Exploring the effectiveness of a novel feedback mechanism within an intelligent tutoring system

Jeremiah Sullins*
Department of Behavioral Sciences, Harding University, USA
Email: jsullins@harding.edu
*Corresponding author

Scotty D. Craig
Department of Technological Entrepreneurship and Innovation Management, Cognitive Science and Engineering Unit, Arizona State University, USA
Email: Scotty.Craig@asu.edu

Xiangen Hu
Department of Psychology, University of Memphis, USA
Email: xiangenh@gmail.com

Abstract: The purpose of this study was to explore the effectiveness of a new feedback mechanism within an intelligent tutoring system called AutoTutor LITE. Participants were randomly assigned to one of three feedback manipulation conditions within the context of complex scientific material: 1) learners’ characteristics curves; 2) random; 3) no feedback. Results revealed that the participants receiving the new feedback mechanism (LCC) showed significantly higher learning gains when compared to the random feedback or no feedback manipulations. Additionally, there were no differences discovered between random feedback and no feedback. Interpretation and implications of results are discussed.

Keywords: intelligent tutoring system; ITS; feedback; learning technologies; cognition; learning.


Biographical notes: Jeremiah Sullins is an Assistant Professor of Psychology in the Department of Behavioral Sciences at Harding University. As a cognitive psychologist, his primary interests are in the learning sciences. His research focus is on learning/educational technologies, with branches extending into emotions, pedagogical strategies, tutorial dialogue/interaction, mechanisms of feedback, and question asking.
Exploring the effectiveness of a novel feedback mechanism

Scotty D. Craig is currently an Assistant Professor in the Human Systems Engineering Program within the Fulton Schools of Engineering at Arizona State University. His research has been conducted in a transdisciplinary area of psychology, education, and technology. His lab, the Cognitive-Based Applied Learning Technology (CoBALT) Lab, seeks to provide cutting edge research at the intersection of human cognition, technology, and the learning sciences which can provide solutions to real world problems within education and training. The lab website, http://www.CoBALTlab.org, provides lists of current projects.

Xiangen Hu is a Dunavant Professor in the Department of Psychology and Department of Electrical and Computer Engineering at The University of Memphis (UoM) and Senior Researcher at the Institute for Intelligent Systems (IIS) at the UoM and Visiting Professor at Central China Normal University (CCNU). His primary research areas include mathematical psychology, research design and statistics, and cognitive psychology. His more specific research interests include general processing tree (GPT) models, categorical data analysis, knowledge representation, computerised tutoring, and advanced distributed learning.

This paper is a revised and expanded version of a paper entitled ‘Exploring the effectiveness of a novel feedback system’ presented at the annual meeting of the Southeastern Psychological Association, New Orleans, Louisiana, 15–18 February 2012.

1 Introduction

Tutoring is one of the most effective methods for knowledge acquisition. This has been repeatedly shown when one-on-one human interaction is compared to traditional classroom settings (Bloom, 1984; Cohen et al., 1982; Graesser et al., 2011; VanLehn, 2011). More specifically, Bloom (1984) reported that students working with expert human tutors obtain achievement levels that are two standard deviations higher than students in conventional instruction. However, research has also shown that most human tutors are often untrained paraprofessionals and peers that do not use sophisticated tutoring strategies (Graesser and Person, 1994). Additionally, research has shown that even ‘outstanding’ human tutors rarely use pedagogical techniques (e.g., Socratic tutoring strategies, modelling, scaffolding, and fading, reciprocal teaching and diagnosis of misconceptions) shown to be effective within the educational literature over the last 20–30 years (Cade et al., 2008). However, human tutors have been shown to be an effective means of promoting learning (Cohen et al., 1982; Fantuzzo et al., 1992; Rogoff, 1990) with effect sizes of $d = 0.79$ (VanLehn, 2011).

