

Gamification of Support for Learning Effective Problem Selection Strategies in Intelligent Tutoring Systems

Yanjin Long

Ph.D. Thesis Proposal, October, 2014

Human Computer Interaction Institute, Carnegie Mellon University

Committee:

Vincent Aleven (Chair) (HCI, CMU)

Kenneth R. Koedinger (HCI & Psychology, CMU)

Jodi Forlizzi (HCI & Design, CMU)

Jesse Schell (Entertainment Technology Center, CMU)

Timothy Nokes-Malach (Psychology, University of Pittsburgh)

Abstract: Many online learning technologies grant students great autonomy and control, which impose high demands for self-regulated learning (SRL) skills. With the fast development of online learning technologies, helping students acquire SRL skills becomes critical to student learning. My proposed work focuses on supporting students' learning of a central SRL skill, making effective problem selection decisions in online learning environments with the aid of certain kinds of learning analytics that are commonly available in many learning technologies. Research has shown that especially younger learners are poor at selecting problems strategically based on their learning status (how much has been learned for different learning units, as generally displayed by the learning analytics). Prior studies mainly targeted on supporting students' problem selection in systems where the scaffolding was in effect, but few studies have tried to teach students the transferrable skills that can be applied in new learning environments. My design centers on teaching two rules of effective problem selection, the mastery rule and the rule for interleaved practice, through design and integration of gamified features in an intelligent tutoring system (ITS), with the goal to foster both learning and enjoyment in the system and transfer of the problem selection skills to new learning environments. I will conduct a classroom experiment to evaluate the effectiveness of the gamified designs on supporting students' learning of the problem selection rules, domain level learning, self-efficacy and enjoyment of learning with the system. The results of my work will shed light on whether and how gamification can be integrated to support learning of transferrable SRL skills in ITSs, and also provide design recommendations for effectively use gamification to support SRL learning, domain level learning and motivation in online learning technologies.

1 Introduction

As an ancient Chinese proverb says, “*Give a man a fish; you have fed him for today. Teach a man to fish; and you have fed him for a lifetime*”. Teaching people how to learn makes fundamental changes in their life-long learning experiences. My research centers on helping students become better self-regulated learners. Theories of Self-Regulated Learning (SRL) take a comprehensive view of the processes involved in academic learning, emphasizing the agency of the learner. For example, Zimmerman (1986) defines SRL as “the degree to which students are metacognitively, motivationally, and behaviorally active participants in their own learning process.” Theories of SRL abound (Pintrich, 2004; Winne & Hadwin, 1998; Zimmerman, 2000); all tend to view learning as repeated cycles with broad phases such as forethought, execution, and evaluation, with learning experiences in one cycle critically influencing those in the next in intricate ways. A number of empirical studies have revealed that the use of SRL processes accounts significantly for the differences in students’ academic performance in different domains of learning (Zimmerman & Martinez-Pons, 1986, Pintrich & De Groot, 1990, Cleary & Zimmerman, 2000), such as reading comprehension, math problem solving, music learning, athletic practice, etc. Consequently, researchers have started to ask whether SRL processes are teachable, so that we can help students to become better self-regulated learners. Many studies have demonstrated that training of SRL processes can enhance students’ academic achievement and motivation (Schunk & Zimmerman, 1998). In a meta-review of intervention studies conducted in primary and secondary schools, Dignath and Büttner (2008) analyzed 84 studies and found the average effect size (standardized mean differences between treatment and control conditions) of the SRL interventions on academic performance to be 0.69.

My proposed work focuses on teaching a critical metacognitive SRL skill, making problem selection decisions while learning with Intelligent Tutoring Systems (ITS). Intelligent Tutoring Systems (ITS) are a type of adaptive online learning environments that support “learning by doing” through scaffolded problem solving practice for individual learners. They have a proven track record of supporting student learning in a wide range of domains (VanLehn, 2011). A common feature of ITS is an Open Learner Model (OLM), which is a type of learning analytics that displays information about students’ learning status (how much/how well they have learned for each type of problems) tracked and assessed by the system’s student model. Specifically, I will design and integrate gamification with an ITS to support students’ learning of actionable rules for making problem selection decisions based on their learning status afforded by the OLM, while enhancing both their enjoyment and domain level learning with the ITS. The interventions also target to facilitate transfer of the learned problem selection rules to be applied in other learning environments that offer learning analytics.

1.1 Theoretical and Practical Benefits of Effective Problem Selection

Selecting problems encompasses two aspects, deciding on which problems to work on and in what order. Research from cognitive science, educational psychology and artificial intelligence in education has substantially established the important role of effective problem selection in enhancing student learning and motivation. Theories such as mastery learning, zone of proximal development (ZPD), and desirable difficulties discuss how to optimally select problems to achieve best learning outcomes in efficient ways. Empirical studies have found that learning with problem selections that are informed by these theories lead to significantly more effective and efficient learning outcomes than learning with randomly selected problems (Metcalf & Kornell, 2005). For example, Corbett (2000) compared the effects on student learning between adaptive problem selection based on cognitive mastery and fixed curriculum in an ITS

for learning Lisp, and found that the implementation of cognitive mastery led to significantly better learning outcomes. In the subject of reading comprehension, Thiede et al. (2003) found that when students strategically chose to re-study the items that they were not good at, they achieved significantly better performance on the post-test. I will discuss two actionable rules for making effective problem selection decisions based on theories and evidence from prior research in section 1.3, and my proposed work will focus on teaching the two specific rules to students in an ITS.

Learning to make effective problem selection decisions also have practical benefits for new generations of learners who have ample opportunities to be exposed to online learning technologies in both formal and informal learning environments. Most of the online learning technologies are self-directed learning environments, offering great autonomy to the students. In recent years, the development of learning analytics in these environments tries to show students their learning status in the systems, for example, in the form of skill meters, progress charts, badges, etc. However, being aware of their learning status does not necessarily mean that the students are able to make the best decisions on what to learn next based on the information. The knowledge for making effective problem selection decisions is grounded in theories and it is highly possible that it needs to be taught to the students explicitly, especially for younger learners. Teaching them the rules in an advanced learning technology, namely, an ITS with an Open Learner Model (OLM), can possibly enable them to transfer the knowledge to be applied in other online learning environments that offer learning analytics. Moreover, even in traditional school learning environments, students receive feedback from teachers on their performance on assignments, quizzes and exams. The rules can also be applied to plan and arrange their learning and review sessions based on such feedback.

In the following section, I first discuss prior research investigating whether students can make effective problem selection decisions in learning, especially with computer-based learning environments (CBLE). Next I review prior interventions that tried to scaffold effective problem selection in CBLE.

