Empowering teachers to design learning resources with metacognitive interface elements

Judy Kay1, Sabina Kleitman2, Roger Azevedo3
1 School of Information Technologies, University of Sydney, Australia
2 School of Psychology, University of Sydney, Australia
3 Department of Educational and Counselling Psychology, Laboratory for the Study of Metacognition and Advanced Learning Technologies, McGill University, Canada

{Judy.Kay,Sabina.Kleitman}@sydney.edu.au, roger.azevedo@mcgill.ca

Abstract. It is becoming increasingly easy for teachers to create many forms of digital learning resources, such as web-based materials, quizzes and games. Extensive research in metacognition provides strong evidence that metacognitive interface elements could make such materials more effective. This chapter argues for the creation of such a set of metacognitive interface elements so a teacher can easily incorporate these into the learning resources they create. We illustrate this in two ways: a series of examples of a simple, but valuable and broadly applicable, interface element which asks learners to assess their confidence; and an overview of a rich and sophisticated example, the Metatutor system. We conclude with a research agenda for creating a toolkit of metacognitive interface elements.

Introduction

Metacognition refers to the executive processes of reflecting on and regulating one’s own thinking, that is, “thinking about thinking” [Flavell, 1979]. These processes, central to learning, are one of the three fundamentals of self-regulated learning [Schraw et al., 2006]. This makes it important for most learning contexts and especially for the many uses of digital learning resources for independent and lifelong learning [Azevedo et al., 2012; Kay, 2008]. Given its importance, it is unsurprising that there is a large body of research about metacognition; for example, see reviews [Dunlosky & Bjork, 2008; Dunlosky & Lipko, 2007]. These provide powerful insights into
how real-world learning can be improved. This metacognition research also has the potential to serve as a foundation for the design of interface elements that can improve learning and be useful in many different learning resources.

It is particularly timely to make it easier for teachers to build upon the existing body of knowledge about metacognition. This is because it is becoming increasingly common for teachers to create their own digital learning resources or to modify or adapt existing ones. This paper argues for the creation of a new class of interface element that is part of the authoring toolkit for digital learning resources. This would pave the way for teachers to readily incorporate these metacognitive interface elements into a learning resource.

For example, within schools, learner management systems have become mainstream and these provide authoring environments for teachers. It has also become increasingly common for teachers to use general tools to create web sites, wikis and blogs. These have important roles outside formal, school and university, settings. They are relevant for a broad range learning needs, including lifelong concerns, such as those associated with learning about health and wellness.

In this chapter, we illustrate how research in metacognition has the potential to provide valuable guidance about interface elements that can support metacognitive processes. This paper is intended for teachers who create learning resources in authoring environments. We also write for designers of environments supporting authoring of learning resources since it would be valuable if they created these elements for authors to include into their learning materials.

In the next section, we introduce one metacognitive interface element, chosen because it is simple to implement within real-life learning environment, easy for the user to understand and has broad applicability. Then we present an overview of MetaTutor, a sophisticated example with multiple, integrated metacognitive interface elements. We conclude with our vision for broad adoption of metacognitive interface elements in learning resources.

2 Metacognitive Elements

One of the challenges in adopting the results of research on metacognition is due to the complexity of the concept [Azevedo and Aleven, 2012]. For example, Azevedo [Azevedo and Hadwin, 2005] identified 33 self-regulatory processes: “(a) planning variables including planning, goal
setting, activating prior knowledge, and recycling goal in working memory; (b) monitoring activities including feeling of knowing (FOK), judgment of learning (JOL), monitoring progress towards goals, content evaluation, identifying the adequacy of information, evaluating the content as the answer to a goal, and self-questioning; (c) learning strategies including hypothesizing, coordinating informational sources, inferences, mnemonics, drawing, summarizing, goal-directed search, selecting new informational sources, free search, rereading, taking notes, knowledge elaboration, finding location in environment, memorizing, reading notes, and reading new paragraph; (d) handling task difficulties and demands including help-seeking behavior, expect adequacy of information, control of context, time and effort planning, and task difficulty; and (e) interest in the task or the content domain of the task.” Most of these involve metacognitive processes. This chapter gives several examples of one class of interface element, then shows a combination of metacognitive processes used by students when using an intelligent tutoring system to develop a conceptual understanding of human biology.

