Chapter 15
Incorporation of Agent Prompts as Scaffolding of Reflection in an Intelligent Learning Environment

Longkai Wu and Chee-Kit Looi

National Institute of Education
1 Nanyang Walk, Singapore
{longkai.wu,cheekit.looi}@nie.edu.sg

Abstract. Recent research has emphasized the importance of reflection for students in an intelligent learning environment. But, researchers have not reached a consensus on the most effective ways to design scaffolding to prompt reflection, nor have they accepted a common mechanism that can explain the effects of scaffolding on reflection. Two types of agent prompts to foster reflection are contrasted in this chapter, both from the perspective of a tutee, differing in their specificity. Generic prompts are content-independent tutee questions, aiming at fostering students' reflection on metacognitive strategies and beliefs regarding their learning-by-teaching activities. Specific prompts, on the other hand, are content-dependent tutee questions that encourage students' reflection on domain-related and task-specific skills, and articulation of their explanatory responses. This chapter describes the design and effect of these two types of agent prompts, adapted to students' learning-by-teaching activities, on the learning outcomes, the elicited levels of reflection, and the self-efficacy of the secondary school students.

15.1 Introduction

The educational research literature suggests that questions prompts, whether from teachers, peers or textbooks, could promote reflection by eliciting explanations. (Rothkopf 1966) investigates the ways in which questions inserted in texts affected subjects' understanding of the texts. (Chi et al. 1994) indicate that questions that elicit self-explanations lead to improved understanding of texts. And students, who provide explanations to other students' questions or explain examples found in their textbooks, seem to strengthen connections among their ideas (Davis 1998). (Moon 2004) further suggests structuring reflection with questions to deepen the quality of reflection.

A. Peña-Ayala (Ed.): Intelligent and Adaptive ELS, SIST 17, pp. 369–391.
springerlink.com

Researchers in Intelligent Learning Environment (ILE) have recognized the importance of incorporating question prompts into ILE design (Hmelo and Day 1999). Question prompts are used as scaffolds to help direct students towards learning-appropriate goals, such as focusing student attention and modeling the kinds of questions students should be learning to ask (Azevedo and Hadwin 2005). Positive evidences are found for question prompts to help students with various aspects, such as, knowledge integration (Davis and Linn 2000) and ill-structured problem-solving processes (Ge and Land 2004; Xie and Bradshaw 2008). Meanwhile, mechanisms for supporting self-explanation, tutorial dialog or reflective dialog (Aleven and Koedinger 2002; Grassser et al. 2001; Katz et al. 2000) have been prevalent in traditional intelligent tutoring systems (ITS), in which the computer plays the role of tutor (e.g., Cognitive Tutor, AutoTutor).

This chapter is concerned with the investigation of agent prompts (i.e., question prompts to initiate learners' reflection in learning, within an agent-enabled learning-by-teaching context). The mechanism of generating agent prompts defines how to assess and model learners' reflective learning-by-teaching activities and metacognitive skills. We intend to explore how a learning partner, acting as the role of inquisitive tutee enabled by the generation of question prompts, might be used to address the challenge of facilitating reflection in a student tutor linked to learning-by-teaching activities. Here, the reflection of the student tutor mainly refers to an intermingled process of knowledge construction and metacognition as a direct result of his engagement in instructional activities inherent to the virtual tutoring process, such as explaining, answering questions from the tutee, correcting errors of the tutee and asking questions to the tutee (Cohen 1986; Garneret, et al. 1971; King 1998). The opportunity for reflection enables the tutor to monitor his own understanding, recognize and repair knowledge gaps and misconceptions, integrate new knowledge with prior knowledge, and generate new ideas for self-evaluation and reflection (Roscoe and Chi 2007).

15.2 Review of Literature

The section firstly reviews how question prompts are used as the scaffolding approach to enhance students' reflective learning. Then we introduce tutee questions as types of question prompts to facilitate reflection and learning. An overview of metacognitive and cognitive strategies and beliefs is subsequently included, which leads the differentiation of generic prompts and specific prompts, in the context of tutee questions, that we are intended to investigate in this study.

15.2.1 Question Prompts

Classroom studies have suggested that prompts fostering reflection could be effective because they provide support for the cognitively complex ways learners think about, feel, and make connections in experience (Davis and Linn 2000). By engaging in reflective activities such as responding to the reflection questions, the learner builds their understanding and locates the significance of his activity in a larger context. Thus he is enabled to observe the meaning, he has taken from the experience and excavate the underlying qualities that make the experience significant. When the learner is prompted to deeper forms of reflection, it also becomes possible for him to identify learning edges, those questions or issues that he is seeking to understand in order to advance his work (Amulya 2004). In doing reflection stimulated by prompts, the learner can unpack the richness of the experience and evaluate which issues emerging from that experience need to be pursued.

15.2.2 Tutee Questions

Meanwhile, recent research also shows the evidence of learning benefits to tutors from tutee's question prompts in the context of peer tutoring. (Cohen et al. 1982) demonstrate empirical evidence of learning gains for tutors compared to non-tutors in the context of peer tutoring. (King et al. 1998) specially study the tutor's explanations and questioning in the tutoring process as the sources for tutor's learning based on high-level question stems (i.e., questions prompting for comparisons, justifications, causes-and-effects, evaluations, etc.). (Graesser et al. 1995) show that the tutee's occasional "deep" questions out of major "shallow" questions can stimulate the tutor's deeper response. (Coleman et al. 1997) demonstrate very similar findings in collaborative learning settings with students using high-level explanation prompts. (Roscoe and Chi 2004) find that in a non-reciprocal and naturalistic (i.e., little or no training) tutoring context, the tutee's questions can motivate tutor explanations and metacognition, and thus have a significant and positive influence on the tutor's learning activities and opportunities.

(Graesser et al. 1995) discuss the kinds of tutee questions that occur during tutoring, which can be divided into shallow and deeper questions. Shallow factual questions ("what" questions) ask definitions or simple calculations while deeper questions ("how" and "why" questions) ask about causal relationships and underlying principles, requiring elaboration, inference and logical reasoning. Peverly and Wood 2001 indicate that deeper questions support learning more efficiently than shallow questions.