1.2 Learner Control over Problem Selection in Computer-based Learning Environments (CBLE)

1.2.1 Can students make effective problem selection decisions, especially in CBLE?

Studies conducted with college students on memory tasks and reading comprehension tasks have shown that the students are able to base problem selection decisions on their own self-assessed learning status (Metcalf, 2009; Thiede et al., 2003). A negative correlation has been found between their Judgment of Learning (how well they have learned about a certain item) and the allocation of study time, which means that adults tend to focus on problems that they judge to be not well learned (Metcalf & Kornell, 2005). However, making such reasonable problem selection decisions has been proven to be difficult for young learners (Metcalf & Kornell, 2003). Studies have found that children tend to make random choices with respect to what they should study (Schneider & Lockl, 2002). Additionally, even with the college students, it is questionable whether they are able to intentionally apply more advanced rules for problem selection such as the interleaving rule (Rule 2) described in section 1.3. Few studies have established that students (even the adult learners) are able to apply advanced problem selection rules other than simply focusing on unlearned topics.

Research on adaptive learning technologies renders further support for this disadvantage of student learning, in that the students are generally found to be unable to make as good problem selection

decisions as the computer algorithms that are developed based on cognitive theories. In a classic experiment in which participants learned vocabulary in a second language, Atkinson (1972) found that the student-select practice condition achieved better learning outcomes than the random-select condition, but was worse than the computer-select condition which adopted a mathematic algorithm that takes into account students' learning status and item difficulty to select practice items for students. In a study with an ITS for SQL, Mitrovic & Martin (2003) found that even the college students with high prior knowledge of SQL were not able to effectively select problems to practice in the tutor. Both the high and low prior knowledge students learned better in the system-select conditions.

In principle, students generally lack the knowledge and skill to make effective problem selection decisions in academic learning environments; although some more mature learners may be capable of using simple rules such as focusing more on the unlearned problems. This may be slightly surprising if we consider how students, even at very young ages, perform in other domains such as sports training. A young child learning to play tennis can probably decide whether s/he needs to spend more time practicing the serve, forehand or backhand after receiving feedback on her/his skills from the coach. It is likely that when it comes to academic learning, students are overwhelmed by the domain level knowledge they need to learn and fail to apply their metacognitive knowledge for selecting problems. Alternatively, they may not be motivated enough to actively plan and make decisions on their learning. Training on problem selection knowledge and skills can coach students in how to make decisions in academic learning, and also affords opportunities for them to perceive the values of applying these rules in helping them learn, boosting their self-efficacy both on the metacognitive and cognitive levels.

1.2.2 Scaffolding for Making Effective Problem Selection Decisions in CBLE

If students are not able to make as good problem selection decisions as the algorithms implemented in the learning technologies, why do we still want them to make the decisions on their own? One answer to the question is to grant them freedom and boost motivation towards learning. Moreover, higher motivation could lead to enhanced learning effects through actively increased learning time spent with the systems, more concentrated processing of the learning materials, etc. However, what may be more important is that the students should be taught to master the advanced problem selection rules that can be applied in different online learning environments even when the scaffolding in the systems are not in effect. Research has been conducted with respect to both aspects of supporting problem selections in CBLE.

A line of work studying how to grant and support students' control in CBLE focuses on scaffolding students' problem selections through visual cues embedded in the designs of the systems. Adaptive navigation support in hypermedia learning environments is among the best examples. Effective designs to support students making decisions on what to attend to next include using headers and site maps, eliminating the links to irrelevant materials, highlighting important topics, etc. Brusilovsky, Sosnovsky, & Shcherbinina (2004) found that with adaptive navigation support in QuizGuide (an adaptive hypermedia learning system), students' participation was increased in the system, as well as their final academic performance. The adaptive navigation support in QuizGuide highlights to the students the important topics and topics that need more practice. Research along this line also generally find that lower prior knowledge students are less likely to benefit from the control offered by the system (Clark & Mayer, 2011), possibly due to the amount of cognitive load caused by the selection processes.

Another line of effort tries to create a shared control over problem selection between students and the system, so as to foster students' motivation while preventing them from making decisions that are

detrimental to learning. For example, Corbalan and colleagues (2008) implemented an adaptive shared control over problem selection in an ITS for health sciences. The tutor adaptively selected a problem type for the students based on factors such as their competence level, task difficulty and rated task load, and provided three problems that only differed with superficial features to let the students select from. This form of shared control led to the same learning outcomes as the full system controlled condition, but did not foster higher interests of using the system.

Efforts have also been made to explicitly teach students rules for effectively selecting problems. In a study which comprises two experiments, students were shown videos of human models who demonstrated how to select problems based on a rule that takes into account past performance and mental effort (Kostons, van Gog, & Paas, 2012). In Experiment 1, the students who watched the video of human models showed significantly better problem selections on the post-tests. However, the results were not replicated in Experiment 2. Mitrovic and Martin (2003) adopted another approach to teach the problem selection strategies through a scaffolding-fading paradigm. In an SQL tutor, the students with low prior knowledge first selected problems with feedback from the system with respect to what the system would have selected for them and why. After they had reached a threshold for learning SQL, the scaffolding was faded, and the students selected their own problems without receiving any feedback. The results indicated that the students were able to better select problems when the scaffolding was in effect, but whether or not they kept making better selection decisions during the fading stage was not measured.

To summarize, there was more work trying to scaffold students' problem selections in computer-based learning environments than explicitly teaching students the rules for effectively selecting the problems. Although some prior work has been successful in scaffolding students' problem selections, it is still an open question whether and how we can support students' learning of problem selection strategies in CBLE while also enhance both their domain level learning and motivation. Moreover, little if any work on SRL and ITS has established improved future SRL skills outside the systems when the scaffolding is not in effect (Aleven, Roll, & Koedinger, 2012; Roll et al., 2014).

1.3 Rules for Making Effective Problem Selection Decisions

In this section I describe two rules for effectively selecting problems that are grounded in theories and supported by empirical evidence. The two rules have been substantially studied and applied to inform instructional design and the design of adaptive problem selection algorithms in learning technologies, but they can also be taught to students to apply in self-regulated learning. My proposed work will focus on teaching students the rules in an ITS, including when and how to apply these rules in problem selections based on their learning status afforded by the OLM.

Table 1. Two Rules for Making Effective Problem Selection Decisions

Rules	When to apply	How to apply
R1: Eliminate mastered problem types from the practice pool	When the problem type currently being worked on is mastered as shown by the OLM	Stop practicing the current problem type
R2: For unmastered problem types, interleave the types based on mastery levels	When there are more than one problem types that are not mastered, and their mastery levels are shown by the OLM	Pick a problem from a different unmastered type from the last practiced problem that has the highest mastery level (if they have the same levels of mastery, just randomly pick one)

Table 1 lists the conditions under which the students should apply the rules and the actionable directions on how to apply them. The two rules act on the problem type level. Each problem type requires a small, common set of knowledge components (Koedinger & Corbett, 2010) or skills. To fully master the skills required for a problem type, students generally need to practice more than one problem in that type. Rule 1 specifies the stop rule for a certain problem type, and also indicates what need to be studied. Rule 2 captures the sequence of practice among the problem types that need to be studied to reach mastery.

