# Interventions during student multimodal learning activities: which, and why?

Beate Grawemeyer
London Knowledge Lab
Birkbeck College
University of London, UK
beate@dcs.bbk.ac.uk

Manolis Mavrikis
London Knowledge Lab
Institute of Education
University of London, UK
m.mavrikis@ioe.ac.uk

Sergio Gutierrez-Santos
London Knowledge Lab
Birkbeck College
University of London, UK
sergut@dcs.bbk.ac.uk

Alice Hansen
London Knowledge Lab
Institute of Education
University of London, UK
a.hansen@ioe.ac.uk

## ABSTRACT
Emotions play a significant role in students' learning behaviour. Positive emotions can enhance learning, whilst negative emotions can inhibit it. This paper describes a Wizard-of-Oz (WoZ) study which investigates the potential of Automatic Speech Recognition (ASR) together with an emotion detector able to classify emotions from speech to support young children in their exploration and reflection whilst working with interactive learning environments. We describe a unique ecologically valid WoZ study in a classroom. During the study the wizards provided support using a script, and followed an iterative methodology which limited their capacity to communicate, in order to simulate the real system we are developing. Our results indicate that there is an effect of emotions on the acceptance of feedback. Additionally, certain types of feedback are more effective than others for particular emotions.

## Keywords
Affect, emotions, intelligent support

## 1. INTRODUCTION
Our aim is to build a learning platform for elementary education which integrates speech recognition for children in order to enable natural communication. This paper reports from on a Wizard-of-Oz study which explores the effect of emotions deduced from speech on different feedback types.

The importance of language as both a psychological and cultural tool that mediates learning has long been recognised; from as early as Vygotsky to modern linguists such as Pinker. From a Human Computer Interaction (HCI) perspective, speech recognition technology has the potential to enable more intuitive interaction with a system, particularly for young learners who reportedly talk aloud while engaged in problem solving (e.g. [11]).

Finally, speech provides an additional cue for drawing inferences on students' emotions and attitude towards the learning situation while they are solving tasks. By paying attention to tone and pitch of speech in conjunction with other auditory signs like sighs, gasps etc., we can provide learners with even more individualized help, by detecting emotions and providing support specifically tailored to the emotional state.

As described in [15] emotions interact with and influence the learning process. While positive emotions such as awe, satisfaction or curiosity contribute towards constructive learning, negative ones including frustration or disillusionment at realising misconceptions can lead to challenges in learning. The learning process includes a range and combination of positive and negative emotions. For example, a student is motivated and expresses curiosity to explore a particular learning goal, however s/he might have some misconceptions and needs to reconsider her/his knowledge. This can evoke frustration and/or disappointment. However, this negative emotion can turn into curiosity again, if the student gets a new idea on how to solve the learning task.

[9] categorised emotions based on facial expressions. These included, joy, anger, surprise, fear, and disgust/contempt. However, these emotions are not specific to learning. [22] classified achievement emotions that arise in a learning situation. Achievement emotions are emotions that are linked to learning, instruction, and achievement. Emotions are classified into prospective, retrospective and activity emotions. They can be positive or negative. For example, a prospective positive emotion is hope for success, while a negative emotion is anxiety about failure. Retrospective emotions are for example, the positive emotion pride or the negative emotion shame, which the student experienced after receiving feedback of an achievement. Activity emotions arise during learning, such as positive emotions like enjoyment, or negative emotions like anger, frustration, or boredom.

We focus on on a subset of emotions identified by Pekrun and Ekman: enjoyment, surprise, frustration, and boredom. We also add confusion as an emotion, which is placed between enjoyment and frustration.

As described in [29] students can become overwhelmed (very confused or frustrated) during learning, which may increase cognitive load for low-ability or novice students. However, appropriate feedback can help to overcome such problems.

Effective support or feedback needs to answer four main questions: *when*, *what*, *how*, and *why*: (i) *when* to provide the support during learning. (ii) It needs to be decided *what* the support should contain; (iii) *how* it should be presented; and (iv) *why* the feedback needs to be provided.

In this paper we focus on *what* (ii) and *why* (iv) support or feedback should be provided based on the student's emotion. In the area of intelligent tutoring systems or learning environments, the only research we are aware of specifically targeting the question of responding to student affect is [29] and [2]. [29] describes how an embodied pedagogical agent is able to provide different types of interventions, such as praising or mirroring the student's emotional state. [2] looks at the effect of cognitive-affective states on student's learning behaviour. In contrast, in this paper, we investigate the impact of emotions on the effectiveness of different feedback types.