15.2.3 Cognitive and Metacognitive Strategies

(Cornford 2002) notes that cognitive strategies and metacognitive strategies are closely related since both of them involve cognition and skill but conceptually they are quite distinct. Cognitive strategies are used to help an individual achieve a particular goal (e.g., understanding a text) while metacognitive strategies are used to ensure that the goal has been reached (e.g., quizzing oneself to evaluate one's understanding of that text) (Cornford 2002). (Weinstein and Meyer 1991) state: "A cognitive learning strategy is a plan for orchestrating cognitive resources, such as attention and long-term memory to help reach a learning goal". They indicate that there are several characteristics of cognitive learning strategies, including that they are goal-directed, intentionally invoked, and effortful that are not universally applicable, but situation specific.

Comparatively, (Schraw 1998) notes that metacognitive strategies appear to share most of these characteristics, with the exception of the last one, since they involve more universal application through focus upon planning for implementation, monitoring and evaluation. It means that metacognitive strategies are not so situation specific but, involve truly generic skills essential for learner, more sophisticated forms of thinking and problem solving.

15.2.4 Generic and Specific Prompts

The generic prompts, or called "general tutee questions" are a series of content-independent questions to lead students to reflect on metacognitive strategies and beliefs in learning and teaching, consider various perspectives regarding their activities, such as "Why should you teach?", "Before starting to teach, can you think about what you are supposed to learn from it?", "What do you learn from me as your tutee?". (VanLehn, Jones and Chi 1992) suggest: "Gneric prompts could increase the chances that individual students will be able to identify gaps in their own understanding, discover deficiencies in their mental models, or generate useful inferences."

The specific prompts, or called "specific tutee questions", on the other hand, are a series of content-dependent questions to lead students to reflect on task-specific and domain-related skills regarding their activities and to articulate their explanatory responses, such as "Can you explain the concepts you just taught me?", "If you query me by asking me a casual question in the below window, you can see how I reason through the concept map that you have taught me. Can you tell me if my reasoning process is correct and give me a further explanation?" Specific prompts appear to be helpful in getting some students to realize that they have a gap in their understanding and may even hint at how to fill the gap (Aleven et al. 2006; VanLehn et al. 1992).

15.3 System Design

This section describes the design and implementation of agent prompts within the Betty's Brain system with both pedagogical and technological considerations. This research has been concerned with the creation of agent prompts that explores new scaffolding approaches in an intelligent learning-by-teaching environment. We puts forth the scheme and architecture of an agent prompts generator which produces two types of prompts to guide students' reflection within a learning-by-teaching environment. The purpose of designing agent prompts generator is to enhance the learning-by-teaching environment, reifying the metacognitive, task-specific and domain-related reflection involved in such activities.

15.3.1 Overview

The development work focused on the generation and incorporation of meaningful agent prompts, which can arouse student's reflective learning-by-teaching activities. The work was built on an existing system, Betty's Brain (Fig. 15.1), a learning-by-teaching agent environment built by the Teachable Agent Group in Vanderbilt University (Biswas et al. 2001). With the ability to learn what the students have taught by concept mapping, Betty's Brain was used to play the role of agent tutee in our research.

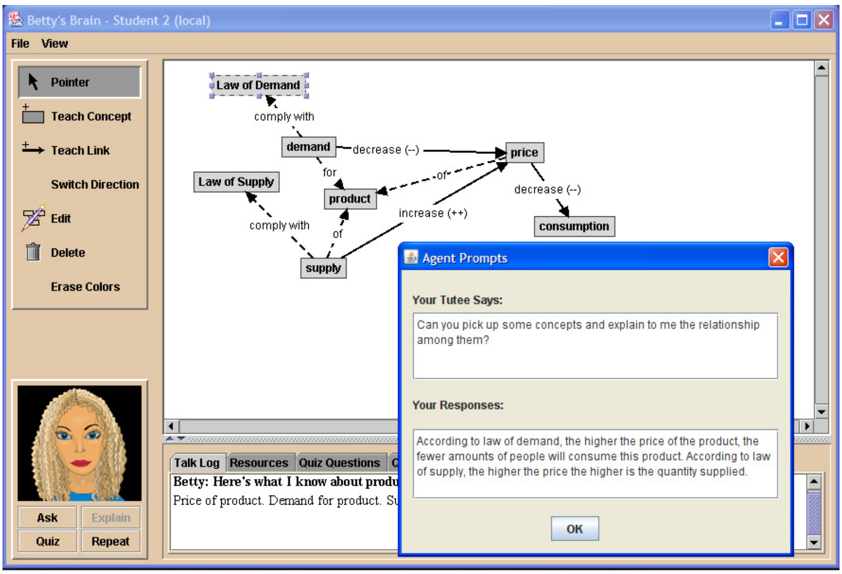


Fig. 15.1 Adapted Betty's Brain (Biswas, et al. 2011) in basic economics

We adapted Betty's Brain to become an inquisitive agent tutee, with the built-in agent prompts generator, in the domain of basic economics. When students interact with the inquisitive agent tutee version of Betty's Brain, they have to respond to the question prompts from the agent system. These question prompts were produced in the agent prompts generator, which could analyze the semantic structure of student concept maps and compare them with expert maps tailored in the domain of basic economics. Our goal was to foster a reflective student-agent learning-by-teaching interaction for better learning outcomes.

15.3.2 Aspects of Consideration

The consideration of generating appropriate question prompts is to guide students in learning-by-teaching activities, by provoking different reflection types in their responses and exploring situational cues and underlying meanings relevant to the context. Students were expected to recognize the importance of reflective activities while, their cognitive load was not significantly increased.

Based on the literature review and the design principles of an agent tutee system, three aspects (Fig. 15.2) were considered, when the agent prompts scheme was designed: Learning-by-Teaching Stages, Reflection Types and Patterns in Student Maps.

Aspects	Value
Learning-by-Teaching Stages →	Familiarization Production Evaluation Post-Task Reflection
Reflection Types →	Double-Loop Reflection Single-Loop Reflection
Patterns in Student Maps →	Missing or Incorrect Expert Concepts Missing or Incorrect Expert Propositions

Fig. 15.2 Aspects of consideration

15.3.3 Learning-by-Teaching Stages

Our learning-by-teaching activities were categorized into four stages as shown in Fig. 15.3, which follow the conceptual stages in practicing tutoring with a metacognition instruction model that focuses on the following metacognitive skills: 1) problem understanding and knowledge monitoring; 2) selection of metacognitive strategies; 3) evaluation of the learning experience (Gama 2004).