1.3.1 Rule for Mastery Learning

Rule 1: Eliminate mastered problem types from the practice pool

Rule 1 draws on the theory of mastery learning and has been applied to develop adaptive problem selection algorithms in several learning technologies, e.g., Cognitive Tutors. Mastery learning emphasizes that the learning targets can be decomposed into small units, and students can proceed to master all the units at their own pace (James, Robert, & Robert, 1990). Atkinson (1972) defines a simple learning model that specifies three transitional states of student learning, *P*, *T* and *U*. In state *P*, the unit is learned and will not be easily interfered by other learning activities, in other words, the unit is “mastered”. State *T* means a learning stage, the target unit is temporarily learned, but is still subject to forgetting. Lastly, the *U* state means the unit is unlearned. Learning and instruction should focus on helping students transition from the unmastered/learning states (*U* and *T*) to the relatively permanently mastered state (*P*), not focusing on the learning units that are already in the mastered state. The extra practice on mastered learning units is considered redundant. A number of studies compared the effects of instructions designed based on mastery learning to fixed curriculums, and found significantly better learning effectiveness and efficiency with the adaptive instructions (Kulik et al., 1990; Corbett, 2000). Learning this rule will help students decide on when to stop practicing a problem type, leading to more efficient learning by preventing them from redundant ineffective practice.

1.3.2 Rule for Interleaved Practice

Rule 2: For unmastered problem types, interleave the types based on mastery levels

Interleaving means that the practice is intermixed rather than blocked, so instead of practicing the different problem types with a blocked sequence like aaabbbccc, practicing with sequences like abcabcabc or abcacab. The superior benefits of interleaved over blocked practice have been demonstrated in various domains, such as learning of motor skills (Hebert, Landin, & Solmon, 1996), vocabulary learning (Cepeda et al., 2006), and math problem solving tasks (Rohrer & Taylor, 2007). Interleaving the problem types for practice is argued to enhance student learning through two mechanisms. Firstly, interleaved practice requires reactivation of the problem solving solution for each problem type more frequently than blocked practice, which could strengthen the long term memory of these solutions and lead to superior learning outcomes (Rau, Aleven, & Rummel, 2013). The frequent reactivation of solutions for different problem types can also help discriminate the solutions under different conditions. For instance, in one study, Tayler and Rohrer (2010) compared the effects of interleaved versus blocked practice on four different types of math problems to find the number of face, corner, edge or angle of a prism. Each problem type requires its own formula to calculate the number. Their results showed that interleaved practice led to significantly better performance on the post-tests where the participants needed to identify which formula to use for which problem type. Secondly, interleaved practice can also contribute to the abstraction or generalization of common solutions across different problem types (Rau et

al., 2013). It enables students to detect the relevance and commonalities between the consecutive problem types, and therefore helps to construct more generalized knowledge for problem solving (De Croock et al., 1998). Rohrer, Dedrick, & Burgess (2014) found that the interleaved practice led to significantly better learning gains compared to blocked practice when the problem types differed with superficial features but essentially required common solutions.

On the other hand, interleaved practice has also been argued as bringing “desirable difficulties” to the students when they encounter new problem types early in the learning process. The desirable difficulties are believed to cause tougher learning process, but produce greater learning outcomes at the end (Taylor & Rohrer, 2010). It is likely that the tougher learning experiences may diminish students’ enjoyment during the learning processes, but eventually foster significant learning gains.

In principle, the proven benefits of interleaved practice inform an effective strategy for making problem selection decisions, which is not generally known by the students. Therefore, it is essential to teach this rule to the students. There are various ways of interleaving the practice for different problem types. Prior research generally compared their own ways of interleaved practice against blocked practice, but few studies had compared the relative effectiveness between different ways of interleaving (Rau et al., 2013).

Therefore, I define my own way of interleaved practice with Rule 2: Students need to interleave the problem types based on the mastery levels of the target unmastered problem types. Specifically, as described in Table 1, each time the students should select a different problem type from the last problem they have practiced, and the selected problem type should have the highest mastery level among the others. In other words, the students should select the problem type that is closest to mastery; arguably, it is also the easiest problem type that is different from the one they just practice. Vygotsky’s concept of Zone of Proximal Development argues that there is a zone that is just above students’ current abilities, and can be reached via scaffolding (Metcalfe & Kornell, 2005). Atkinson’s transitional states model (1972) also pointed out that students should be directed to focus on the learning units that are in the T (learning) state first rather than the U (unlearned) state, as the learning units in the T state is closer to mastery. Both theories advocate that students should practice the problem types from easier to more difficult relative to their abilities. Therefore, Rule 2 is informed by these theories and may lead to better learning outcomes than randomly interleaved practice. In addition, practice with Rule 2 may also help mitigate the frustrations generally experienced by interleaved practice, leading to smoother learning processes.

1.4 Supporting SRL Skills in Gamified Intelligent Tutoring Systems

Games are famous for their affordances to successfully engage players. In recent years, there has been a substantial amount of research trying to integrate game elements into educational technologies, in other words, to gamify the systems for better engagement while maintaining the academic effectiveness of the systems. Gamification examples commonly seen in online learning environments include using badges or stars to reward learners, leveling up to proceed through the topics, presenting a leaderboard for the learning community, etc. As adaptive learning environments, ITS share several features with successful games, such as problem solving oriented environments, feedback on performance, individualized learning paths, etc., which lays foundation for integration of gamification features with ITS. Gamification of ITSs has achieved some success in fostering student learning and motivation. For example, a narrative-centered game-based ITS for science learning has generated the same learning outcomes as a nongame tutor (Rowe, Shores, Mott, & Lester, 2010). Another game-based tutoring environment trying to teach middle school

and high school students mathematical concepts such as Cartesian coordinates, symmetry, and iteration found the integration of game elements fostered higher engagement of learning and facilitated the students to use the system beyond class time (Boyce & Barnes, 2010).

Gamified ITSs have also been used to support students' SRL skills. Games are almost always self-directed environments where the players make decisions on their next moves based on the feedback and constraints afforded by the games. Successful designs of games lead students to work toward a goal under rules enforced by the system, and constantly encourage and motivate them by offering different incentives. These characteristics of game-based environments afford advantages to support SRL training. McNamara, Jackson and Graesser (2010) summarized several game elements that might be integrated with ITSs to foster students' self-regulation, self-efficacy, engagement and interests, including feedback, incentives, task difficulty, control and environments. In their game-based tutoring system for reading comprehension (Jackson & McNamara, 2013), gamified designs of these five features were integrated with the tutor to facilitate self-explanations (which is also an important SRL skill) during the reading process. Results from their experiments revealed that the game-based tutor led to comparable learning outcomes as the nongame version, and fostered higher motivation and enjoyment of using the system.