The structure of the paper is as follows: The next section overviews related work on detecting and adapting to emotions in the educational domain. This is followed by a description of the Wizard-of-Oz study, which investigated the effect of emotions on different feedback types. We then discuss the different feedback types. After this, we provide results and discuss the results of the study in respect to adaptive support based on student's emotion. We conclude by outlining directions for future research.

## 2. BACKGROUND

Different computational approaches have been taken into account in order to detect emotions. These include for example, speech-based approaches (e.g. [6, 27]), using information from facial expressions (e.g. [14]), keystrokes or mouse movements [10], physiological sensors (e.g. [16, 28, 21]), or a combination of these [7].

In the area of education [5] developed a model of emotions (Dynamic Bayesian network) based on students' bodily expressions for an educational game. The system uses six emotional states: joy, distress, pride, shame, admiration and reproach. A pedagogical agent provides support according to the emotional state of the students and the user's personal goal, such as wanting help, having fun, learning maths, or succeeding by oneself.user's personal goal, such as wanting help, having fun, learning maths, or succeeding by oneself.

Another example, is [25] who also used Bayesian Networks to classify students' emotions. Here biophysical signals, such as heart rate, skin conductance, blood pressure, and EEG brainwaves, for the classification of emotions. These include: interest, engagement, confusion, frustration, boredom, hopefulness, satisfaction, and disappointment.

As described earlier, [29] developed an affective pedagogical agent which is able to mirror students' emotional state, or acknowledge a student's emotion if it is negative. They use hardware sensors and facial movements to detect students emotion. The system discriminates between seven emotions: high/low pleasure, frustration, novelty, boredom, anxiety, and confidence. Different machine learning techniques were applied for the classification, including Bayesian Networks and Hidden Markov models.

[17] developed a physics text-based tutoring system called ITSPOKE. It uses spoken dialogue to classify emotions. Acoustic-prosodic and lexical features are used to predict student emotion. They apply boosted decision trees for their classification. Three emotion types are detected: negative, neutral and positive emotions.

Another example is the AutoTutor tutoring system [7], which holds conversations with students in computer literacy and physics courses. The system classifies emotions based on natural language interaction, facial expressions, and gross body movements. The focus is on three emotions, namely frustration, confusion, and boredom. The classification is used to respond to students via a conversation.

Most of the related work in the educational domain focusses on detecting emotions based on different input stimuli, ranging from spoken dialogue to physiological sensors. However, little research has been done on how those detected emotions can be used in a tutoring system to enhance the learning experience. One exception is [29] who describes how an affective pedagogical agent can support students in particular emotional states. Additionally, [2] investigated the impact of student's cognitive-affective states on how they interacted with the learning environment. They found that certain types of emotions, such as boredom, were associated with poor learning and gaming the system. In contrast, we investigate the implications of emotions for different feedback types. We conducted a WoZ study where different kinds of feedback were provided to students in different emotional states. The next section describes the WoZ study in more detail.

### 2.1 Aims

One of our research aims is to provide adaptive feedback to students during a learning activity which enhances the learning experience by taking into account students' emotion. We were specifically interested in the following questions, which we aimed to address in the WoZ studies:

- Is there an effect of different emotion types upon reaction towards feedback?
- Which interventions were most successful given a particular emotional state?

In order to address these questions we ran an ecologically valid WoZ study which investigated the effect of emotions on different feedback types at different stages of the task.

### 2.2 Methodology

The studies reported on this paper are part of a methodology referred to as Iterative Communication Capacity Tapering (ICCT). This can be used to inform the design of intelligent support for helping students in interactive educational applications [18]. During the first phase, the facilitator gradually moves from a situation in which the interaction with the student is close, fast, and natural (i.e. face-to-face free interaction) towards a situation in which the interaction is mediated by computer technologies (e.g. voice-over-ip or similar for voice interaction, instant messaging or similar for textual interaction) and regularised by means of a script. In the second phase, the script is crystallized into a series of intelligent components that produce feedback in the same way that the human facilitator formaly did. The gradual reduction of communication capacity and the iterative nature of the process maximise the probability of the computer-based support being as useful as the facilitator's help. In this paper, we are already starting the second phase, i.e. gradually replacing humans by a computer-based system. Experts ('wizards') are not physically near enough to the students to observe them directly, and therefore must observe them by indirect mediated means: the students' voice was heard by using microphones and headsets and their screen was observed by a mirror screen. The wizards did not have direct access to the students' screens (so e.g. could not point to anything on the screen to make a point), could not see the students' faces (for facial cues), and could not communicate to students by using body language, only by means of the facilities provided by the wizard-of-oz tools that resemble those of the final system.