It has been shown that under the correct circumstances feedback can be an effective means of increasing student learning (Crooks, 1988; Black and Wiliam, 1998) and is a method that all human tutors implement to some degree (Bloom, 1984; McKendree, 1990; Shute, 2008; VanLehn, 2011). There are many definitions for the term feedback. However, for the current purposes, feedback will be defined as all comments provided to a student following an appraisal of a student response (Sadler, 2010). According to some research, for feedback to be truly effective the following criteria must be present:
complementing students on the strength of their work

telling them gently about their deficiencies

where those deficiencies occurred and their nature

telling students what would have improved their submitted production

pointing them to what could be done next time they complete a related type of response (Sadler, 2010).

Unfortunately, in order to provide all of the previously mentioned criteria, a massive amount of time and resources must be devoted to every individual student, which is unrealistic. The good news is that research has shown the potential benefits that can be provided by intelligent tutoring systems (ITSs) (VanLehn, 2011).

As mentioned earlier, Bloom (1984) discovered that students who have access to one-on-one human tutoring outperform conventional instruction by two standard deviations, which Bloom referred to as the two-sigma problem. However, a more recent review has shown that Bloom’s initial projection might be inflated. VanLehn (2011) performed a systematic review of the tutoring literature and found that human tutoring, while still very effective, was at a much lower level of $d = 0.79$. This effect size was comparable to the effect size of ITSs at $d = 0.76$. As a result of this, the goal was to design educational environments that could be as effective as human tutors but were also affordable and able to be disseminated to a large population. For example, there is substantial evidence that one such ITS called AutoTutor improves learning (Graesser et al., 2004; VanLehn et al., 2007) compared to reading a textbook for an equivalent amount of time; the mean effect size is 0.8 over 18 studies in physics and computer literacy. So it is very conceivable that ITSs can provide high quality and affordable individualised instruction to numerous students at the same time. However, one question that warrants further investigation is: What mechanism within ITSs are accounting for student learning gains?

2 **Feedback types and implications to ITSs**

Feedback used in the context of educational settings is typically regarded as being of utmost importance when dealing with knowledge and skill acquisition (e.g., Azevedo and Bernard, 1995; Bangert-Drowns et al., 1991; Corbett and Anderson, 1989). Feedback has also been considered a significant factor not only with knowledge but also motivation in the context of learning (Narciss and Huth, 2004). Cohen (1985, p.33) stated that “feedback is one of the more instructionally powerful and least understood features in instructional design”. Although a large number of research acknowledge the importance of feedback, unfortunately the corpus of research consists of many conflicting findings and inconsistent results.

Nicol and Macfarlane-Dick (2006) examined the benefits of formative feedback and how this feedback affects students’ self-regulated learning. Formative feedback is defined as “information communicated to the learner that is intended to modify his or her thinking or behavior to improve learning” [Shute, (2008), p.153]. Seven principles of good formative feedback were proposed in the relation to self-regulated learning:
Help clarify what good performance is. This principle suggests that students can only achieve learning goals if they understand those goals, assume some ownership of them, and can assess progress.

Facilitates the development of self-assessment (reflection) in learning. In other words, teachers need to create more structured opportunities for self-monitoring and the judging of progression to goals.

Delivers high quality information to students about their learning. This refers to the need for teachers to provide the feedback rather than peer feedback. The reason for this according to Nicol and Macfarlane-Dick (2006) is that teachers provide feedback in which students can evaluate their progress and check out their own internal constructions of goals, criteria and standards. The argument is made that teachers are more able to provide good quality feedback. Good quality feedback is defined here as “information that helps students troubleshoot their own performance and self-correct: that is, it helps students take action to reduce the discrepancy between their intentions and the resulting effects” (p. 9).

Encourage teacher and peer dialogue around learning. According to research, students often have trouble truly understanding the feedback that they receive from tutors. One way of alleviating this potential problem is to conceptualise feedback as a dialogue rather than information delivery.

Encourages positive motivational beliefs and self-esteem. It is believed that motivation and self-esteem play an important role in learning and assessment (Dweck, 1999). Because of this, teachers have the power to have a positive or negative effect on motivation and self-esteem. For example, influencing goal setting of students, along with praising and focusing students on the feedback they receive.