The prior literature illustrates the potential of using gamification features to support students' SRL in ITS, but substantially more work needs to be done to establish how the different features can be designed to support different SRL skills. In my proposed work, I focus on designing gamification features to support students' learning of problem selection rules, and evaluate the effectiveness of the designs with a classroom experiment.

1.5 Research Questions

In my proposed work, I will investigate:

1. Whether we can teach the two rules for making problem selection decisions with the assistance of an OLM in an ITS through gamified support for student control over problem selection?
2. Will the gamified support for student control over problem selection lead to better domain level learning outcomes and enhanced learning motivation, as compared to a system-controlled ITS?
3. Whether the students will be able to transfer the two rules of problem selection to apply in a new tutoring environment that has an OLM but without the gamified support?
4. Will the transferred problem selection skills with student control lead to better future domain level learning in the new tutoring environment, as compared to a system-controlled ITS?

To address these research questions, my proposed work will consist of two parts: user-centered design of a gamified ITS for supporting the learning of the two problem selection rules; and a classroom experiment that will evaluate the effectiveness of the designs and provide answers to the research questions. My proposed work should contribute to different strands of research. It contributes to educational research by trying to teach students an important SRL skill, making effective problem selection decisions, in online learning technologies. Few studies with ITS have successfully produced transfer effects of scaffolding SRL skills in new learning environments, and gamification is a relatively new approach to support SRL skills. My work should help investigate whether and how gamified features can be designed and integrated with ITS to teach students transferrable SRL skills that can be applied in a new learning environment when the scaffolding is not in effect. My proposed work also contributes to HCI by

producing a set of design recommendations for how to design gamified ITS with user control to foster higher enjoyment and better learning outcomes.

Before I describe my proposed work in more details, I first discuss my prior work that had focused on improving the experience of problem selection with shared student/system control in an ITS for equation solving. The prior interventions also targeted at enhancing students' learning outcomes and enjoyment while the scaffolding for problem selection was in effect, but did not teach students the rules for problem selection and measure the transfer effects in new learning environments, which are the goals of my proposed study. Nevertheless, the prior results shed light on the designs of support for student control over problem selection in ITS, and informed the design of my proposed experiment.

2 Prior Work

My prior work has explored the designs to support two forms of shared student/system control over problem selection in an ITS for equation solving. Experiments 1 and 2 investigated a Broad Shared Control, in which, for each problem practiced by the student, the student decides on the problem type/level and the tutor picks a specific problem from that level. On the other hand, Experiment 3 implemented a Strict Shared Control, in which the tutor decides on the problem level and the students are provided with a list of specific problems to choose from.

In this section, I describe the specific designs of the two forms of shared control, the scaffolding implemented to support students' experiences with the shared control over problem selection, the findings of student-select problem sequences with the Broad Shared Control, and results from the Experiments that highlighted the effectiveness of the equation solving tutor and motivated my proposed work.

2.1 Experiments 1 & 2: Supporting Broad Shared Control over Problem Selection with an Open Learner Model in a Linear Equation Tutor

2.1.1 *Lynnette 1.0* – An Equation Solving Tutor

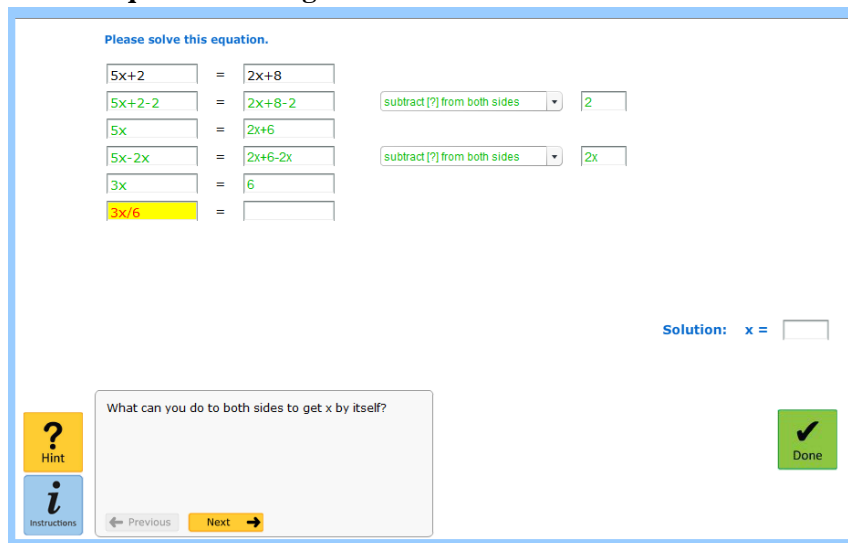


Figure 1. The problem solving interface of *Lynnette 1.0*

I used an ITS for equation solving as the platform. The tutor was first designed and built by Maaiké Waalkens, using the Cognitive Tutor Authoring Tools (CTAT), as an example-tracing tutor (it was not named as *Lynnette* then). The tutor teaches five different types of linear equations (see Table 2). It had been used in a classroom study and was proved to be effective on improving student learning of equations (Waalkens,). The original equation solving tutor has a built-in simple OLM that displays students' learning status with skill bars, but it was not shown to the students in that version. The tutor was then redesigned and used in Experiments 1 and 2, as *Lynnette 1.0*. Figure 1 shows the main interface of *Lynnette 1.0*: in addition to solving the equations, students need to self-explain each main step. The tutor provides step-by-step guidance for each problem with hints and feedback.

Table 2. Five types of equations in *Lynnette 1.0*

Equations	Example	Level/Problem Type
One Step	$x+5 = 7$	Level 1
Two Steps	$2x+1=7$	Level 2
Multiple Steps	$3x+1=x+5$	Level 3
Parentheses	$2(x+1)=8$	Level 4
Parentheses, more difficult	$2(x+1)+1=5$	Level 5

2.2.2 Broad Shared Control over Problem Selection in *Lynnette 1.0*

Table 3 summarizes the steps I went through to redesign *Lynnette 1.0*, including redesign of the built-in OLM (Long & Aleven, 2013a). The overall goals of the design process were to explore how much control we may give to the students over problem selection without impairing the effectiveness of the tutor on equation solving, and how we can redesign the OLM to enhance their experience of using the control. The supported student control in the tutor should lead to better learning outcomes and higher enjoyment of using the tutor. However, the design did not focus on teaching students the rules to make problem selection decisions.