## 2.3 Participants and Procedure

After returning informed consent forms signed by their parents 60 Year-5 (9 to 10-year old) students took part in a series of sessions with the learning platform configured for learning fractions through structured tasks from the intelligent tutoring system, together with more open-ended tasks offered by the exploratory learning environment. The sessions were designed to first familiarise all students with the environment, and then to allow them to undertake as many tasks as possible (in a study which has goals outside the scope of this paper). In parallel, we were running the WOZ study by asking two students in each session to work on different computers as described below. In total 12 students took part in the WOZ study but due to data errors we were able to analyse the interaction of only 10 students. At the end of the session the students who participated in the WOZ joined in a focus group discussing their experience with the learning platform. We were particularly interested in students' opinions about the different feedback types provided.

## 2.4 Classroom setup

The ecological validity of the study was achieved by following the setup depicted in Figure 1, 2 and Figure 3. The classroom where the studies took place is the normal computer lab of the school in which most of the computers are on tables facing the walls in a II-shape, and a few are on a central table. This is the place where the WOZ study took place, while, for ecological validity, the rest of the class was working on the other computers. The students were only told that the computers in the central isle were designed to test the next version of the system and were thus also responding to (rather than just recording as the rest of the computers) their speech. The central isle has two rows of computers, facing opposite directions, and isolated by a small separator for plugs etc. In the central isle the students worked on a console consisting on a keyboard, a mouse, and a screen. Usually, those components are connected to the computer behind the screen; for these studies, they were connected to a laptop on the wizards' side of the table. This allowed the wizard to observe what the students were doing. As the learning platform is a web-based system, and all the students' see is a web browser, the operating system and general look-and-feel of the experience was equivalent to the one that the rest of the students were using. When the wizards wanted to intervene, they used the learning platform's WOZ tools to send messages to the student's machine. These messages were both shown on screen and read aloud by the system to students, who could hear them on their headset.

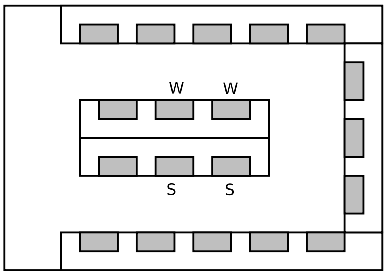


**Figure 1: The layout. The Wizard-of-Oz studies took place on the central isle while the rest of the students worked on a version of the system which only sequences tasks and provides minimal support.**

**Figure 2: The classroom. The children being wizarded in front with wizards at the back.**

## 2.5 The wizard's tools

In line with the ICCT methodology mentioned above, the wizards restricted their 'freedom' in addressing the students by employing a pre-determined agreed script in which the expected interventions had been written. Figure 4 shows a high-level view of this script, the end-points of which require further decisions also agreed in advance in a protocol but not shown here for simplicity. In this study, we limited ourselves to written interventions that could be selected from an online document appropriate for being read aloud by the system. There were no other kinds of interventions (such as sounds, graphical symbols on screen etc.). The intervention had a set of associated conditions that would fire them thus resembling very closely the system under development.

## 2.6 Feedback types

As outlined in the script (figure 4) different types of feedback were presented to students at different stages of their learning task. The feedback provided was based on interaction via keyboard and mouse, as well as speech.

From an HCI perspective speech production and recognition can provide potentially more intuitive interaction. In

