence.* New York: Bantam books.

Graesser, A., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., and Cholson, B. (2006). Detection of emotions during learning with AutoTutor, 285-290. In

proceedings of the 28th Annual Conference of the Cognitive Science Society, CogSci 2006. Vancouver, Canada: Cognitive Science Society 28th annual meeting.

Grant, E., and Spivey, M. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science* 12(5): 462-466.

Hess, E. H, and Polt, J. M.(1964). Pupil size as related to interest value of visual stimuli. *Science* 132(3423): 349-350.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research* 60: 549–57.

Immordino-Yang, M. H., and Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education, *Journal of Mind, Brain and Education* 1(1): 3-10.

Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. *Science* 159(3756): 1583-1585.

Kapoor, A., and Picard, R. W. (2005). Multimodal affect recognition in learning environments. In Proceeding of the 13th annual ACM international conference on Multimedia, 677-682. Singapore: ACM.

Köhler, W. (1959). *The task of gestal psychology.* Prineton NJ: Princeton University.

Kort, B., Reilly, R., and Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy- building a learning companion. In T. Okamoto, R. Hartley, Kinshuk and J. P. Klus (Eds.) Proceedings IEEE International conference on Advanced Learning Technology: Issues, Achievements and Challenges, 43-48. Madison, Wisconsin: IEEE Computer Society.

Knoblich, G, Ohlsson, S, and Raney, G. A. (2001). An eye movement study of insight problem solving. *Memory and Cognition,* 29(7): 1000-1007.

Lindström, P, Holmqvist, K. (in preparation).

Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation and cognition. *Educational Psychology Review* 18: 307-314.

McQuggan, S. Lee, S. and Lester, J. (2007) Early Prediction of Student Frustration. In: Paiva, A., Prada. R., and Picard, R. eds. *Proceedings of the Second International Conference on Affective Computing and Intelligent Interaction (ACII 2007), 698-709.* Springer-Verlag.

Meyer, D., and Turner, J. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychology Review* 18: 377-390.

Picard, R. W., and Bryant Daily, S. (2005). Evaluating affective interactions: Alternatives to asking what users feel. In Proceedings of the CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches, Portland.

Pons, F., Hancock, D. R., Lafortune, L., and Doudain, P-A. eds. (2005). *Emotions in learning.* Aalborg: Aalborg Universitetsforlag.

Schutz, P. A., and Pekrun, R. eds. (2007). *Emotions in education.* Educational Psychology Series: Academic Press: Burlington.

Schutz, P. A., Hong, J. Y., Cross, D. I., and Osbon, J. N. (2006). Reflections on investigating emotion in educational activity setting. *Educational Psychology Review* 18: 343-360.

Trout, J. D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science* 69: 212-233.

Wertheimer, M. (1959). *Productive thinking.* New York: Harper.

Yarbus, A. L. (1967). *Eye movements and vision.* New York: Plenum Press.