Table 3. An overview of the design process for *Lynnette 1.0*

Design Processes	Research Approaches
1. Paper Prototyping with 3 8 th grade students	HCI/User-centered design
2. High Fidelity Prototyping with 4 6 th and 8 th grade students	HCI/User-centered design
3. Building a working version of <i>Lynnette</i> for an initial classroom evaluation	N/A
4. Classroom Experiment 0 with 98 8 th grade students	Experimental educational research, educational data mining
5. Building <i>Lynnette 1.0</i>	N/A
6. Classroom Experiment 1 with 62 7 th grade students	Experimental educational research, educational data mining
7. Classroom Experiment 2 with 245 7 th and 8 th grade students	Experimental educational research, educational data mining

Broad Shared Control

Results from the user-centered design process suggested that students needed scaffolding to make decisions on what to practice. All participants admitted during the prototyping sessions that they might keep selecting easy problems if they were completely free to select problems by themselves. As a result, I decided to let the tutor lock the mastered levels once it deems the students have reached mastery for all

the skills in a level (as shown in Figure 2 for level 1 and level 2). Students could only select to get problems from unmastered levels, but they were free to select the levels in any order. Once they selected a level, the tutor assigned a new problem to them from that level. All the problems in the same level entail the same set of skills for equation solving (e.g. add/subtract a constant from both sides), and would only be practiced once. Figure 2 shows the newly designed problem selection screen that implements the Broad Shared Control.

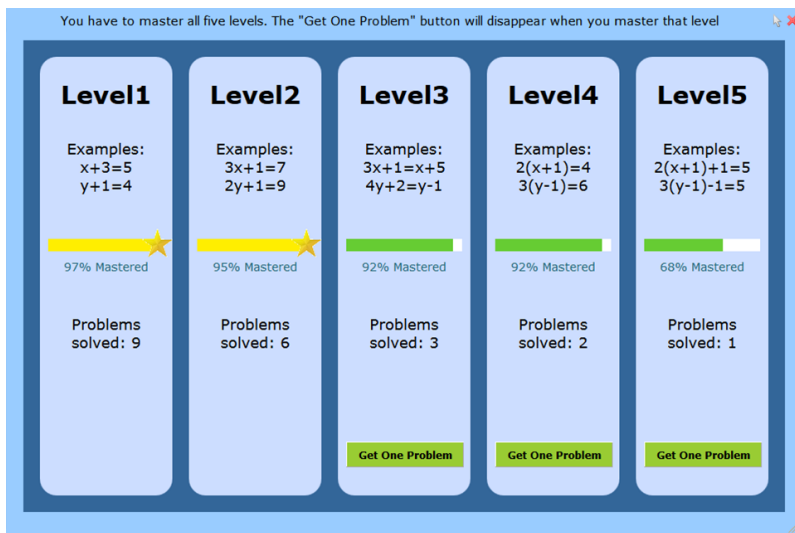


Figure 2. The problem selection screen of *Lynnette 1.0*

The redesigned Open Learner Model (OLM) that supports the Broad Shared Control

Decisions on what to practice in *Lynnette 1.0* was scaffolded by the tutor (by locking the mastered levels), while I redesigned the original OLM to support students’ own decisions on how they would order their practice of the five types of equations. Prior literature has shown that students were not good at making accurate self-assessment on their learning status (Long & Alevan, 2013b), which is arguably the foundation for making effective problem selection decisions. Alternatively, OLM can serve as a substitute for students’ self-assessment, which offers accurate information regarding students’ learning status assessed by the system. Figure 3 shows the new OLM in *Lynnette 1.0*. There were three main features: self-assessment prompts on learning progress, delaying the update of the skill bars until students have answered the self-assessment prompts, and showing overall progress on the problem type level. The OLM was shown on the problem solving interface at the end of each problem to create a short session for self-assessment with feedback (from the update of the skill bars) on their current learning status before the students proceeded to select the next level. The learning status on the problem type level was also displayed on the problem selection screen to assist their decision making. However, no instructions were

provided to the students regarding how to refer to the OLM when they make problem selection decisions.

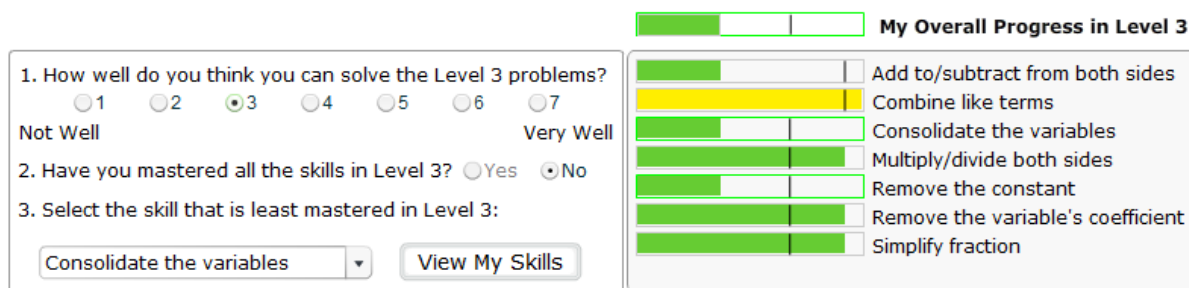


Figure 3. The Open Learner Model (OLM) in *Lynnette 1.0*

As stated earlier, Experiments 1 and 2 did not focus on teaching students the rules for problem selection. Rather, the designs tried to scaffold their control by enriching the use of the OLM, and preventing them from making suboptimal decisions by letting the system lock the mastered levels. The purposes of the classroom experiments were to find out whether the inclusion of the redesigned OLM and the freedom of control over the problem sequences would lead to better domain level learning outcomes and higher enjoyment of using the tutor.

2.1.3 Classroom Experiment 1

The experiment had a 2x2 factorial design, with independent factors OLM (whether or not the redesigned OLM) and PS (whether the students had Broad Shared Control or problem selection was fully system-controlled) (Long & Aleven, 2013c). 62 7th grade students from 3 advanced classes taught by the same teacher at a local public school were randomly assigned to one of four conditions: 1) OLM+PS; 2) OLM+noPS; 3) noOLM+PS; and 4) noOLM+noPS. For the two noPS conditions, there was only one “Get One Problem” button on the problem selection screen, and the tutor assigned problems to the students to reach mastery sequentially from level 1 to level 5. In other words, the two system-controlled conditions would follow a sequentially blocked practice for the five levels, which is the common practice for many ITSs. On the other side, the students in the two PS conditions were free to select whether they would follow a blocked or interleaved practice to reach mastery for the five levels. All participants completed a paper pre-test on their abilities to solve the five types of equations on the first day of the study. They then worked with one of the four different versions of *Lynnette 1.0* in their computer labs for five class periods on five consecutive days. Lastly all students completed a paper post-test to measure their learning gains on solving linear equations.

Overall the students improved significantly from pre to post-tests, affirming the effectiveness of *Lynnette 1.0* in supporting students’ equation solving. A two-way ANOVA with the two factors (OLM and PS) found a significant main effect of OLM on students’ post-test scores, suggesting that the inclusion of the OLM led to better domain level learning outcomes. However, no significant main effect was found for PS. Due to the small size of the sample in this experiment, I decided to run a replication experiment later in the same school year to further investigate the effects of the new designs, as well as to study how students would select their problem sequences with the control.