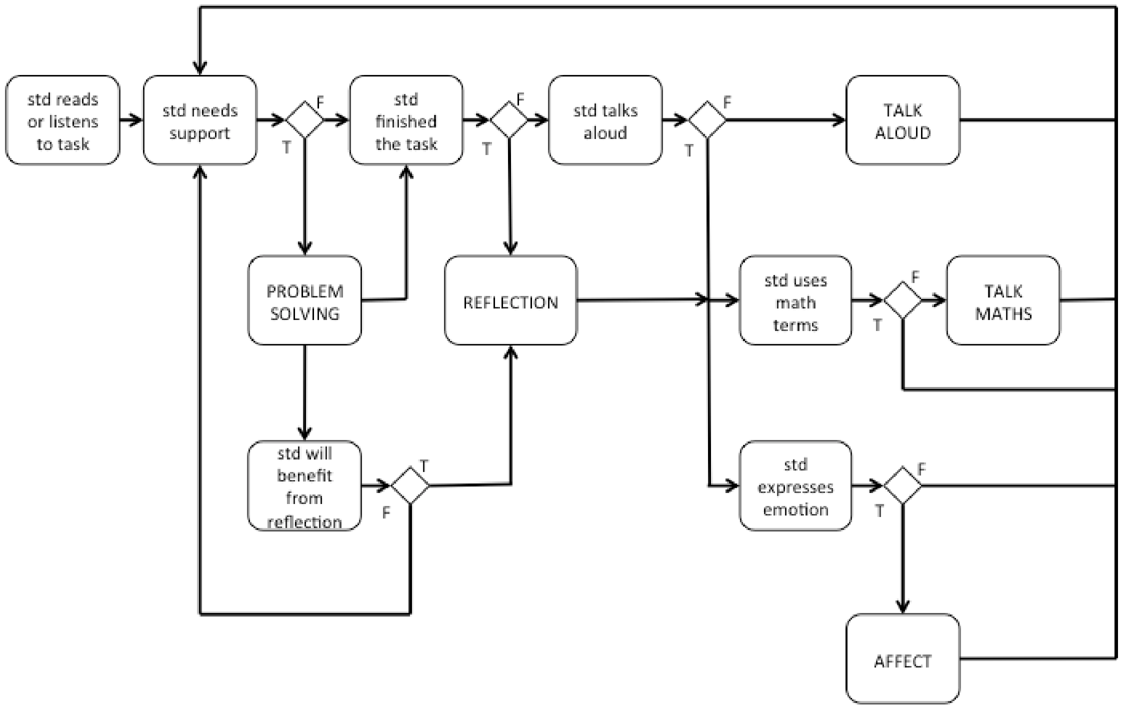


**Figure 4: Flowchart representing the wizard's script for support.**

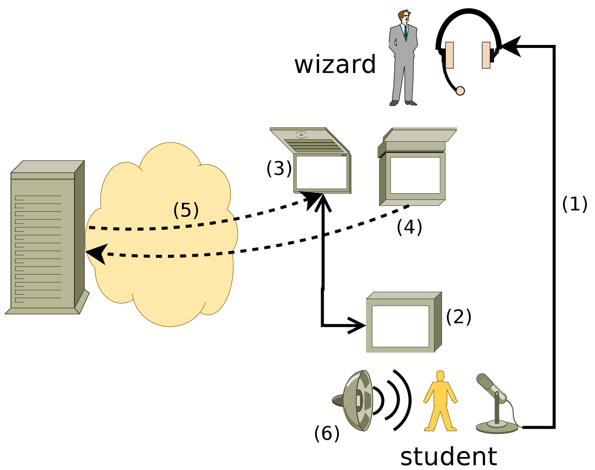


**Figure 3: Wizard-of-oz setup. Each student speaks on a headset (mic) which is connected to the wizard's headset (1). The student interacts with a console (i.e. keyboard, mouse, screen) connected to a laptop on the wizard's side (2,3) so that the latter can witness their interaction. The wizard can send messages (4) by using some ad-hoc wizard tools. These messages arrive at the student laptop (5) and are shown on the screen of the student's monitor and read aloud on the student's headset (6).**

particular, spoken language input can enable students to communicate verbally with an educational application and thus interact without using human interface devices such as a mouse or keyboard. The following different feedback types were provided:

- **PROBLEM SOLVING - task-dependent feedback**
  This feedback based mainly on the interaction with mouse and keyboard with the learning environment. Here the feedback involved providing support in solving a particular maths problem.
- **TALK MATHS - using particular domain specific maths vocabulary**
  The importance of students' verbal communication in mathematics in particular becomes apparent if we consider that learning mathematics is often like learning a foreign language. Focusing, for example, on learning mathematical vocabulary, [3] encouraged students to talk to a partner about a mathematical text to share confusions and difficulties, make connections, put text into their own words and generate hypotheses. This way, students were able to make their tentative thinking public and continually revise their interpretations.
- **AFFECT - affect boosts**
  As described in [29] affect boosts can help to enhance student's motivation in solving a particular learning task. Higher motivation also implies better performance.
- **TALK ALOUD - talking aloud**
  With respect to learning in particular, the hypothesis

that automatic speech recognition (ASR) can facilitate learning is based mostly on educational research that has shown benefits of verbalization for learning (e.g., [1, 3, 20]).