- Stage 1 Familiarization, understanding, and planning: This stage contains two types of reflective activities that include self-assessment of the understanding of the domain knowledge and difficulties and self-selection of metacognitive strategies,
- Stage 2 Production, teaching, presenting answer, and answers: This stage is devoted to enable students to teach the agent tutee what they have learned to the agent tutee by constructing concept maps and monitoring the agent tutee's understandings,
- Stage 3 Evaluation, evaluating the performance: This stage provides students with opportunities to evaluate the performance of the agent tutee, as well as their own performance,
- Stage 4 Post-Task Reflection, Reflecting on Learning-by-Teaching Experience: This stage is oriented to promote post-practice reflection on the learning-by-teaching processes and the strategies implemented.

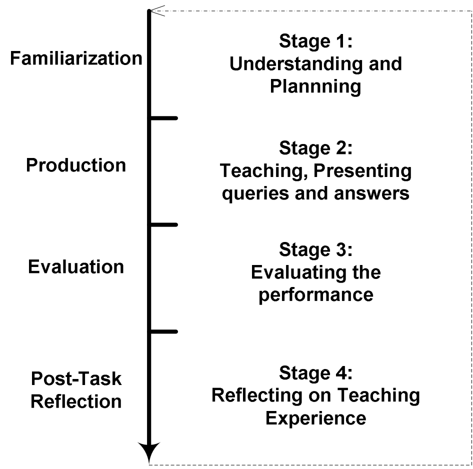


Fig. 15.3 Stages of learning-by-teaching activities

15.3.4 Reflection Types

Considering both the theoretical and practical perspectives, we design agent prompts to foster two major types of reflection for students in the learning-by-teaching environment.

- Generic prompts fostering double-loop reflection are content-independent, stimulating students to monitor their learning-by-teaching processes and consider various perspectives regarding their learning-by-teaching activities. Double-loop reflection focuses on the examination and reflection of the theory or perspective in use,
- Specific prompts fostering single-loop reflection are content-dependent, providing students with a structure through the learning-by-teaching process and lead them to complete a specific cognitive task and articulate their explanatory responses. Single-loop reflection refers to increasing efficiency of an objective, which is task oriented and is about the design of the process to retain reliability.

The content of sample generic prompts containing metacognitive strategies and beliefs, which was developed from the adaptation of Reflection Assistant Model (Gama 2004) attempting to simulate a tutee's perspective, is depicted as follows.

- List of strategies for assessing beliefs:
  - What do you think about teaching and who is it for?
  - Why should you teach?
  - What do you think about what you are supposed to learn from teaching?
  - What do you learn from me as your tutee?

- List of strategies for monitoring understanding:
  - Can you read your learning objectives more than once?
  - Can you read the learning objectives to separate the important parts?
  - Can you think of a related learning task you have already done and use it as an example?
  - Before starting to teach, can you think about what you are supposed to learn from it?
  - Can you read the learning objectives and determine which parts you don't understand well?
  - Can you review the basic concepts that are not clear, before you begin to teach?
  - Can you set a goal, and think about the steps to reach this goal?
- List of strategies for controlling errors:
  - Can you stop and review each part in the map to see if a mistake has been made?
  - Can you reread the resources to check for missing important parts?
  - Can you change strategies if you feel lost and confused and don't seem to move anywhere?
- List of strategies for revising:
  - Can you think about a way of checking to see if your map is correct?
  - Can you review all things done to make sure nothing is missing?
  - Can you reread the learning objectives and resources and ask if the map really meets the description in the learning objectives and resources?

The content of sample specific prompts containing domain-related and task-specific skills, which was developed from the adaptation of Teachable Agent (Leelawong 2005), was partly illustrated as follows.

- Read the on-line resources to learn:
  - Can you check the on-line resources for more information and tell me more?
- Request explanation on concepts or propositions:
  - Can you explain the concepts you just taught me?
  - Can you pick up two concepts from that section and explain to me the relationship among them?
- Query to teach better:
  - A good teacher asks students questions to make sure that they understand things correctly. You can ask me by clicking on the *Ask* button underneath the window. Let me know if my answer is useful by offering a description.

- Query causal questions in order to teach:
  - If you ask me a casual question, you can see how I reason through the concept map that you have taught me. Can you tell me if my reasoning process is correct and give me a further explanation?
- Ask for quiz in order to teach:
  - I have learned something from you. Please require me to take a quiz. Can you give an evaluation comment on my quiz performance?

15.3.5 Patterns in Student Maps

Our intention is to develop an evaluation scheme for student concept maps, which supports and facilitates the identification and categorization of faulty propositions in a concept map.

We classify learners' errors based on the pattern categories that are presented below, including missing or incorrect expert concepts and expert propositions. Moreover, this categorization is used as the basis for the construction of agent prompts that reflect the different types of student errors.

- Missing expert concepts:
  - The student omits specific concepts (which are considered fundamental concepts of the subject matter) from their maps. The usual omissions of specific concepts lead us to the conclusion that the student manifests incomplete understanding.
- Incorrect Expert Concepts:
  - The students make mistakes on specific concepts (which are considered fundamental concepts of the subject matter) in their concept maps. The usual mistakes of specific concepts lead us to the conclusion that the student manifests incorrect understanding.
- Missing Expert Propositions: The student uses specific relationships between two or more concepts, which are not false but they do not correctly/fully address the relation of these concepts in the context of the subject matter. He does not relate two or more concepts denoting their relationship. These cases are considered as an evidence of incomplete understanding.
- Incorrect Expert Propositions: The student relates two or more concepts that should not be related, and/or with incorrect relationships that lead to clearly false propositions. The mistakes of expert propositions lead us to the conclusion that the student manifests incorrect understanding.

15.3.6 Architecture of Agent Prompts Generator

In Fig. 15.4, the architecture of Agent Prompts Generation System is depicted with three major components involved, namely the Agent Prompts Generator, the Map Comparer, and the Stage Detector.

15.3.6.1 Agent Prompts Generator

The Agent Prompts Generator monitors the student's concept mapping activities (i.e., the student teaches Betty to tailor a concept map in the agent environment) and plays the role of coordinator in the system. It mainly receives the results from the Map Comparer, selects proper prompts from the repository of question prompts and sends them to the Reflective Dialogue window for students to respond to. The student tutor receives these prompts in the Reflective Dialogue window and responds while, teaching the agent by modeling in a Concept Map Editor.