2.1.4 Classroom Experiment 2

Experiment 2 replicated the procedures in Experiment 1, except that the pre and post-tests were shortened (they were too long for the students in Experiment 1) and a questionnaire on enjoyment was added to the

post-test. 245 7th and 8th grade students from 16 classes (8 advanced classes and 8 mainstream classes) of 3 local public schools participated in Experiment 2. They were taught by 6 teachers.

Effects of the two factors on learning outcomes and enjoyment

This experiment also found, overall, a significant improvement on equation solving from pre to post-tests. ANCOVA (controlling for teachers) analyses found no significant main effects for OLM or PS on students' learning gains from pre- to post-tests. Also, no significant main effects of the two factors were found on enjoyment of using the systems. However, a significant interaction between OLM and PS was found on students' learning gains from pre to post-tests. Pairwise contrasts with Bonferroni Corrections revealed that the OLM+PS condition learned significantly more than the noOLM+PS condition. In other words, when students were allowed to select the levels with the Broad Shared Control, the students who had access to an OLM learned significantly more about equation solving than their counterparts who did not. This finding on domain level learning possibly indicates that when students were granted control over problem selection, the presence of the OLM helped reduce their cognitive load and led to better learning outcomes. Although there were no instructions regarding how to use the information from the OLM to help make problem selection decisions, the students might naturally try to look for such information when they were required to make choices. The absence of the OLM meant that they had to recall and self-assess their learning status, which might be frustrating and consequently diminished their learning. It is also likely that having control over problem selection nudged students to pay more attention to the OLM, which led to deeper reflection at the end of each problem with the self-assessment prompts and update of the skill bars that consequently lead to enhanced learning outcomes. In short, the results indicate that OLM is an important tool for supporting problem selection and enhancing domain level learning in a learning environment where students are granted control over problem selection.

Student-select interleaved versus blocked practice

Experiment 2 also affords opportunities to study how students would freely select their own problem sequences with the Broad Shared Control without any instructions. Of the 245 students in Experiment 2, 120 students were in the two PS conditions. Tutor log data revealed that 61 out of the 120 students (50.8%) selected the same strictly blocked practice from level 1 to level 5 as what was implemented in the two noPS conditions with full system control. This might be partly due to the design of the interface, which positions level 1 to level 5 from left to right sequentially (as shown in Figure 2). It is also likely that the students were more familiar with the sequentially blocked practice that is commonly seen in their textbooks. On the other hand, 59 out of the 120 students (49.2%) selected interleaved sequences with varying ways of interleaving. I measured the degree to which these interleaved sequences differed from the sequentially blocked sequence by counting the number of reverse orders they had as compared against the blocked sequence. The results revealed that the degree of differences were generally small for the 59 students. In other words, they still by and large followed the sequentially blocked sequence. Most of the time, what might have happened was that the student tried to get one or two problems from higher levels, realized they were hard and went back to follow the more intuitive blocked sequence from lower to higher levels. In principle, the students were much more inclined to select an arguably suboptimal practice order, the blocked practice, in *Lynette 1.0*.

I also investigated whether the student-select interleaved sequences led to different effects on student learning and enjoyment as compared to the student-select blocked sequence. ANCOVA (controlling for

teachers) analyses with a factor as whether they selected a blocked or interleaved sequence revealed no significant main effect of this factor on their learning gains from pre to post-tests. This could be largely due to the fact that the interleaved sequences did not differ much from the blocked sequence. However, a significant main effect was found for the self-reported enjoyment on post-test, as the students who selected the blocked sequence reported significantly higher enjoyment. This is consistent with the theories and prior findings about interleaved practice, which argue that it causes tougher and more frustrating learning process for the learners. The students who selected an interleaved sequence might encounter more difficulties when they were practicing the higher level problems early in the learning process, and the frustrating experience led to less enjoyment. This is supported by the log data analyses, as on average, the students who selected interleaved sequences made more errors per step, spent more time on each step, and requested more hints per step. The difference on the number of hints was significant, which was consistent with our informal observations in classrooms. When students got to a new level and encountered difficulties when solving a new type of problems, they relied on the hints. In short, the interleaved sequences selected by the students did not lead to significant difference on domain level learning outcomes. On the other hand, the tougher experiences that resulted from the interleaved sequences still appeared to cause less enjoyment of using the tutor.

To summarize, when students were given control over the sequences of problem types for practice, they were more inclined to select blocked rather than interleaved practice. The fact that no significant difference on learning outcomes was found for whether students practiced with a student-select blocked or interleaved practice does not convincingly conclude that interleaved practice was not more effective, given the student-select interleaved sequence did not differ much from the blocked practice. Therefore, we should still expect to see the superior benefits of interleaved practice on learning outcomes if the students were taught to better interleave their problem practice. There also needs to be interventions that help to mitigate the frustrations and loss of enjoyment caused by the interleaved practice.

2.2 Experiment 3: Gamification of Strict Shared Control Over Problem Selection in a Linear Equation Tutor

Experiment 3 studied whether gamification could be integrated with ITS to boost students' motivation and learning outcomes in a learning environment where they were granted some control over problem selection (Long & Alevan, 2014). This experiment mainly focused on the motivational benefits of gamification, therefore, I restricted the amount of control students could have to ensure the same practice sequence of the problem types, and studied the effects of two gamification features on student enjoyment and learning outcomes with the ITS.

Table 4. New five types of equations in *Lynnette 2.0*

Equations	Example	Level/Problem Type
One Step	$x+5 = 7$	Level 1
Two Steps	$2x+1=7$	Level 2
Multiple Steps 1	$3x+4=x$	Level 3
Multiple Steps 2	$3x+1=x+5$	Level 4
Parentheses	$2(x+1)+1=5$	Level 5

2.2.1 *Lynnette 2.0* – A Tablet based Equation Solving Tutor



Figure 4. The problem solving interface of *Lynnette 2.0* on a Samsung Galaxy Tablet

We (the CTAT team and I) designed and implemented *Lynnette 2.0* as a rule-based Cognitive Tutor that runs on Android tablets, implemented with CTAT. The problem-solving interface (as shown in Figure 4) was redesigned from *Lynnette 1.0* to fit the use on tablet computers. An additional “Undo” function was implemented to allow students to undo their steps. The students could undo the correct steps that have already been accepted by the tutor, in case they wanted to use a different strategy in the midst of solving a problem. Overall, as a rule-based tutor, *Lynnette 2.0* was very flexible in terms of allowing alternative strategies and skipping intermediate steps. Moreover, based on the results from Experiments 1 and 2, the five levels of equations were slightly reorganized. As shown in Table 4, the former Level 3 was separated into two levels (the new Level 3 and Level 4), given it was shown from the data that the students had particular difficulties with equations that have variables on both sides. Former Level 4 and Level 5 were combined into a new Level 5.