The possible verbalization effect could be enhanced with ASR since cognitive load theory [26] and cognitive theory of multimedia learning [19] predict that a more natural and efficient form of communication will also have positive learning gains.

The few existing research studies have found mixed results with respect to whether the input modality (speaking vs. typing) has a positive, negative or no effect on learning. In [8], for example, the authors investigated whether student typing or speaking leads to higher computer literacy with the use of AutoTutor. They reported mixed results that highlight individual differences among students and a relationship to personal preferences and motivation.

- **REFLECTION - reflecting on task performance and learning**
For further consideration is the research about self-explanation; an efficient learning strategy where students are prompted to verbalize their thoughts and explanations about the target domain to make knowledge personally meaningful. Previous research [13] found that the amount of self-explanation that students generated in a computer environment was suppressed by having learners type rather than speaking and the studies. Moreover, some students are natural self-explainers while others can be trained to self-explain [24]. Even when self-explanation is explicitly elicited, it can be beneficial [4] but requires going beyond asking students to talk aloud by using specific reflection prompts [24].

Self-explanation can be viewed as a tool to address students' own misunderstandings [4] and as a 'window' into students' thinking. While it may be early days for accurate speech recognition to be able to highlight specific errors and misconceptions, undertaking carefully-designed tasks can help identify systematic errors that students make. For example, [12] explores how naming and misnaming involves logic and rules that often aid or hinder students' mathematical learning and relate to misconceptions.

A lack of mathematical terminology can also be noticed and prompts made to students to use appropriate language as they self-explain.

Table 1 shows examples of the different feedback types. We were interested to explore how emotions impact on the effectiveness of those different feedback types.

# 3. RESULTS

From the WoZ study we recorded students' screen display and their voices. From this data, we annotated emotions and whether students reacted to feedback.

For the annotation of the emotions and students reactions towards the feedback, we used a similar strategy as described in [23] where dialog between a teacher and a student was

| Feedback type | Example |
|---|---|
| AFFECT | It may be hard, but keep trying. If you find this easy, check your work and change the task. |
| TALK ALOUD | Remember to talk aloud, what are you thinking? What is the task asking you to do? |
| TALK MATHS | Can you explain that again using the terms denominator, numerator? |
| PROBLEM SOLVING | You can't add fractions with different denominators. |
| REFLECTION | What did you learn from this task? What do you notice about the two fractions? |

**Table 1: Examples of feedback types**

annotated according to different feedback types. Also,[2] describe how they coded different cognitive-affective states based on observations of students interacting with a learning environment. Similarly, we annotated student's emotion and if they reacted for each type of feedback provided. Another researcher went through the categories and any discrepancies were discussed and resolved before any analysis took place.

In total 170 messages were sent to 10 students. The raw video data was analysed by a researcher who categorised the emotions and feedback messages. Table 1 shows the different types of messages send to students and the emotions that occurred while the feedback was given. It can be seen that most frequent messages were reminders to talk aloud (66). This was followed by problem-solving feedback (55), and feedback according to students emotions (31). The least frequent messages relates to reflection (13) and using maths terminology (5).

It is not surprising that most of the problem solving feedback was provided when students were confused (35 out of 55). Most of the affect boosts were provided when students enjoyed the activity (15 out of 31), closely followed by students' being confused (11 out of 31). Most of the reflection prompts were given when students enjoyed the activity (10 out of 13). Talk aloud reminders were mainly given when students were confused (30 out of 66). Talk maths prompts were mainly given when students enjoyed the task (3 out of 5) or when they were confused (2 out of 5).

The emotions that were detected by students when feedback was provided and whether students reacted can be seen in figure 5.

Students reacted to all of the feedback when they were bored or surprised (100%).This was followed by reactions to feedback when students were confused (83%) or enjoyed the activity (81%). Students responded the least if they were frustrated (69%).