Provides opportunities to close the gap between current and desired performance. According to researchers, two important questions that need to be asked are ‘is the feedback of the best quality?’ along with ‘does it actually bring about direct changes in student behaviour?’. Many researchers have focused on the first question while largely neglected the second. Teachers not only need to focus on providing high quality feedback to students but also set up evaluations in order to test that the feedback is actually changing learners’ behaviours.

Provides information to teachers that can be used to help shape the teaching. Good feedback is not only beneficial for learners but also beneficial for teachers. For example, by giving feedback to the learner and assessing the understanding of the feedback, teachers are able to tailor their future teaching to the specific needs of the student.

Shute (2008) provided a thorough review on formative feedback. Although it is beyond the scope of this article to go into all of the recommendations provided by Shute, there are a few ‘take away’ findings that warrant mentioning. More specifically, Shute discovered that formative feedback is similar to a ‘good murder’ in that it must comprise of three important components:
Overall, despite considerable variability, Shute (2008) discovered through a meta-analysis that formative feedback generally improves learning compared to control conditions. In other words, “formative feedback has been shown in numerous studies to improve students’ learning and enhance teachers teaching to the extent that the learners are receptive and the feedback is on target (valid), objective, focused and clear” (p.182). The previously mentioned studies suggest that feedback is a beneficial strategy for student learning.

However, as can be seen from the previous paragraphs, providing well designed feedback is a laborious process on the part of the teacher. In recent decades, ITSs have been developed in order to alleviate the load that can potentially be placed on the individual teacher who may not have the time/resources available for every individual student. In general, ITSs have been shown to be effective in a variety of domains. However, one area that has received recent attention is the use of feedback within these systems.

For example, Lin et al. (2013) explored ‘simple feedback’ versus ‘elaborate feedback’. The study was a 2 × 2 (agent present/agent absent; simple feedback/elaborate feedback) factorial design. According to Lin et al. (2013), simple feedback consisted of verifying right or wrong after the participant responded to a multiple choice question (e.g., ‘yes, that’s correct’ or ‘no, that’s wrong’). Elaborate feedback not only included whether the answer was right or wrong but also gave a justification as to why the answer was right or wrong. For example, “Your answer was wrong because temperature is a measure of the average kinetic energy of the particles in a substance, not a process of energy transfer” (p.243).

Results revealed a significant interaction. More specifically, participants that were placed in the ‘agent present/simple feedback’ condition scored significantly lower than the participants in the ‘agent present/elaborate feedback’ condition suggesting that the degree of feedback (as opposed to simply the presentation of feedback) influences depth of processing.

Roll et al. (2011) investigated whether immediate metacognitive feedback on students’ help seeking behaviour could help students acquire better help seeking skills. Metacognitive feedback refers to feedback that is triggered by students’ learning behaviour and not by the accuracy of their responses at the domain level (p.268). Participants were randomly assigned to one of two different conditions. In one condition participants interacted with ‘help tutor’ which was an enhanced version of the Geometry Cognitive Tutor. In this condition, participants received immediate metacognitive feedback on their help seeking behaviour. In the second condition, participants interacted with an unmodified cognitive geometry tutor (control condition). In regards to help seeking behaviour, participants who were in the ‘help tutor’ condition differed from the control condition on several actions. Participants in the ‘help tutor’ condition made significantly fewer help seeking errors (26% vs. 36%). There was also a difference in the level of hints that participants asked to see. Participants in the ‘help tutor’ condition asked to see the bottom out hint only 48% of the time compared to the participants in the control condition who asked to see the bottom out hint 70% of the time. Overall, the
results from the study demonstrated that for common types of errors metacognitive feedback can be used to improve students’ behaviour in the context of an ITS.