2.2.2 Strict Shared Control over Problem Selection in *Lynnette 2.0*

In Experiment 3, the student control was restricted in the sense that they were not allowed to select the levels which decided their practice sequence of the problem types. Rather, with the new Strict Shared Control, they needed to complete the lower levels to unlock the higher levels, and the tutor locked the lower levels once they were mastered (there would always only be one unlocked level on the interface). Therefore, the Strict Shared Control enforced the same full system-controlled problem type sequences as in Experiments 1 and 2, which was the sequentially blocked practice. On the other hand, students were granted another level of control over which specific problems they could select within a problem type. As shown by the right image in Figure 5, the students were presented with a list of problems that they could select from. The problems were supposed to require the same set of skills. Two gamification features were integrated with the Strict Shared Control, re-practice and rewards, which are both very commonly seen in commercial games. The left image in Figure 5 shows the rewards students could earn at the end of each problem, depending on whether they had completed that problem, the number of errors they made and the number of hints requested. The rewards were also displayed next to the problems on the problem selection screen (as shown on the right of Figure 5). Student could earn an extra trophy for perfect

problem solving. Re-practice means that the students were allowed to re-do the problems they had completed before, and the rewards could be updated based on their re-practice performance.

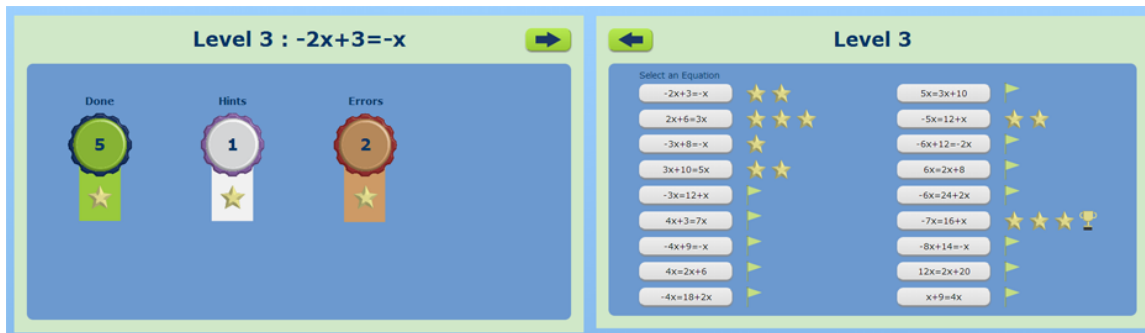


Figure 5. Problem summary screen with rewards (left) and problem selection screen (right) in *Lynnette 2.0*

2.2.3 Classroom Experiment 3

Experiment 3 was conducted to investigate whether the gamified shared control could lead to enhanced enjoyment and learning outcomes. 161 7th and 8th grade students from 15 classes (3 advanced classes and 12 mainstream classes) of 3 local public schools participated in the experiment. They were taught by 5 teachers. Experiment 3 had a 2x2+1 design, with two independent factors as 1) whether or not the students were allowed to re-practice the completed problems, and 2) whether the students were shown performance-based rewards. I also included an ecological comparison condition, which was a standard version of *Lynnette 2.0* that had full-system control (no in-between screens as shown in Figure 5). With the standard tutor, students just kept receiving problems from the system. All five conditions followed the same procedure as Experiments 1 and 2. They all completed a paper pre-test on the first day of the study, learned with one of the five versions of *Lynnette 2.0* for 5 class periods, and took a post-test and an enjoyment questionnaire on the last day of the study.

Overall the five conditions improved significantly on equation solving from pre- to post-tests. However, the results revealed no significant difference on equation solving or self-reported enjoyment between the four gamified *Lynnette 2.0* versions and the standard *Lynnette 2.0*. In other words, the gamified shared control led to comparable learning outcomes as the system-controlled tutor, but did not foster higher enjoyment of using the tutor. Among the four gamified *Lynnette 2.0* versions, the main effects of Re-Practice and Rewards were also not significant for equation solving or enjoyment. However, an interesting significant interaction was found between Re-practice and Rewards on equation solving. When students were allowed to re-practice the completed problems, those who were given rewards did significantly worse on the post-test than their counterparts who did not see the rewards. Further tutor log data analyses revealed that the students who were given rewards revisited significantly more completed problems, and the ratio of revisited problems correlated negatively with their post-test performance. These findings suggest that the performance-based rewards encouraged students to re-practice completed problems to earn more stars and trophies, but the re-practice of previously completed problems was detrimental to learning.

2.3 Conclusions and Design Implications from Experiments 1, 2, and 3

Experiments 1, 2 and 3 produced two *Lynnette* tutors that are both effective on teaching students equation solving skills, which is an important topic in Algebra for middle school students. The three experiments

mainly aimed at supporting the experience of student control through different scaffolding, but did not focus on teaching students the problem selection skills. Therefore, they did not measure whether students learned to make good problem selection decisions on their own, after the scaffolding was removed, which is what my proposed work will address. Nevertheless, the findings from the three experiments motivated my proposed work, and provided tentative design implications.

In general, Experiments 1 and 2 showed that student control over problem selection without guidance and instructions to the students with respect to how to select the problem types may not be effective. The students admitted that they would keep selecting easy problems if they were given full control, which violates Rule 1 (Eliminate the mastered problem types from the practice pool), and they were much more inclined to select blocked practice than interleaved practice, which violates the directions of Rule 2 (For unmastered problem types, interleave the types based on mastery levels). These findings highlight the importance of teaching students the rules that they can both use in the system and transfer to other learning environments. Instead of scaffolding student control by limiting the amount of control they can have, the system should grant them more control and allow them to learn from practicing problem selection with properly designed instruction and feedback, so that they can learn the transferrable rules of problem selection that can be applied even when the scaffolding and feedback is not in effect.

Experiment 2 also highlights OLM as an important tool for supporting problem selection and domain level learning in environments with student control. The information regarding learning status offered by learning analytics like an OLM may help lower the cognitive load for the students when they have to make decisions on problem selection, and also mitigate the detrimental influence on their decisions due to inaccurate self-assessment on their learning status. However, the students should also be taught about how to base their problem selection decisions on the learning status afforded by the OLM.

The results regarding student-select interleaved versus blocked practice in Experiment 2 also highlight the importance of teaching students a good way of interleaving the practice, in order to fully foster the effects of interleaved practice on students' learning outcomes. As argued in section 1.3, Rule 2 specifies an interleaving rule that is grounded in psychological theories, and it should also help to mitigate the frustrations and loss of enjoyment that might be caused by randomly interleaved sequences, by guiding students to practice the problem types that are closer to their zone of proximal development.

Lastly, although gamification in Experiment 3 did not lead to significant difference on students' enjoyment, it illustrated the effectiveness of using rewards as simple as stars and trophies to nudge middle school students' decisions on problem selection. However, the use of rewards in Experiment 3 encouraged a suboptimal strategy (re-practice the completed problems) for problem selection, which impaired student learning. Therefore, in my proposed new designs, the use of rewards needs to be aligned with the instructional goal as to encourage desirable problem selection behaviors. With well aligned designs, gamification has great potential to guide students into desirable behaviors. My proposed work will focus on exerting the effects of gamification in this regard and also keep aiming at fostering higher enjoyment of using the system, which has not yet been established in Experiment 3.