Looking in more detail at emotions and whether students reacted to the different feedback types, figures 6, 7, and 8 show the percentage of student's reaction towards feedback type for enjoyment, confusion, and frustration.

| | emotion | | | | | |
|---|---|---|---|---|---|---|
| Feedback type | enjoyment | boredom | confusion | frustration | surprise | total |
| PROBLEM SOLVING | 8 | 3 | 35 | 8 | 1 | 55 |
| TALK MATHS | 3 | 0 | 2 | 0 | 0 | 5 |
| AFFECT | 15 | 2 | 11 | 3 | 0 | 31 |
| TALK ALOUD | 21 | 1 | 40 | 4 | 0 | 66 |
| REFLECTION | 10 | 1 | 1 | 1 | 0 | 13 |
| Total | 57 | 7 | 89 | 16 | 1 | 170 |

**Table 2: Feedback types, including emotion that occurred while the feedback was provided.**

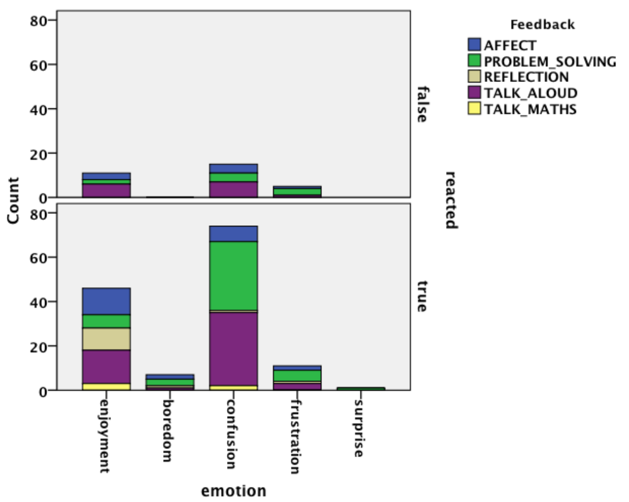


**Figure 5: Student's reaction according to feedback types and emotion.**

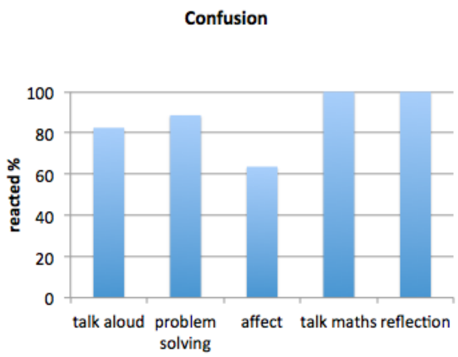


**Figure 7: Students' reaction according to feedback types if they were confused.**

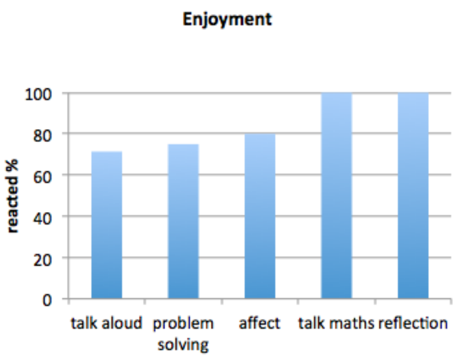


**Figure 6: Students' reaction according to feedback types if they enjoyed the activity.**

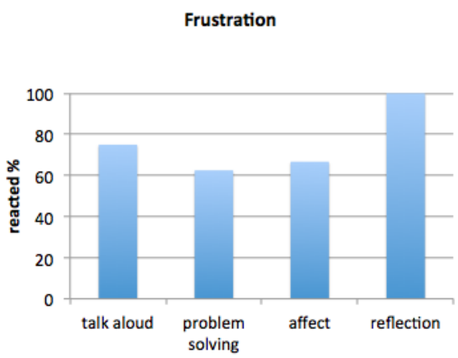


**Figure 8: Students' reaction according to feedback types if they were frustrated.**

It is interesting to see that while students enjoyed their activity, they responded very well to talk maths (100%) or to reflect on what they have done (100%). The least reaction was given if students were prompted to talk aloud (71%).

If students were confused they responded well again on talk maths (100%) or reflection prompts (100%), followed by problem solving feedback (89%). Surprisingly, least reactions were given when affect boosts were provided (64%).

If students were frustrated most reactions were given for reflection (100%) and prompts to talk aloud (75%). Least responses were given if problem solving feedback was provided (63%).

# 4. DISCUSSION

The key findings with respect to impact of emotions on the effect of feedback types are listed below in relation to our research aims.

## 4.1 Is there an effect of different emotion types upon reaction towards feedback?

The results show that for certain types of emotions, such as boredom, any type of feedback is reacted to. This indicates that students may welcome a distraction from their learning and react to feedback if they are bored. As boredom indicates a reduction in learning [2], the feedback provided to students when they are bored should aim to motivate and support the student to continue with the learning task.