In another study, McKendree (1986, 1990) examined the efficiency of content feedback within the context of an ITS for geometry proofs. McKendree (1986) investigated a geometry tutor using error-flagging and error-remediation feedback and found that the remediation group performed better than the error-flagging group at immediate evaluation. As a follow-up to this, McKendree (1990) compared a generic error-flagging feedback, which simply stated that an error was present, with more context-specific error explanations that stated conditions present in the error and/or possible inferences along a solution path. An example of error flagging was, ‘This rule does not apply with the premises you have chosen’, whereas a more context specific explanation was, ‘To solve this problem, it requires selecting three congruent sides. You have only chosen two congruent sides along with 1 other side’. As a result of these feedback manipulations, McKendree found that her participants in the context-specific feedback conditions retained benefits at long term evaluation, demonstrating fewer errors in geometry proofs.

Research by Conrad et al. (2005) investigated the use of progress feedback within survey methodology. Respondents in these surveys are expected to answer a series of questions on the survey but may become less motivated when they have no idea of how much progress they are achieving. This runs the risk of respondents dropping out and not completing the experiment/survey. The research conducted by Conrad et al. (2008) has focused on decreasing the dropout rate and improving retention by means of progress feedback. Conrad et al. (2005) reported that respondents often make an initial and lasting impression of their progress while answering a set of questions. Respondents who experienced an initial slow rate of progress were more likely to ‘break off’ and not complete the survey than those who experienced a steady or initially quick rate of progress. A second study, also reported in Conrad et al., suggested that an intermittent display of progress in the late phase of the experiments may provide the most benefits and fewest costs, when compared to steady progress or on-demand progress feedback.

Additionally, Jackson and Graesser (2007) examined the effects of both content feedback and progress feedback in the context of an ITS called AutoTutor. More specifically, participants were trained using one of four different versions of AutoTutor to assess the effects of both content feedback and progress feedback. The four different versions of AutoTutor were as follows:

1. content and progress feedback condition
2. content feedback condition
3. progress feedback condition
4. no feedback.

Results revealed that the presence or absence of content feedback was responsible for almost all of the significant effects in the experiment. Additionally, it was discovered that the presence or absence of progress feedback had little to no effect of student learning.

As can be seen from the previously mentioned studies, the research on feedback is inconsistent regarding exactly what makes feedback beneficial for learners. However, it does seem that overall feedback is beneficial for both learners and teachers. The purpose of this study is to explore a new feedback system (learner’s characteristics curves – LCC)
implemented within an ITS, AutoTutor Learning and Instruction in Training Environments (LITE).

3 AutoTutor and AutoTutor LITE

AutoTutor is a fully automated tutor that holds conversations with learners in natural language and that simulates the dialogue moves of human tutors (Graesser et al., 2005). AutoTutor presents students with a series of challenging problems (or main questions), each requiring approximately 3 to 7 sentences of information for a correct answer. When presented with a problem, students typically respond with answers that are between one word to two sentences in length. In order to guide students in their construction of an improved answer, AutoTutor actively monitors students’ knowledge states and engages learners in a turn-based dialogue. AutoTutor adaptively manages the tutorial dialogue by providing feedback, pumping the learner for more information, giving hints, correcting misconceptions, answering questions, and summarising answers. As students move through their learning sessions, AutoTutor adheres to constructivist theories of learning by guiding students to actively construct answers to difficult questions, as opposed to simply presenting the solution to the students.

In order to correctly answer a main question, students typically have to construct dozens of dialogue turns. After each student dialogue turn, AutoTutor compares the learner’s response to expectations and common misconceptions for that main question. AutoTutor leads the student through the relevant expectations and assists their learning by using a specific dialogue structure referred to as a hint-prompt-assertion cycle. This cycle is designed to gradually shift the cognitive burden from student to the tutor. Specifically, AutoTutor first presents the student with a hint that requires the student to provide the majority of the information. If the student does not supply an adequate answer, AutoTutor then prompts the student for more information. If the student is still unable to provide the correct answer desired by the prompt, AutoTutor provides an assertion. If the student provides a satisfactory answer at any point during the hint-prompt-assertion cycle, AutoTutor exits the cycle and moves to the next expectation. If the student provides a satisfactory answer that covers the last remaining expectation, then AutoTutor asks the student to restate a complete answer and subsequently provides a summary of what an ideal answer would look like. For a detailed discussion of AutoTutor’s dialogue mechanisms and strategies, see Graesser et al. (2005).