2.1 Feeling of confidence

Confidence levels are an important component of Feeling of Knowing as they reflect the degree of certainty about the accuracy of one’s own performance while doing a task. They index key metacognitive experiences that are closely tied to self-monitoring, learning, self-regulation and decision making [Allwood and Granhag, 1999, Efklides, 2008, Flavell, 1979, Koriat, 2012, Stankov and Kleitman, 2008]. This section introduces interface elements that enable learners to consider and report their confidence levels in their answers to tasks in a learning environment.

With the premise that “confidence controls action” [Gilovich et al., 2002]:248, measures of confidence in one’s own knowledge, opinions and answers have been used in a wide range of scientific and real-life domains including self-regulated learning, memory, metacognition, decision making, intelligence, eyewitness testimonies, neuroscience, and perception. (See [Dunlosky and Metcalfe, 2008, Koriat, 2012] for reviews.) Typically, a person is asked to express their confidence about their answers, opinions, decisions, perceptions during on-task cognitive activity, thus prompting immediate reflection on current performance. (See [Allwood and Granhag, 1999, Kleitman, 2008, Koriat, 2012] for reviews.)

These confidence levels, captured on a variety of seemingly unrelated cognitive tasks (opinions, mathematics, verbal, reasoning testing stimuli), are demonstrated to converge into the construct of
Confidence [see Kleitman, 2008; Stankov & Kleitman, 2008 for reviews]—a broad psychological trait spanning a variety of cognitive domains, and reflects the habitual way in which people assess the accuracy of their decisions, memories, and opinions.

Kleitman [2008] also showed that both over- and under-confidence, tends to be a stable characteristic of a person. Both can be damaging. “No problem in judgment and decision making is more prevalent and more potentially catastrophic than overconfidence” [Plous, 1993, p. 217]. On the other hand, under-confidence can also pose problems [Want and Kleitman, 2006], including anxiety and detrimental tendencies such as procrastination and negative thinking. Within formal settings, this makes it valuable for a teacher to learn whether a student tends to be over- or under-confident. For a learner, there is potential benefit in becoming self-aware, based on formative assessment of their confidence, for example in terms of the Feeling of Knowing. We now introduce examples of simple interface elements that enable a learner to think about their certainty about their response to questions in a learning interface.

For instance, in a recent study, Kleitman and Costa (2012) included a metacognitive interface element in the quizzes for a statistics subject. This was part of a tool for formative assessment, the Statistical Metacognitive Instrumentation Quizzes (Stats-mIQ). It was designed for a typical learning environment, as part of a senior compulsory for psychology students undergraduate statistics subject. The students see this as a series of quizzes, each comprised of multiple choice questions. The metacognitive element can be seen in the middle of Figure 1. It asks students to rate how confident they are that their answer was correct. Afterwards, they see feedback on their answers as can be seen at the bottom of the figure. After completing the whole quiz, the students were provided with more feedback on their overall accuracy, confidence (overall, and separately for correct and incorrect answers) and bias levels for this and the other quizzes (and attempts).

Demonstrating the effectiveness of the quizzes, the final exam mark was positively predicted by the total number of quiz attempts, a composite accuracy and confidence score, and students' prediction of their exam mark. The metacognitive feedback was shown to be of real benefit, subjective and objective, for the students struggling with the content of the course. That is, poorly performing students reported that they found the confidence procedure were very useful for their learning. Importantly, even a small degree of engagement with quizzes improves the accuracy of the exam mark prediction (minimizing mark-prediction bias). Although replication and extension is required, these
findings highlight the important role metacognitive feedback plays in an authentic tertiary education setting.