Also in most of the cases students reacted to the feedback when they were confused. This implies that students welcome feedback that will help them to get out of their confused state. In designing feedback for learning environments students should be provided with feedback that enables them to overcome their confusion, such as task-dependent problem solving feedback, or feedback to reflect on their learning, which might help to identify and overcome misconceptions.

Additionally, students mainly reacted to feedback when they were enjoying their activity. This is an interesting finding, as in theory this seems to interrupt their learning flow. Here, it seems students' motivation is high and they did not mind being interrupted. Students particularly reacted positively on feedback to reflect.

In contrast, when students were frustrated, they reacted to feedback in only 69% of the cases. This indicates that frustration can reduce motivation and may also increase cognitive load. Here feedback that might help to decrease the frustration, such as reflecting on the difficulty of the learning task might help to motivate the student.

## 4.2 Which interventions were most successful given a particular emotional state?

The results indicate that for different emotional states, different feedback types are more effective than others.

It is interesting to see that although students enjoyed their activity and reacted to feedback in 81% of the cases, response to talk aloud was only 71%. This was similar when students were frustrated (75%). In contrast when students were confused in 83% of the cases students followed the recommendation to talk aloud. It looks like as if talking aloud might help to identify the problem and might resolve the confusion.

The highest reaction was given to problem solving feedback if students were confused (89%). This is not surprising as students were happy to receive help to perform the task. However, in only 75% of the cases was problem solving feedback reacted to while students enjoyed the activity. This might be because they were interrupted in their learning flow and they needed to switch to a new strategy of answering the learning task based on the problem solving feedback. The number drops even more when students were frustrated (63%). As discussed above, students' motivation might be low when frustrated and also there might be increased cognitive load. Providing problem solving feedback when students are frustrated does not seem to be a very effective strategy.

Providing affect boosts was most effective when students enjoyed their activity (80%). In contrast, students only reacted to affect boosts in 67% of the cases when they were frustrated or 64% when they were confused. From the focus group with the students it emerged that although some students did not react to the emotional boosts when they were confused or frustrated, they liked the encouragement, and that it helped with their motivation to continue to work on the particular learning task.

Providing prompts to talk maths and reflection were very effective across the emotion types. Despite the fact that 5 talk maths prompts and 13 reflection prompt were provided, students seemed to respond to them very well whether confused or frustrated. This implies that reflecting on one's own strategy of solving a task is motivating even if confused or frustrated. We noticed that it may also helped students to identify misconceptions or lead to new ideas on how to solve the learning task.

# 5. CONCLUSION AND FUTURE WORK

We explored the impact of students' emotional state upon different feedback types. The results indicate that certain types of feedback are more effective then others according to the emotional state of the student. While for some emotional states, such as boredom, a variety of feedback types worked well, for other emotional states, like frustration, only a few types of feedback seem to be effective.

We are now developing and integrating the automatic speech and emotion recognition in our learning platform. Additionally the adaptive support that is able to provide the different feedback types for particular emotional states is under development. At the next stage of our research we are interested to explore how the presentation of the feedback (e.g. high or low intrusive) affects students being interrupted in performing the task and if the presentation has an effect on reaction towards the feedback.

# 6. ACKNOWLEDGMENTS

This research has been co-funded by the EU in FP7 in the iTalk2Learn project (318051). Thanks to all our iTalk2Learn colleagues for their support and ideas.

# 7. REFERENCES


[1] M. Askeland. Sound-based strategy training in multiplication. *European Journal of Special Needs Education*, 27(2):201–217, 2012.

[2] R. S. J. d. Baker, S. K. D'Mello, M. T. Rodrigo, and A. C. Graesser. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *Int. J. Hum.-Comput. Stud.*, 68(4):223–241, apr 2010.

[3] R. Borasi, M. Siegel, J. Fonzi, and C. Smith. Using transactional reading strategies to support sense-making and discussion in mathematics classrooms: An exploratory study. *Journal for Research in Mathematics Education*, 29:275–305, 1998.

[4] M. Chi. Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser, editor, *Advances in instructional psychology*, pages 161–238. Mahwah, NJ: Lawrence Erbaum Associates, 2000.

[5] C. Conati and H. MacLaren. Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, 2009.

[6] R. Cowie, E. Douglas-Cowie, B. Apolloni, J. Taylor, A. Romano, and W. Fellenz. What a neural net needs to know about emotion words. *Computational Intelligence and Applications*, pages 109–114, 1999.