It is important to emphasise that AutoTutor is not merely a scripted information delivery system, but rather is adaptive to the student’s knowledge and the dialogue history. AutoTutor provides the student with feedback (positive, neutral positive, neutral, neutral negative, and negative) after the student expresses information in a conversational turn. AutoTutor selects hints, prompts, and assertions in a dynamic fashion that is sensitive to the responses of the student.

The AutoTutor LITE system is based on AutoTutor. One challenge of the original AutoTutor system was its scalability due to its dependence on language analysers. The version of AutoTutor LITE used in this study is a minimalist implementation of AutoTutor. It only included an AutoTutor style interface and interaction with a lightweight language analyser. This provides the learner with a streamlined tutorial interaction that relying on tutor hints and feedback for tutoring coherent brief chunk of information called a Shareable Knowledge Object or SKO (Hu et al., 2014).
Similar to AutoTutor, Autotutor LITE interacts with students using natural language and is most effective when the learning objectives are qualitative/conceptual. AutoTutor LITE requires users to construct an answer to the question. A typical system interaction starts with a general seed question. The system evaluates the student’s answer and asks follow up questions, which it selects based on the student model. AutoTutor LITE provides feedback and selects the next questions based on the four indices of learners characteristics curves (discussed in subsequent sections). The current implementation of AutoTutor LITE uses extended weighted keyword matching and latent semantic analysis. See Figure 1 for an example screen capture of AutoTutor LITE.

**Figure 1**  Example screen capture of AutoTutor LITE (see online version for colours)

3.1 Learners’ characteristics curves

The light-weight language analyser implemented in AutoTutor LITE is used to create a simple student model called LCC. There are four curves collectively describing what students ‘know’ about the expected answer of the seed question. The four curves are measures of student’s input in four different indices: relevant-new, relevant-old, irrelevant-new, and irrelevant-old. The four values are computed for each of the student’s contribution:

- relevant-new: relevant to the answer and was not included in the previous answers
- irrelevant-new: irrelevant to the answer and was not included in the previous answers
- relevant-old: relevant to the answer and was included in the previous answers
- irrelevant-old: irrelevant to the answer and was included in the previous answers
• irrelevant-new: irrelevant to the answer and was not included in the previous answers.

The four numbers are computed from the light-weight language-analyser. AutoTutor LITE provides feedback and selects the next questions based on a function of these four indices. LCC is a simplification of the student’s model in comparison to other sophisticated ITS implementations. In AutoTutor LITE, LCC is enough for the tutor to offer appropriate feedback and question selection. For example, the non-decreasing trend of relevant-new can be an indication of active construction of answers, while positive values of irrelevant-new can be an indication of knowledge deficits for the answers.

The quality of LCC is dependent on the quality of language analyser implemented. The current implementation of AutoTutor LITE uses extended weighted keyword matching and latent semantic analysis. LSA has been widely used in information retrieval (Dumais et al., 1988; Graesser et al., 2002b), similarity measurements between texts (Hu et al., 2007), text cohesion analysis (Foltz et al., 1998), and ITSs (Graesser et al., 2000, 2002a).

The basic idea is as follows and depicted in Figure 2. First, instructional content (metadata from SCOs, text from the instructional content, questions, and expected answers) is represented in the form of semantic vectors. Second, the student’s contribution is encoded using the same semantic engine. The similarity score is computed between the student’s input and the stored answers. Finally, feedback is presented so the student knows if the response to the tutor’s question was relevant to the answer.