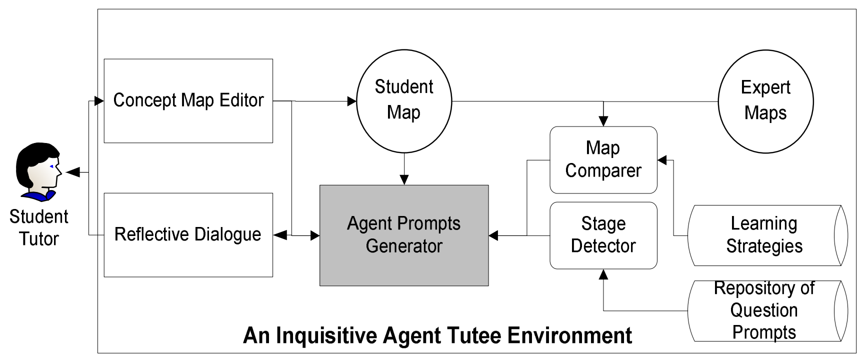


Fig. 15.4 Architecture of Agent Prompts Generator

15.3.6.2 Map Comparer

The map comparer detects patterns and assesses student performance by comparing student map to an integrated expert map. It selects proper prompting strategies for an individual student from the repository of learning strategies, which stores tactics and strategies specified to the agent environment (e.g., examining agent's understanding, observing the agent's independent performance, etc.), based on his concept mapping practice.

An overlay assessment method is used to evaluate students by matching their student maps with expert maps. To be accurate for measurement, an integrated expert map is considered by integrating a number of expert maps developed by separate experts (teachers and researchers). We adopt a fuzzy integration technique (Chen et al. 2001) based on fuzzy set discipline that attempts to produce an integrated expert map that could be superior to any of the individual expert map.

The algorithm for detecting patterns in student concept maps, with a breadth-first search through the expert map (Kornilakis et al. 2004), begins at the central concept ("Demand" concept in our case). A queue is used to collect the concepts that have not yet been searched. The algorithm appears as in the next pseudo code.

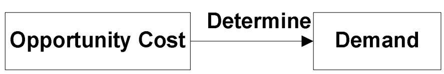
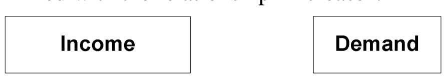
Combining this with the breadth-first nature of the algorithm, we can be certain that it will always be possible to find the student concept node corresponding to the expert concept node in the expert map.

The algorithm is also guaranteed to end either after finding a pattern or after confirming that the student map matches the expert map.

The correspondence between map patterns and agent prompts was pre-defined based on the design principles within an Adaptive Feedback Framework (Gouli et al. 2005). Table 15.1 shows the samples of patterns found in student maps and appearing as corresponding agent prompts stored in the repository of question prompts.

```
Insert central concept in queue
While the queue is not empty repeat
v <-the first concept in the queue
v' <-the concept corresponding to v in the expert map
E <-the set of all links coming out of student map
E' <-the set of all links coming out of expert map
For every link e' in E'
  d' <-the destination concept of e'
  If e' not in E
     If there exists a link in E ending on d'
       If there exists a link in the expert map from n'
                to d' other than e'
          Return Missing Expert link
       Else
         Return Incorrect Expert Link
     Else
       Search the student map for d'
       If found
          Return Missing Expert Link
       Else
          Return Missing Expert Concept
  Else
    e <-the link corresponding to e' in the student map
    If the destination concept of e is not d'
       If there exists a concept other that d'
           in the expert map connected to n' with a link e'
          Return Incorrect Expert Concept
       Else
          Return Missing Expert Concept
          For every link e in E
            If e not in E'
               Return Missing Expert Link
```

Table 15.1 Agent Prompts Generation based on Pattern Detection

Pattern	Description
Missing expert concepts	Missing concept and its relationships when specific concepts defined in concept map are omitted from the student map
	E.g. The concept of "Income" and its links with the concepts of "Demand" are missing from the student map.
	
	Generic prompt: "*Can you review all things to check for missing important parts to teach me and give me an explanation*?"
	Specific prompt: "*Do you consider that you could teach me the concept of 'Income' and give me an explanation*?"
Incorrect expert concepts	A concept is related to an incorrect concept which should be replaced with another concept.
	E.g. In the proposition "Demand *affects* Elasticity", the concept "Elasticity" should replace the concept "Electricity".
	
	Generic prompt: "*Can you stop and review each part in the map to see if you have made a mistake and give me an explanation*?"
	Specific prompt: "*Do you want to reconsider the concept of 'Electricity' you have taught me and give me an explanation*?"
Incorrect expert propositions	Two concepts are related even though they should not be.
	E.g. The proposition "Opportunity Cost *determines* Demand" is incorrect, as the concepts "Opportunity Cost" and the "Demand" are not related. So, the link "determine" should be omitted.
	
	Generic prompt: "*Can you stop and review each part in the map to see if a mistake has been made and give me an explanation*?"
	Specific prompt: "*Do you want to reconsider the link of 'Determine' between 'Opportunity Cost' and 'Demand' and give me an explanation*?"
Missing expert propositions	The proposition of two expert concepts is missing on the student concept map.
	E.g. The concepts "Income" and the "Demand" are not related although they should be linked with the relationship "increase".
	
	Generic prompt: "*Can you review all things done to make sure nothing is missing and give me an explanation*?"
	Specific prompt: "*Do you want to consider teaching the link between 'Income' and 'Demand' again and give me an explanation*?"

15.3.6.3 Student Performance Assessment

We use an overlay assessment method proposed by Chang et al. (2005) to compare student concept maps with expert concept maps in order to assess student performance. By comparing concept maps drawn by the students with the expert concept maps, the students' comprehension of each proposition can be determined.

A student's comprehension has one of the following learning states: the proposition is learned (complete expert concepts and expert propositions), partially learned (missing expert concepts or missing expert propositions), unlearned (no expert concepts or no expert propositions), or the student has a misconception about the proposition (incorrect expert concepts or incorrect expert propositions).

The learning state of the student revealed in his map is used to grade student's performance and determine the frequency of agent prompts delivery. Once the similarity between student map and expert map is high enough, the agent prompts are removed just as traditional scaffolds were faded.

The proposed method considers propositions based on weighted concept maps.