[7] S. D'Mello, S. Craig, B. Gholson, S. Franklin, R. Picard, and A. Graesser. Integrating affect sensors in an intelligent tutoring system. In *Affective Interactions: The Computer in the Affective Loop Workshop at 2005 International Conference on Intelligent User Interfaces*, pages 7–13, 2005.

[8] S. K. D'Mello, N. Dowell, and A. Graesser. Does it really matter whether student's contributions are spoken versus typed in an intelligent tutoring system with natural language? *Journal of Experimental Psychology; Applied 2011*, 17(1):1–17, 2011.

[9] P. Ekman. An argument for basic emotions. *Cognition & Emotion*, 6(3-4):169–200, 1992.

[10] C. Epp, M. Lippold, and R. Mandryk. Identifying emotional states using keystroke dynamics. In *2011 Annual Conference on Human Factors in Computing Systems*, pages 715–724, 2011.

[11] J. H. Flavell, F. L. Green, E. R. Flavell, and J. B. Grossman. The Development of Children's Knowledge of Inner Speech. *Child Development*, 68(1):39–47, 1997.

[12] H. A. Furani. Misconceiving or misnaming?: Some implications of toddlers' symbolizing for mathematics education. *Philosophy of Mathematics Education Journal*, 17, 2003.

[13] R. G. M. Hausmann and M. T. H. Chi. Can a computer interface support self-explaining? *Cognitive Technology*, 7(1):4–14, 2002.

[14] R. Kaliouby and P. Robinson. Real-time inference of complex mental states from facial expressions and head gestures. In *IEEE International Conference on Computer Vision and Pattern Recognition Workshop*, 2004.

[15] B. Kort, R. Reilly, and R. Picard. An affective model of the interplay between emotions and learning. In *IEEE International Conference on Advanced Learning Technologies*, number 43-46, 2001.

[16] P. Lang, M. Greenwald, M. Bradley, and A. Hamm. Look at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30:261–273, 1993.

[17] D. Litman and K. Forbes-Riley. Predicting student emotions in computer-human tutoring dialogues. In *42nd Annual Meeting on Association for Computational Linguistics (ACL '04), Association for Computational Linguistics*, 2004.

[18] M. Mavrikis and S. Gutierrez-Santos. Not all wizards are from Oz: Iterative design of intelligent learning environments by communication capacity tapering. *Computers & Education*, 54(3):641–651, Apr. 2010.

[19] R. E. Mayer and R. Moreno. Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1):43–52, 2003.

[20] N. Mercer and C. Sams. Tecahing children how to use language to solve maths problems. *Language and Education*, 20(6):507–528, 2007.

[21] F. Nasoz, K. Alvarez, C. Lisetti, and N. Finkelstein. Emotion recognition from physiological signals for presence technologies. *International Journal of Cognition, Technology and Work, Special Issue on Presence*, 6(1):4–14, 2003.

[22] R. Pekrun. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *J. Edu. Psych. Rev.*, pages 315–341, 2006.

[23] K. Porayska-Pomsta, M. Mavrikis, and H. Pain. Diagnosing and acting on student affect: the tutor's perspective. *User Modeling and User-Adapted Interaction*, 18(1):125–173, Feb. 2008.

[24] M. Roy and M. T. H. Chi. The self-explanation principle in multimedia learning. In R. E. Mayer, editor, *Cambridge handbook of multimedia learning*, pages 271–286. New York: Cambridge University Press, 2005.

[25] L. Shen, M. Wang, and R. Shen. ffective e-learning: Using "emotional" data to improve learning in pervasive learning environment. *Educational Technology & Society*, 12(2):176–189, 2009.

[26] J. Sweller, J. G. van Merrienboer, and G. W. Paas. Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10:251 – 296+, 1998.

[27] T. Vogt and E. André. Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition. In *Multimedia and Expo (ICME05)*, pages 474–477, 2005.

[28] E. Vyzas and R. Picard. Affective pattern classification. In *AAAI Fall Symposium, Emotional and Intelligent: The Tangled Knot of Cognition*, pages 176–182, 1998.

[29] B. Woolf, W. Burleson, I. Arroyo, T. Dragon, D. Cooper, and R. Picard. Affect-aware tutors: recognising and responding to student affect. *Int. J. Learning Technology*, 4(3-4):129–164, 2009.