Figure 2   AutoTutor LITE logic model (see online version for colours)

The entire process can be understood in terms of keyword matching (Hu et al., 2007). Consider a case where a student responds to a tutor’s question in multiple turns. In other words, the student might not provide the complete answer in a single response. Therefore, the system would need to make determinations on the subsequent information provided. This information could be either old (o) or new (n) information. It could also be relevant (r) to the topic or irrelevant (i). If we denote the sequence of contributions as \( s_i \), \( i = 1, \ldots, I \), for every contribution, the tutor would give feedback based on the four different types of information for each contribution from the student.
It is understandable that a human tutor would offer positive feedback when a student is providing new and relevant (N-R) contributions. Furthermore, if a student is actively constructing relevant answers, one would see a non-decrease value for the cell (N-R) in a sequence of responses. In the same fashion, other cells can be used as an indication of a student’s knowledge. For example, an increasing value for the (N-I) would indicate the lack of relevant knowledge. We call them the student’s characteristics curves (LCC).

One of the challenges of an ITS is to create a student model (Graesser et al., 2001) that adequately assess a student’s knowledge. For example, an experienced human tutor can estimate how much a student knows or does not know by evaluating a student’s answers to key questions. The human tutor can provide feedback to help a student actively construct responses that are relevant to the questions asked. AutoTutor LITE used LCC as a student model and to offered appropriate feedback. For a more detailed discussion regarding the mechanics behind LCC please refer to Wolfe et al. (in press, 2012).

4 Current experiment

Does the LCC methodology within AutoTutor LITE demonstrating an effective feedback mechanism for an ITS? In other words, is this visual, direct, ‘just in time’ feedback beneficial to learning, compared to various feedback controls?

Because LCC is able to determine the relevance of a learner’s contributions (both previous and current) and use this to provide appropriate feedback, it is hypothesised that learning would be improved using this system. To establish support for this hypothesis, feedback based off of LCC would need to both show that it improves learning and that it is the specific LCC based feedback that is adaptive to the learner that is causing the learning and not any feedback. This hypothesis was tested using a repeated measures design with feedback based on the LCC model, a random feedback control and a no feedback condition. The no feedback condition allows for learning predictions of the occurrence of learning to be tested. The random feedback condition indicates if any observed learning would occur with any additional feedback. These allows for the establishment of an adaptivity effect for LCC.

4.1 Experiment predictions

Based on previous research, there are two possible outcomes to the current study. According to the content matters hypothesis (Jackson and Graesser, 2007; McKendree, 1986, 1990), no differences would be expected among the LCC feedback condition, the random feedback condition, and the no feedback condition.

Conversely, according to the progress matters hypothesis (Conrad et al., 2005), we would expect that learning in the LCC feedback condition to be significantly greater than the learning in both the random feedback condition and the no feedback condition.
5 Methods

5.1 Participants

Participants consisted of 17 students who were recruited through the university’s subject pool. Participants took part in this study in order to fulfill a course requirement.

5.2 Methods and procedures

5.2.1 Experimental design

The study was a within subject design which contained a pretest phase, a training phase, and a posttest phase. During the implementation of the design, participants received three different types of feedback: LCC, random feedback, or no feedback that were incorporated into three AutoTutor LITE modules. Each module was a different topic: bioluminescence, phenylthiocarbamide (PTC), and matter. The order in which participants received the content along with the feedback that accompanied the content was all counterbalanced. For example, participant one might get bioluminescence with LCC feedback during session one, matter with random feedback for session two, and PTC with no feedback for session 3.

Participants received three different types of feedback depending on which condition they were currently working in. The first condition called LCC feedback consisted of the full working version of AutoTutor LITE that contained feedback based on the LCC recommendation. A second condition called random feedback consisted of a version of AutoTutor LITE in which students received random feedback regardless of the student’s input. A third and final condition called no feedback consisted of a version of AutoTutor LITE in which students received no feedback regarding their input. Participants received information during their AutoTutor LITE interactions on the following topics: bioluminescence, matter, and PTC. This base content did not change thus keeping the conditions content equivalent with only feedback type changing.