Let $G_e = (V_e, E_e)$ be an expert concept map. If $(v_i, v_j) \in V_e$ and $e_{ij} \in E_e$, then (v_i, e_{ij}, v_j) represents a proposition in G_e if the relation link e_{ij} connects two concept nodes v_i and v_j. Any proposition (v_i, e_{ij}, v_j) can be compared with the propositions in a student concept map. From the resulting comparison, it is possible to decide if the proposition (v_i, e_{ij}, v_j) is learned, partially learned, unlearned, or if the student has a misconception. The following procedure shows how the comparison is performed:

- If there is a proposition (v_i, e^*_{ij}, v_j) in the student concept map, then
  - If $e^*_{ij} = e_{ij}$, then (v_i, e_{ij}, v_j) is learned.
  - If $e^*_{ij} = \phi$, then (v_i, e_{ij}, v_j) is partially learned.
  - If $e^*_{ij} \bullet e_{ij}$, then the student has misconception about (v_i, e_{ij}, v_j).
- If there is a proposition (v_j, e^*_{ij}, v_i) in the student concept map, then
  - If $e^*_{ij} = e_{ij}$ or $e^*_{ij} = \phi$, (v_i, e_{ij}, v_j) is partially learned.
  - If $e^*_{ij} \bullet e_{ij}$, the student has misconception about (v_i, e_{ij}, v_j).
- If the proposition (v_i, e^*_{ij}, v_j) or (v_j, e^*_{ij}, v_i) does not exist in the student concept map, then (v_i, e^*_{ij}, v_j) is not learned.

In order to quantify the similarity between student map and expert map, all student propositions scores according to the student's learning state are calculated and then the similarity index S is computed by the means of equation 15.1.

$$S = \frac{\sum_{i,j} score(v^*_{p_i})}{\sum_{v_{p_i} \in V_p} w(v_{p_i})}, 0 \le S \le 1 \tag{15.1}$$

In this equation, v_{p_i} is a proposition in the expert map and $w(v_{p_i})$ is its weight. The $score(v^*_{p_i})$ is the score assigned to the proposition $v^*_{p_i}$. After calculating the scores for all student propositions, the similarity index S is achieved and used to measure how similar the student's knowledge structure is to the expert's. The larger the index, greater is the similarity. Once S equals 1, the agent prompts is totally removed to let students fully concentrate on the concept mapping activities.

15.3.6.4 Stage Detector

Apart from Map Comparer, the Agent Prompts Generator also needs the Stage Detector to detect the student's learning-by-teaching stage as discussed in the previous sections, which includes Familiarization, Production, Evaluation and Post-Task Reflection, and stimulate student in different cognitive and metacognitive aspects of reflection in learning.

The Stage Detector detects the stage in which the student is involved and retrieves and selects the appropriate agent prompts within the repository of question prompts. Examples of question prompts stored in the repository adapted to the four stages are the following.

- Agent prompts in the familiarization stage:
  - Generic prompts: What do you think about teaching and who is it for?
  - Specific prompts: How will you comment on the materials you will teach me at the beginning?
- Agent prompts in the production stage:
  - Generic prompts: How do you want me, as your tutee, to deal with you?
  - Specific prompts: It seems you have taught me several concepts. Can you choose some to explain to me?
- Agent prompts in the evaluation stage:
  - Generic/specific prompts: What is the most important thing you have tried to teach me?
- Agent prompts in the post-task reflection stage:
  - Generic prompts: What is your thinking after teaching Betty?
  - Specific prompts: Can you provide your further comments on the evaluation on Betty (advantages, weaknesses, expectations, etc…)?

15.4 Empirical Study

This study is to investigate whether the proposed framework is able to help learners in reflection and learning.

15.4.1 Participants and Procedure

Participants were 33 students from two grade levels (level 1 and 2) in two local secondary schools (ages ranged from 13 to 15), who took part in the experiments on a voluntary basis for two-hour sessions within 2-week period (Table 15.2).

Table 15.2 Procedure of Empirical Study

Phases	Activities	Description
Phase 1	Pre-test	MSLQ and Knowledge Pre-test
Phase 2.1	Tutoring: Familiarization	Get familiar with materials and simulated tutee
Phase 2.2	Tutoring: Production	Teach simulated tutee by concept mapping
Phase 2.3	Tutoring: Evaluation	Check the performance of simulated tutee
Phase 2.4	Tutoring: Post-Task Reflection	Reflect upon own performance
Phase 3	Post-test	MSLQ and Knowledge Post-test (1 week later)

They were randomly assigned to one of the three conditions to study elementary economics of demand and supply. Economics is a theoretical and applied domain, seldom studied in secondary school and rarely adopted as the domain in ILE research. The domain materials were provided to participants before the sessions.

In short, 29 students (76%), 20 female (69%) and 9 male (31%) completed all activities within the 2-week period of the trial, resulting in the next division over the three conditions: no prompts (NP) condition as control group: n = 10, specific prompts (SP) condition: n = 10 and generic prompts (GP) condition: n = 9.

During the tutoring phases, participants were working with the simulated tutee system to teach what they learnt from materials by constructing concept maps. The NP group (n=10) worked with the basic version of simulated tutee without prompts. The SP group (n=10) worked with the version embedded with specific prompts. The GP group (n=9) worked with the version embedded with generic prompts. Both the SP and GP groups were required to write down their reflection statements in the dialog window to respond to the simulated tutee prompts to proceed with their tutoring activities. A sample of response statements from participants to two types of prompts are as follows:

```
…
[Simulated tutee detects decreasing of missing expert propo-
sitions in the production phase] ¨
Can you pick up some concepts and explain to me the relation-
ship among them? (Specific Prompts)
[SP Student] According to law of demand, the higher the price
of the product, the fewer amounts of people will consume this
product. According to law of supply, the higher the price the
higher is the quantity supplied
…
[Simulated tutee detects start of the post-reflection phase]
What is your thinking after teaching me? (Generic Prompts)
```

```
[GP Student] You are a curious student by asking a lot of
questions to me. But sometimes, I don't quite understand what
you are asking me to do. I need to learn more about demand
and supply to teach you better.
…
```

We further categorized the participants into High and Low group according to their self-efficacy scores in their MSLQ (Motivated Strategies for Learning Questionnaire) pre- and post- test. Participants scored above the mean self-efficacy score in MSLQ pre-test were included in the High group and the rest were included in the Low group.

Zimmerman (2000) notes that self-efficacy has emerged as a highly effective predictor of students' motivation and learning. As a performance-based measure of perceived capability, self-efficacy differs conceptually and psychometrically from related motivational constructs, such as outcome expectations, self-concept, or locus of control. We selected nine questions from the MSLQ developed by Pintrich and DeGroot (1990), in which Questions 2, 6, 8, 9, 11, 13, 16, 18, and 19 are serving the purpose of determining one's self-efficacy. These questions helped us to avoid subject bias or subject characteristics threat by clarifying students' tendencies in learning activities.