In another study [Gluga et al., 2012a,b], confidence ratings were used for two purposes. They served to evaluate a learner’s perception of their progress, based on confidence judgments at each stage of learning. They were also valuable input to the authors of the materials since they gave clear evidence about the appropriateness and difficulty levels of their learning stimuli. The interface element is illustrated in Figure 2. This is an online tutorial system designed to help teachers learn how to use the Bloom taxonomy to classify examination tasks in terms of the level of learning that they can assess. In the example shown, the system makes use of examples from computer programming fundamentals. The system has the following stages:

- It asks learners to assess their initial knowledge of Bloom—original feeling of knowing.
- Then, for each of the six Bloom levels, it presents a brief description and an example of a task. Learners self-assess their ability to correctly classify tasks at each Bloom level—their prediction of performance.
- The third stage involves working through a series of examples (see Figure 2) where the learner reads the example and selects the classification they consider most appropriate. In the figure, this is Synthesis. They also indicate their confidence, in this case 95% (on-task confidence). As shown in the figure, this interface asks learners to also indicate the reason for their classification and any reasons for uncertainty. Afterwards, the interface presents the remaining information visible in the figure: the expert classification and its explanation. If this differs from the learner’s classification, they are invited to explain that.

This interface element provided a valuable means to evaluate the effectiveness of the material in each part of the tutorial. Results indicate that for each learning category (ie the six Bloom levels), on-task judgments were higher than prediction judgments, indicating that participants felt more competent to apply the material learned after completing the tutorial.

To evaluate appropriateness of the learning stimuli, the on-task confidence levels were matched with the on-task accuracy scores (their difference being the Bias scores). It is well-established that misleading items result in high degrees of overconfidence for most learners, irrespective of their levels of knowledge [Gigerenzer et al., 1991, Juslin, 1994]. A high degree of overconfidence for a task strongly suggest it may be ‘tricky’ or misleading (see [Harvey, 1997] for a review), suggesting a need to reconsider this choice of example in the
future. Indeed, reviewing the material involved and the free comments provided by study participants, it became apparent that this was indeed a tricky and contentious case. This indicates the way that confidence interface elements, easily incorporated in such learning resources, can serve as a valuable diagnostic tool.

In another recent study, Jackson and Kleitman [2012] used confidence levels to investigate the role of confidence in decision-making tendencies. Their Medical Decision-making Test (MDMT) required participants to diagnose patients with fictitious illnesses and to indicate their confidence in the accuracy of their diagnoses. Participants were then asked to make a decision: either to administer treatment matching the diagnosis (direct), or request a blood test to accurately identify the patients' illness state (see Figure 3). Four novel reliable individual decision-making tendencies were established: ideal decision-making (patients cured without delays 100% of the time); adequate decision-making (patients cured at least 50% of the time); wasteful decision-making (50% of patients die due to unnecessary testing); and fatal decision-making (patients die due to incorrect diagnosis and treatment). Confidence predicted these tendencies incrementally after taking account of diagnostic accuracy, cognitive ability, personality, cognitive styles, gender, and age. These results demonstrate that confidence judgments play an important role in decision-making, and are valuable where individual differences in decision-making are of interest.

Confidence ratings could easily be incorporated within the many computerised learning contexts which require learners to provide answers/decisions/opinions. The judgments may be in terms of percentages (see Figure 1), using sliding scales (see Figure 2) or using verbal expressions e.g., Not sure, Just Guessing to Very Sure (see Allwood, Granhag, and Jonsson, 2006 for a review). The starting point of the scale must reflect the number of choices available to the learner. For example, a multiple choice question with five, four or two response options would have a minimum confidence level of 20%, 25% or 50% respectively (i.e., 100/5; 100/4; 100/2); even if the participant was guessing, there is a 1 in 5/4/2 chance that the answer is correct. With open-ended questions, the scale should start at zero, as the learner is not provided with the answers and has to generate them. The objectivity of these confidence ratings can be verified by comparing them with actual performance (‘confidence’ minus ‘accuracy’), giving immediate and reliable ‘bias’ scores (Stankov, 1999). Several other scores, or calibration indices, can also be collected (e.g., discrimination – an ability to discriminate between correct and incorrect answers; see [Yates, 1990] for a review) for each individual, for each performance and across many learning tasks. These would be collected with the aim
of fostering learners’ self-awareness of the state of their knowledge. When assessed over several occasions, over- and under-estimations reflect individual differences in habitual over- and under-confidence [Stankov, 1999] or habitual difficulties with discrimination [Jackson and Kleitman, 2012]. Importantly, the bias scores can also be used for the test’s diagnostic purposes, evaluating items to exclude ‘misleading’ or ‘tricky’ ones (see [Gluga et al., 2012a, Harvey, 1997, Kleitman and Stankov, 2001] for reviews).