Each module started with a brief two minute information delivery on the topic. At the end of each module, AutoTutor LITE would ask the participant ‘Why do not you tell me a little bit about what you have learned today?’ Participants would type their answer. The student’s answer would be processed against a preexisting ideal answer using semantic space based natural language processing technique based on the LCC algorithm. The feedback that the students received was in graphical form throughout the duration of all three modules. No additional feedback was presented. (See Figure 3, for example, output presented to students). During the tutoring interaction, participants would type their answer until they were satisfied with their scores. Each time the answer was changed, participants received a new bar representing the current feedback. So, participants could track progress over time.
As can be seen in the previous figure, participants are able to view their feedback in graph form. The feedback received contained information on the number of attempts to answer each question along with the total coverage for the problem along with the current score. With each participant entry, they were able to see their score based on four different types of information: new and relevant information, new and irrelevant information, old and relevant information, and old and irrelevant information. In the LCC feedback condition, all of the graphical feedback that the participant received was correct and based on their actual input. In the random feedback condition, the graphical feedback that the student received was random and did not take into consideration the participant input. Finally, in the no feedback condition, participants saw no graphical feedback on graphs. In other words, a student would be randomly assigned to an AutoTutor LITE module covering information on bioluminescence. The student would hear a brief AutoTutor LITE monologue lasting approximately two minutes. Following the monologue, AutoTutor LITE would then prompt the student to write down as much as they could regarding the previously covered information. Depending on the type of feedback associated with the module, the student would either receive accurate coverage feedback in graphical form (LCC), a random graphical output not actually related to the
coverage of material (random), or no feedback at all based on their response (no feedback).

- **Pretest phase.** During the pretest phase, participants were brought into the lab (one participant at a time) and sat in front of a laptop computer. They were then read instructions for the study. Following the instructions, all participants were given a pretest that assessed prior science knowledge. The prior knowledge science test consisted of 20 multiple-choice questions and assessed students on the following areas: math, biology, chemistry, research methods and earth science. Some example questions from the prior knowledge science test can be seen in Table 1.

- **Training phase.** Following the completion of the prior knowledge science questions, all participants were given instruction on how to interact with the system as well as how they should interpret their feedback from the diagrams displayed after a student’s input. Following the training phase students were assigned to three different versions of Autotutor LITE.

- **Posttest phase.** Following the completion of all three AutoTutor LITE modules, participants completed a 20 item multiple-choice posttest assessing their knowledge on the information they received over the course of the module.

**Table 1** Example prior knowledge questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Option</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which of these has a positive charge and is found in the nucleus of an atom?</td>
<td>a. Neutrons</td>
<td>0.32</td>
</tr>
<tr>
<td>b. Protons</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>c. Electrons</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>d. Elements</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Which unit of measurement is the largest?</td>
<td>a. Yard</td>
<td>0.32</td>
</tr>
<tr>
<td>b. Foot</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>c. Metre</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>d. Millimetre</td>
<td>0.32</td>
<td></td>
</tr>
</tbody>
</table>

**6 Results**

A repeated measures ANOVA was performed on the participant’s posttest scores to determine differences between the three feedback types. Analysis revealed a significant difference among the three different types of feedback $F (2, 30) = 4.120, p < 0.05$. More specifically, students who received feedback via LCC ($M = 1.29$ (32% correct), $SD = 1.21$) significantly outperformed students who received random feedback ($M = 0.94$ (23% correct), $SD = 0.82$) and no feedback ($M = 1.00$ (25% correct), $SD = 0.93$). There was no difference between no feedback and random feedback.
7 Discussion

The current study found that feedback based on the LCC student model led to greater participant learning than participants receiving no feedback or random feedback. This indicates that the LCC student model was effective for improving learning (LCC > no feedback) and that the adaptivity that the model provided was the driving factor for the learning not just the added content (LCC > random feedback). So, the LCC model offers a simple metric for increasing interactive feedback into educational technologies such as ITSs.