15.4.2 Impact of Question Prompts on Knowledge Pre-/Post-Test

Fig. 15.5 uses error bars of the pre-test and post-test scores to represent the pre-test/post-test scores across groups. It indicates that there was a tendency for learners in all three experimental groups to achieve approximately the same level in the pre-test while both the two prompted conditions (GP and SP) outperformed the non-prompted condition in the post-test significantly (the result of ANOVA test is $F(2, 25) = 19.55, p < .05$).

We also compared the pre-test-to-post-test effect sizes (Cohen's d) of the three conditions. As seen in the Table 15.3, the two prompted conditions yielded an average effect size of 2.84, outperforming the non-prompted condition ($d = 2.15$). The difference between the SP group ($d = 2.37$) and the GP group ($d = 3.30$) is also prominent.

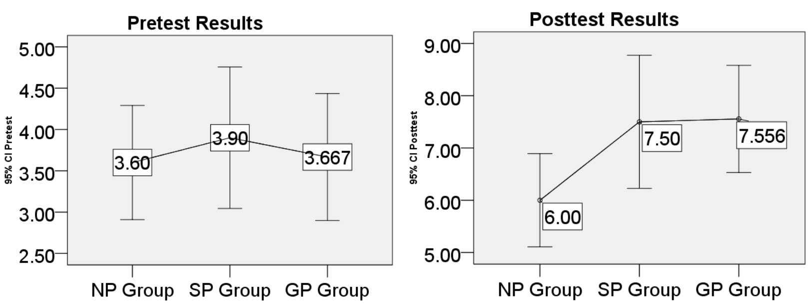


Fig. 15.5 Pretest and Posttest Scores

Table 15.3 Result of Knowledge Pre-/Post-Test

Groups	N	Pre-test (Mean/SD)	SE Post-test (Mean/SD)	Cohen's *d*
NP Group	10	3.60 (0.97)	6.00 (1.25)	2.15
SP Group	10	3.90 (1.20)	7.50 (1.78)	2.37
GP Group	9	3.67 (1.00)	7.56 (1.33)	3.30

15.4.3 Impact of Question Prompts on Self-Efficacy Development

To test students' development of self-efficacy, we compared the students' pre-to-post scores in the pre-test on self-efficacy. The data is reported in Table 15.4. The effect size computed as Cohen's *d* suggests that, although most groups experienced little progress in self-efficacy during such a relative short-term experimental study, the GP Low group has experienced a prominent progress ($d = 6.27$) than others. A Tukey's test, performed to compare the difference between groups, also reveals that there is a significant difference between the GP Low group and the Control Low group ($MD = 21.75$, $p = 0.003$). It suggests that generic prompts could support students with low self-efficacy to develop their level of self-efficacy through the learning activities.

15.4.4 Impact of Self-Efficacy on Adoption of Question Prompts

Table 15.5 shows the result of domain knowledge pretest to posttest. It shows the effect size (Cohen's *d*) of each group from knowledge pretest to posttest. Students evaluated as having high efficacy receiving generic prompts acquired the most progress (Cohen's $d = 4.17$). Such a result implies that students with high efficacy, who received generic prompts could achieve better learning outcomes than other groups of students.

Table 15.4 Result of Knowledge Pre-/Post- Test

Groups	N	SE Pre-test (Mean/SD)	SE Post-test (Mean/SD)	Cohen's *d*
Control Low	5	23.00 (10.36)	27.00 (10.36)	0.39
Control High	5	43.89 (7.08)	45.00 (8.00)	0.15
GP Low	4	20.50 (4.20)	48.75 (4.78)	6.27
GP High	5	47.00 (7.00)	49.40 (4.87)	0.40
SP Low	5	25.40 (8.93)	35.40 (8.82)	1.13
SP High	5	45.20 (8.29)	43.40 (6.07)	-0.24

Table 15.5 Result of Domain Knowledge Pre-/Post- Test

Groups	N	Pre-test (Mean/SD)	Post-test (Mean/SD)	Cohen's *d*
GP High	5	4.25 (0.96)	8.25 (0.96)	4.17
GP Low	4	3.20 (0.84)	7.00 (1.41)	3.27
SP High	5	4.00 (1.58)	7.60 (2.30)	1.82
SP Low	5	3.80 (0.83)	7.40 (1.34)	3.22

15.4.5 Impact of Question Prompts on Elicited Reflection Levels

A tripartite coding scheme, adapted by Ortiz (2006) from the categorization scheme proposed by Surbeck et al. (1991), was adopted to analyze the participant' response statements to the agent prompts. This scheme, includes categories of reaction, elaboration, and contemplation, from lowest to highest, adopts the perspective that the nature of the stimulus to reflect will impact the quality of the reflection (Moore and Whitfield 2008), which conforms to what this work study.

An initial observation of the transcript excerpts of prompt-response exchange showed that the GP participants exhibited a higher level of reflection than the SP participants. We noted that the GP participants gradually tended to present more completive response statements than the SP participants after being exposed to the question prompts for a period of time. A sample transcript of participants' response statement is outlined as follows to demonstrate both SP participant and GP participant responded to the purposefully designed same question prompts but showed different levels of reflection in their response statements.

[Simulated Tutee] What is the key thing you try to teach me?
[SP Participant] What is economics? (Reaction: Report facts)
[GP Participant] The universality of these links, causes and implications. (*Contemplation: Indicate a shift in thought*)
[Simulated Tutee] What is the key thing you learn from me?
[SP Participant] Demand, supply and price are interrelated and contribute to the idea of economics in total. (*Elaboration: Provide interpretive analysis of course concepts*)
[GP Participant] What to do when approaching economics in real life. (*Contemplation: Indicate a shift in attitude about himself*)
[Simulated Tutee] Do you learn something from me?
[SP Participant] Yes, I learn how to organize my thoughts better. (*Contemplation: Indicate a shift in attitude about himself*)
[GP Participant] How to teach better and more clearly. (*Contemplation: Indicate a shift in attitude about himself*).

Two raters were involved in the analysis of response statements. To begin, the first rater (the first author) analyzed a sample of ten response statements to become confident with the coding scheme. Then, the first rater explained the coding scheme to the second rater (a researcher specialized in educational research).

The second rater analyzed the first ten response statements as well and the differences between the two raters were discussed to differentiate between *reaction*, *elaboration,* and *contemplation*.