3. Interface Elements to Facilitate Metacognition during Complex Understanding

This section of the chapter deals with the design of interface elements to assess and foster several metacognitive processes used by students during learning with MetaTutor. MetaTutor is a multi-agent, adaptive hypermedia learning environment, designed to train, model, and foster students’ self regulated learning (SRL) while learning about the human circulatory system. The primary goal underlying the design of MetaTutor is to investigate how advanced learning technologies (ALTs) can adaptively scaffold learners’ SRL while they learn about complex biological topics. MetaTutor is grounded in a theory of SRL that views learning as an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognitive and metacognitive processes in the service of those goals. More specifically, MetaTutor is based on several theoretical assumptions of SRL that emphasize the role of cognitive, affective, metacognitive, and motivational (CAMM) processes. Moreover, there is a fundamental assumption that learners have the potential to monitor and regulate their CAMM processes while developing a conceptual understanding of the science topic (e.g., the human circulatory system). Although all students have the potential to regulate, few students do so effectively, possibly due to inefficient, or a lack of cognitive, emotional or metacognitive strategies, and knowledge.

As a learning tool, MetaTutor has a host of features that embody and foster self-regulated learning (see Figure 4). These include four pedagogical agents (PAs) who guide students through the two-hour learning session and prompt them to engage in planning, monitoring, and strategic learning behaviours. In addition, the agents can provide feedback and engage in a tutorial dialogue in order to scaffold students’ selection of appropriate sub-goals, accuracy of metacognitive judgments, and use of particular learning strategies. The system also uses natural language processing (NLP) to allow learners to express metacognitive judgements such as JOLs (e.g., "I do not understand this paragraph on systemic circulation"), FOK (e.g., "I do not recall reading about the valves in the heart"), content evaluation (e.g. "this section in
not related to my current learning sub-goal), and monitoring of progress towards goals (by using a colour overlay on the current sub-goal list to indicate to students how much content they have covered that is directly related to their current sub-goal). In addition, the interface elements allow learners to click on the SRL palette and then choose their level of understanding on a 6-item Likert-scale before receiving a quiz. They can also use the interface to summarize a static illustration related to the circulatory system. Additionally, MetaTutor collects information from user interactions with it to provide adaptive feedback on the deployment of students’ SRL behaviours. For example, students can be prompted to self-assess their understanding (i.e., system-initiated judgment of learning [JOL]) and are then administered a brief quiz. Results from the self-assessment and quiz allow pedagogic agents to provide adaptive feedback according to the calibration between students’ confidence of comprehension and their actual quiz performance.

As illustrated in Figure 4, the system’s interface layout also supports SRL processes. For example, an embedded SRL palette provides the opportunity for students to initiate an interaction with the system according to the SRL process selected (e.g., summarize their understanding of the topic). Overall, in line with its theoretical foundations, MetaTutor supports and fosters a variety of SRL behaviours including prior knowledge activation, goal setting, evaluation of learning strategies, integrating information across representations, content evaluation, summarization, note-taking, and drawing. Importantly, it also scaffolds specific metacognitive processes such as judgments of learning, feelings of knowing, content evaluation, and monitoring progress towards goals. Overall, these interface elements have been shown to enhance students metacognitive awareness and monitoring of their learning processes and enhanced their ability to use various learning strategies to enhance their conceptual understanding (Azevedo et al., 2010, 2011, in press).