The results from the current study suggest that providing feedback via LCC is effective as a source of feedback while learning complex scientific material. In other words, the data seem to provide support for the previously mentioned learners characteristics curves hypothesis. This is based on the fact that random feedback did not perform as well as LCC feedback although both of these types of feedback provide nothing more than a learner’s progression through the material. This result may be due the fact that learners were cognisant of the fact that the ‘progress feedback’ they were receiving during the random feedback module was not a true representation of their progress through the module. In turn, they may have assumed something was wrong with the system and therefore ‘zoned out’ during this portion of the learning session.

Contrary to what earlier research suggests feedback alone, as indicated by the random feedback condition, is not enough to promote learning. Based on the observed results, it appears to be equivalent to not providing feedback. This provides more evidence that feedback needs to be targeted to the error and immediate to be effective (Shute, 2008).

Additionally, the results from the current study support that content feedback may not be a necessary step in order for learning to occur. Simply viewing the visual, direct, ‘just in time’, progress feedback, participants were able to significantly outperform their counterparts in the feedback control conditions. This is great news for educators who may not have the resources or time to provide individualised content feedback to every student in an overcrowded classroom. The results from the current study are also good news for researchers who are interested in building systems capable of delivering individual instruction to large groups of learners. For example, the cost of building an ITS that is capable of providing accurate, direct, content feedback has been estimated to be in the millions of dollars.

Furthermore, the results from the current study support earlier studies that feedback in general is an important component in learning. More specifically, when students are left on their own to master a concept, it is possible that no learning will occur or they may cling to old misconceptions.

For example, Muller (2008) was interested in the question of whether videos lead to meaningful learning in the context of physics. Participants completed a pretest, and then were randomly assigned to different video conditions, and then completed a posttest that was identical to the original pretest. The results revealed no learning differences from pretest to posttest. However, the students claimed that the videos were easy to understand and were confident that they understood the material. It was speculated that the videos were perhaps too passive based on student responses to the videos. In fact, some students said “I was listening but I wasn’t really paying attention”, during the debriefing. Further results revealed that student did not correctly remember what was in the videos. In
particular, student misconceptions that existed prior to watching the video persisted after the correct information was presented in the video.

Additionally, a study conducted by Azevedo et al. (2008) explored differences in student learning when students learned in the presence of a tutor (externally facilitated self-regulated learning) or when students learned in isolation (self-regulated learning). Participants were told to learn all they could in 40 minutes about the human circulatory system within the context of a hypermedia environment. The only difference between the two conditions was the fact that participants in the externally facilitated self-regulated learning condition have access to a tutor. The tutor was allowed to, among other things, prompt students to use several effective strategies, such as summarising, coordinating informational sources, hypothesising, drawing, and using mnemonics. In other words, the tutor was allowed to provide feedback to the participants. Results revealed that participants that had access to a human tutor (i.e., feedback) learned significant more knowledge regarding the circulatory system. Furthermore, participants in the externally facilitated self-regulated learning condition showed more advanced learning strategies during their interaction with the hypermedia environment.

Further research is warranted in order to explore under what other circumstances progress feedback is beneficial for learning. It is possible that providing progress feedback may only be productive for learning when the learner task does involve complex scientific material (i.e., STEM fields). Furthermore, progress feedback may only be beneficial for students that enter the learning session with high domain knowledge of the content area to be learned. In other words, it may take more targeted content type feedback when a learner does not have the foundation to build upon while learning new material.

8 Conclusions

When designing information delivery systems, it might not be of crucial importance to provide learners with content feedback as opposed to true progress feedback (i.e., learners characteristic curves). In other words, learning can occur just by informing students of their progress through the system using learners’ characteristics curves. This is good news for researchers that do not have the resources to invest in a system that is capable of providing content feedback (e.g., ITSs). Although the exact dollar amount is unknown regarding how much it costs to develop one of these systems, we do know that it is estimated to be in the millions. However, the fact is that developing an environment that is able to provide progress feedback is cost efficient and easily accessible.

References