After that, the two raters analyzed the remaining response statements independently. Cohen's kappa was used to estimate the level of agreement between the 2 raters, taking the agreement occurring by chance into account. Calculated with Cohen's Kappa, the agreement between the two raters appeared to be good, $\kappa = 0.653$ (A Cohen's kappa value between 0-0.4 is considered poor, 0.4-0.6 fair, 0.6-0.75 good, and 0.75 outstanding).

A combined qualitative and quantitative analysis of participants' response statements to simulated tutees' question prompts (a total number of 286 response statements were analyzed) showed the difference in the levels of reflection between groups (Table 15.6). An ANOVA test shows a statistical significant difference between the groups as to reactive statements (F (1, 17) = 36.747, p <.05) and contemplative statements (F (1, 17) =19.472, p < .05). The number of elaborative statements was not significantly different between the groups. Such a result shows that the participant of GP group, whether with high or low efficacy, was more likely to respond with contemplative statements representing a higher level of reflection. Comparatively, the participants of SP group, whether with high or low efficacy, responded more with reactive statements representing a lower level of reflection which means they pay more attention to report issues with no development than the GP group.

15.5 Conclusions

The results suggest that the use of agent tutee as an active and inquisitive learning partner to raise meaningful questions could be helpful to students learning by reflective teaching. The preliminary results, quantitative and qualitative, suggest that the generation of agent prompts as computer-based scaffolds, when adapted to student learning-by-teaching activities, could be useful in particular ways, such as improving students' learning outcomes and eliciting higher levels of reflection.

Table 15.6 Result of Response Statements Analysis

Levels of Reflection	GP (Mean/SD)	SP (Mean/SD)	ANOVA-Test[1]
Reaction	6.00 (1.41)	9.60(1.17)	36.75*
Elaboration	7.56 (1.54)	9.50(2.17)	4.19
Contemplation	9.00 (1.80)	5.10 (2.82)	19.47*

*, p < .05

Specially, this study suggests that generic prompts could possibly be beneficial in the development of self-efficacy, thus, leading to fostering self-directed learning. However, there remain significant challenges that must be overcome before a system similar to the one used in this study, with agent prompts, can be incorporated into regular classroom use. Generally, the goal of supporting students to learn needs a shift from a notion of leading to one of facilitating and enabling. This means designing agent tutee systems that are not necessarily recipients of information but, rather, a facilitator to promote reflection. Instead of attempting to create agent tutee systems that will always know the correct answer, designers need to invent agent tutee systems that encourage students to do both reflection-in-action and reflection-on-action (Schön 1987) and attend more on double-loop reflection (i.e., think out of the box) (Argyris and Schön 1996).

A future research direction might be to investigate the design of current learning-by-teaching system to involve intelligent mixes of generic prompts and specific prompts. For example, an intelligent mix might give students specific prompts for their first experiences with prompting, and then fade into generic prompts.

An alternative improvement of system design would be tailoring the prompts given on the basis of individual student characteristics like self-efficacy. Students with high efficacy might be given the generic prompts, whereas, students with low efficacy might receive the intelligent mix of generic and specific prompts.

Possible future work also includes exploring the relationship between agent prompts and flow. Flow is a state of the mind, individuals reached that made them very productive in the task that they perform (Csikszentmihalyi 1978). As Csikszentmihalyi notes, persons in flow feel completely involved in, and focused on their task because of their curiosity or training. They feel great inner clarity, know that the activity is doable, do not notice time passing, and feel intrinsically motivated. Flow is a highly activated state of the mind that puts the person performing a task in control. Putting a learner in flow some of the time would be a highly appreciable goal for learning environments (Katzlberger 2005).

However, agent prompts inherently interrupts the learner's learning tasks at-hand. To achieve flow, one key issue is to facilitate incorporating and fading agent prompts for the students at the right moment. More refined approaches are needed to analyze, when the students are embarking on a new activity and when they have finished that activity and are almost ready to move on to the next steps. The answers will provide an in-depth understanding of intellectual enjoyment that students encounter, when using agent tutee systems as learning partners.

Acknowledgments. We thank Dr. Gautam Biswas in Vanderbilt for providing the Teachable Agent software.

References

Aleven, V., Koedinger, K.R.: An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor. Cognitive Science 26, 147–179 (2002)

Aleven, V., Pinkwart, N., Ashley, K., Lynch, C.: Supporting self-explanation of argument transcripts: specific v. generic prompts. In: Proceedings of ITS Workshop of Intelligent Tutoring Systems for Ill-Defined Domains, pp. 47–55 (2006)

Amulya, J.: What is Reflective Practice? (2004), http://www.itslifejimbutnotasweknowit.org.uk/files/whatisreflectivepractice.pdf (accessed April 18, 2009)

Argyris, C., Schön, D.: Organizational learning II: Theory, method and practice. FT Press, Mass (1996)

Azevedo, R., Hadwin, A.F.: Scaffolding self-regulated learning and metacognition-Implications for the design of computer-based scaffolds. Instructional Science 33, 367–379 (2005)

Biswas, G., Schwartz, D., Bransford, J. (TAG-V): Technology support for complex problem solving: From SAD Environments to AI. In: Feltovich, F. (ed.) Smart Machines in Education, pp. 71–98. AAAI Press, Menlo Park (2001)

Brockbank, A., McGill, I.: Facilitating reflective learning in higher education. Society for Research into Higher Education. Open University Press, Buckingham (1998)

Chang, K.-E., Sung, Y.-T., Chang, R.-B., Lin, S.-C.: A new assessment for computer-based concept mapping. Educational Technology & Society 8(3), 138–148 (2005)

Chen, N.S., Kinshuk, Wei, C.W., Liu, C.C.: Effects of matching teaching strategy to thinking style on learner's quality of reflection in an online learning environment. Computers & Education 56(1), 53–64 (2011)

Chen, N.S., Wei, C.W., Wu, K.T., Uden, L.: Effects of high level prompts and peer assessment on online learners' reflection levels. Computers & Education 52(2), 283–291 (2009)

Chen, S.W., Lin, S.C., Chang, K.E.: Attributed concept maps: fuzzy integration and fuzzy matching. IEEE Transactions on Systems, Man, And Cybernetics 31(5) (2001)

Chi, M.T.H., de Leeuw, N., Chiu, M., LaVancher, C.: Eliciting self-explanations improves understanding. Cognitive Science 18, 39–477 (1994)

Cohen, J.: Theoretical considerations of peer tutoring. Psychology in the Schools 23, 175–186 (1986)

Cohen, P.A., Kulik, J.A., Kulik, C.C.: Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal 19(2), 237–248 (1982)