4 Summary and Conclusions

There has been considerable research exploring how best to incorporate metacognitive interface elements into learning environments so that they enhance learning. Some of that work has explored how to explicitly teach metacognitive skills, for example [Aleven and Koedinger, 2000, Azevedo and Cromley, 2004, Gama, 2004, Hacker et al., 1998, Roll et al., 2011]. Indeed, that work has emphasised the need for explicit teaching, arguing that many students do not use metacognitive skills effectively and independently [Aleven and Koedinger, 2000, Aleven et al., 2006, Azevedo & Witherspoon, 2009, Roll et al., 2011, Gama, 2004].
How does this relate to the simple confidence rating interface element? It is easy to create an interface that requires learners to answer questions about their confidence before they can progress. This partially avoids some of the problems identified in previous work where students simply did not use metacognitive interface elements that were available [Aleven and Koedinger, 2000, Azevedo and Cromley, 2004]. However, it does not guarantee that learners will take this seriously. In addition, calibration is important here. Just because we provide interface elements for them to assess their confidence, does not mean that will make accurate metacognitive judgments.

One valuable approach is to explicitly teach metacognitive skills and knowledge. This may be done simply, as in Stat-mIQ, which showed summary information about the student’s accuracy in self-assessment of their knowledge. Equally, teachers can explain the value of improving this metacognitive skill. These low-tech approaches can help learners calibrate their confidence scores and potentially help them to see greater value in doing the extra work of rating confidence.

There are many benefits that follow from standardisation of interface elements, and these apply for those we propose for metacognitive interface elements. When visually consistent interface elements appear in many learning resources, learners will recognise them. This is helpful in terms of usability and learnability of the interface. Learners may also appreciate that it may offer similar benefits to those they came to value in the past. Standard widget-like interface elements will make it easy for the teacher (or other designers of learning resources) to include them. In this case, it is important that infrastructure also gives useful feedback to teachers so that they can see the ways that learners are using the metacognitive elements. This should be available in a form that is useful for individual tutoring of the student and in aggregate form so the teacher can learn about the effectiveness of the learning material (as in our ProGoSs example).

This chapter has argued that there is potential value in creating metacognitive interface elements. We chose to focus on just one of these so that we could illustrate it with several examples. Our MetaTutor overview gave a taste of a far broader broad range of metacognitive interface elements [Azevedo et al., 2011]. We envisage that there is a role for more of these elements, too. The body of metacognition research points their importance. We propose several strands of a research agenda to exploit this research.

- A technical stand will create interface elements that can operate as plugins to formal learning environments such as LMSs as well as general tools like wikis.
• Another technical strand must deal with creation of infrastructure that integrates the data from interface elements to provide long term feedback to learners and teachers.
• At a very different level, there is a need for communication between practitioner teachers and researchers so that these elements can meet the needs of teachers, fitting into their teaching practice and building their understanding of the teaching metacognitive skills.

We envisage that these interface elements have a role that crosses learning contexts, from formal learning in schools and universities to broader lifelong learning. For example, personal informatics involves self-tracking and self-rating [Wolf, 2010]. Across these contexts, many people, including teachers in formal settings, create learning resources in many forms. These have not yet begun to include metacognitive interface elements. This chapter has presented examples illustrating potential metacognitive interface elements that could become part of authoring toolkits, so that many teachers can use them to create more effective learning resources.

---

Consider the model summary table (below) obtained from carrying out linear regression analysis in SPSS. Which of the following statements is correct?

<table>
<thead>
<tr>
<th>Model Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), nr3, nr2

(a) The variance in the DV which is accounted for by the IV(s) is 70.9%.  
(b) The variance in the DV which is accounted for by the IV(s) is 50.3%.  
(c) The variance in the IV(s) which is accounted for by the DV is 25.3%.  
(d) The variance in the IV(s) which is accounted for by the DV is 70.9%.

Correct: The proportion of variance (in percentages) in the DV accounted for by the IV(s) given by 100*R^2, here = 0.503.
Fig. 1. Example of a quiz question format, with confidence rating and showing the feedback for the correct response.

Fig. 2. Tutorial system for learning how to classify examination questions in terms of the Bloom taxonomy, with learners assessing their self-confidence in the classification.
Fig. 3. Example of a quiz question format, with confidence rating and showing the feedback for the correct response.

Fig. 4. Annotated screenshot of MetaTutor (Azevedo et al., in press).
Bibliography


Jackson, A.S., and S. Kleitman. To cure, to kill, or to test? Decision-making