Coleman, E.B., Brown, A.L., Rivkin, I.D.: The effect of instructional explanations on formal learning from scientific texts. The Journal of the Learning Sciences 6(4), 347–365 (1997)

Cornford, I.R.: Learning-to-learn strategies as a basis for effective lifelong learning. International Journal of Lifelong Education 21, 57–368 (2002)

Csikszentmihalyi, M.: Intrinsic rewards and emergent motivation. In: Lepper, M.R., Greene, D. (eds.) The Hidden Costs of Reward. Lawrence Erlbaum Associates, Hillsdale (1978)

Davis, E.A.: Scaffolding students' reflection for science learning. PhD Thesis. University of California, Berkeley, CA (1998)

Davis, E.A.: Prompting middle school science students for productive reflection: Generic and directed Prompts. The Journal of The Learning Sciences 12(1), 91–142 (2003)

Davis, E.A., Linn, M.: Scaffolding students' knowledge integration: Prompts for reflection in KIE. International Journal of Science Education 22(8), 819–837 (2000)

Gama, C.: Metacognition in Interactive Learning Environments: The Reflection Assistant Model. In: Lester, J.C., Vicari, R.M., Paraguaçu, F. (eds.) ITS 2004. LNCS, vol. 3220, pp. 668–677. Springer, Heidelberg (2004)

Gartner, A., Kohler, M., Riessman, F.: Children teach children: Learning by teaching. Harper & Row, New York (1971)

Ge, X., Land, S.M.: A conceptual framework of scaffolding ill-structured problem solving processes using question prompts and peer interactions. Educational Technology Research and Development 52(2), 5–27 (2004)

Gouli, E., Gogoulou, A., Papanikolaou, K.A., Grigoriadou, M.: An adaptive feedback framework to support reflection, guiding and tutoring. In: Magoulas, G., Chen, S. (eds.) Advances in Web-based Education: Personalized Learning Environments, pp. 178–202. Information Science Publishing, New York (2005)

Graesser, A.C., Person, N.K., Magliano, J.P.: Collaborative dialogue patterns in naturalistic one-to-one Tutoring. Applied Cognitive Psychology 9, 495–522 (1995)

Graesser, A.C., VanLehn, K., Rose, C., Jordan, P., Harter, D.: Intelligent tutoring systems with conversational dialogue. AI Magazine 22, 39–51 (2001)

Hmelo, C., Day, R.: Contextualized questioning to scaffold learning from simulations. Computers & Education 32, 151–164 (1999)

Katz, S., O'Donnell, G., Kay, H.: An approach to analyzing the role and structure of reflective dialogue. International Journal of Artificial Intelligence and Education 11, 320–333 (2000)

Katzlberger, T.: Learning by teaching agents. PhD Thesis, Vanderbilt University (2005)

King, A., Staffieri, A., Adelgais, A.: Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. Journal of Educational Psychology 90(1), 134–152 (1998)

Kornilakis, H., Grigoriadou, M., Papanikolaou, K.A., Gouli, E.: Using WordNet to support interactive concept map construction. In: ICALT (2004)

Kuhn, D., Udell, W.: The development of argument skills. Child Development 74(5), 1245–1260 (2003)

Leelawong, K.: Using the learning-by-teaching paradigm to design intelligent learning environments. PhD Thesis, Vanderbilt University (2005)

Mason, B., Bruning, R.: Providing feedback in computer-based instruction: What the research tells us (2003), `http://dwb.unl.edu/Edit/MB/MasonBruning.html`

Moon, J.: A Handbook of reflective and experiential learning. Routledge, London (2004)

Moore, J., Whitfield, V.F.: Musing: A way to inform and inpire pedagogy throught self-reflection. The Reading Teacher 61(7), 586–588 (2008)

Mory, E.: Feedback research. In: Jonassen, D.H. (ed.) Handbook of Research for Educational Communications and Technology, pp. 919–956. Simon & Schuster Maxmillan, New York (1996)

Ortiz, J.: Reflective practice and student learning in the introductory interpersonal communication course. Technical report, Maricopa Institute for Learning (2006)

Peverly, S.T., Wood, R.: The effects of adjunct questions and feedback on improving the reading comprehension skills of learning-disabled adolescents. Contemporary Educational Psychology 26(1), 25–43 (2001)

Pintrich, P.A., Smith, D.A.F., Garcia, T., McKeachie, W.J.: A manual for the use of the motivated strategies for learning questionnaire. Technical report, The University of Michigan (1993)

Roscoe, R.D., Chi, M.T.H.: The influence of the tutee in learning by peer tutoring. In: Forbus, K., Gentner, D., Regier, T. (eds.) Proceedings of AMCSS, Chicago, pp. 1179–1184 (2004)

Roscoe, R.D., Chi, M.T.H.: Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. Review of Educational Research 77(4), 534–574 (2007)

Rothkopf, E.: Learning from written instructive materials: An exploration of the control of inspection by test-like events. American Educational Research Journal 3, 241–249 (1966)

Sandoval, W.: Conceptual and epistemic aspects of students´ scientific explanations. The Journal of the Learning Sciences 12(1), 5–51 (2003)

Schön, D.A.: Teaching artistry through reflection-in-action. Educating the Reflective Practitioner. Jossey-Bass Publishers, San Francisco (1987)

Schraw, G.: Promoting general metacognitive awareness. Instructional Science 26, 113–125 (1998)

Surbeck, E., Han, E.P., Moyer, J.E.: Assessing reflective responses in journals. Educational Leadership 48(6), 25–27 (1991)

VanLehn, K., Jones, R.M., Chi, M.T.H.: A model of the self- explanation effect. Journal of the Learning Sciences 2(1), 1–60 (1992)

Weinstein, C.E., Meyer, D.K.: Cognitive learning strategies and college teaching. New Directions for Teaching and Learning 45, 15–26 (1991)

Xie, K., Bradshaw, A.C.: Using question prompts to support ill-structured problem solving in online peer collaborations. International Journal of Technology in Teaching and Learning 4(2), 148–165 (2008)

Zimmerman, B.J.: Self-Efficacy: An essential motive to learn. Contemporary Educational Psychology 25(1), 82–91 (2000)

Abbreviations

ANOVA	Analysis of Variance
ILE	Intelligent Learning Environment
ITS	Intelligent Tutoring System
GP	Generic Prompts
MD	Mean Difference
MSLQ	Motivated Strategies for Learning Questionnaire
SP	Specific Prompts