3 Proposed Work

In this section I describe the two parts of my proposed work, 1) the design of the gamified support for learning the two rules for making problem selection decisions based on learning status provided by an OLM; and 2) a classroom experiment that addresses my main research questions.

3.1 Proposed Design of the Gamified Support for Learning of the Problem Selection Rules

The overall goal of the design is to help students learn transferrable skills of making effective problem selection decisions in online learning environments that offer learning analytics, e.g. an OLM. Specifically, the proposed design of the gamified supporting for learning problem selection rules builds on findings from the prior Experiments, and focuses on three features: explicit instruction on the rules, feedback on students’ problem selection decisions, and incentives for desirable problem selection behaviors. Experiments 1 and 2 suggest the inclusion of explicit instruction on how to make problem selection decisions based on learning status is necessary. Also, the use of rewards was found to influence students’ behaviors in Experiment 3, even if these behaviors were detrimental to learning; it is reasonable to assume that with better design, rewards can encourage desirable behaviors.

One salient distinction from my prior work is that the students will be supported to have broader control over problem selection with the proposed new designs of *Lynnette 3.0*. Table 5 shows the amount of student and system control in different versions of *Lynnette*. In *Lynnette 1.0* and *Lynnette 2.0*, Rule 1 (mastery rule) was enforced by the system, but it is critical for students to learn this useful rule and apply it on their own. Experiment 2 also highlighted the importance of learning Rule 2 (the interleaved rule) to strategically sequence the practice of different problem types to exert best learning outcomes and enjoyment. Both rules operate on the problem type level, therefore, *Lynnette 3.0* will grant students control over what problem types to practice and in which order, but not asking them to select specific problems within a problem type. In principle, *Lynnette 3.0* will offer broader control to students with appropriate instruction on the problem selection rules, and provide opportunities for the students to practice these rules with feedback and incentives, aiming to teach them the transferrable skills of making effective problem selection decisions in online learning environments that offer learning analytics.

Table 5. Student and System Control in Different Versions of *Lynnette*

	Which problem type/level to work on	Which specific problem (within a problem type) to work on	When to stop practicing a problem type
<i>Lynnette 1.0</i>	Student Control	System Pick	System locks the mastered levels
<i>Lynnette 2.0</i>	System Control	Student Pick	System locks the mastered levels
<i>Lynnette 3.0</i> (proposed work)	Student Control	System Pick	Student Control

The design of the three gamified features, explicit instruction, feedback and incentives focus on teaching the problem selection rules only. Once students select a type of equations and get a specific problem from the system, they will practice it with the normal tutoring environment, with all the ITS features such as instant correctness feedback and on-demand hints. The OLM in *Lynnette 3.0* will be slightly redesigned to primarily reflect students’ learning status on the problem type level (instead of mainly emphasizing on the

skill level as in *Lynnette 1.0*). The students are going to make decisions on which problem type/level they want to work on next, hence the learning status should be aligned with the decisions they need to make.

3.1.1 Explicit Instruction

The instruction will be embedded at the beginning of the tutor session. It can be a short animated video that explains the two rules that the students are going to use to select problems in the tutor. The instruction will also communicate to the students the overall goal of the learning session as to master all types of equations and earn most points (the point economy will be explained below) in the tutor by correctly applying the rules that are being taught.

3.1.2 Feedback

There will be two kinds of feedback on students' problem selection decisions, tutor points and explanatory feedback.

Point Economy.

A point economy will be created and integrated to serve as one kind of feedback on students' problem selection decisions. Students will be given a certain number of points to begin with, and depending on how well they select the problems based on the two rules, they can earn/lose points. The design of the mechanisms of the point economy will be iteratively tested and improved with user testing with real students to ensure it encourages the desirable behaviors and can be easily understood by the students.

Explanatory Feedback.

Prior research showed that incorporation of explanatory feedback in game-based learning environment could lead to more effective learning outcomes (Clark & Mayer, 2011). One study found that explanatory feedback offered by a learning agent led to better learning than simple correctness feedback, and did not hurt students' enjoyment of using the system (Moreno, 2004). Therefore, my design will also incorporate explanatory feedback messages when students keep making suboptimal decisions. For example, if after reaching mastery for a level (as will be shown by the OLM), the students still keep selecting problems from that level, then after the 3rd problem they select after mastery, the system will pop up a message reminding them that they are not making good choices, and need to switch to other levels to earn more points. In addition to reminding and explaining the reasons to students for their suboptimal behaviors, the feedback messages can also be integrated with incentives to encourage and reward students.

3.1.3 Incentives

The purpose of the incentives is mainly to accompany the feedback (both the points and the explanatory feedback), and help make up for possible frustrations due to the rough problem solving experience that might still be caused by the interleaved practice based on Rule 2, fostering higher enjoyment of using the system. The effects of the feedback may be further strengthened by the presence of incentives, such as trophies. However, the designs of the incentives should not distract the students from the main learning tasks and need to be aligned with the learning goals.

3.1.4 Proposed Design Process

I will start by creating a point economy that aligns with the problem selection rules and learning goals, and create the instructional video that explains the learning goals, two rules and mechanisms of the point economy. Next, I will create prototypes for the OLM and the designs of interfaces that display the

feedback and incentives, and conduct user testing with middle school students. The main focus of the user testing will be to find out 1) whether the students will be able to understand the learning goals, problem selection rules and point economy in the system; 2) how they react to earning and losing points in the system; 3) how they interpret the information from the OLM and explanatory feedback messages; and 4) any feedback on the designs and flows of the interfaces. I plan to have two rounds of prototyping, one with lo-fi prototypes early in the designing stage, and one with high-fi prototypes. Each round of user testing should have at least 8-10 students, including both girls and boys, and also students with different academic abilities.

3.2 Proposed Classroom Experiment

3.2.1 Experimental Design

To address my research questions, the classroom experiment will have three conditions, as shown in Table 7. Condition 1 will be learning with *Lynnette 3.0* that has the new designs of gamified support for learning of Rule 1 and Rule 2. Condition 2 will be a controlled condition that has full system control over problem selection, which enforces Rule 1 and Rule 2. I will also include an ecological control condition (Condition 3), which enforces Rule 1 and sequentially blocked practice of different problem types (the same as the system-controlled conditions in prior Experiments). As argued earlier, such blocked practice is commonly implemented in ITSs and textbooks, where the practice problems for the same problem type are grouped together. The comparison between the two control conditions can help confirm whether the benefits of interleaved over blocked practice still stand for the interleaved practice specified in Rule 2.

Table 6. Conditions of the Proposed Experiment

Conditions	Problem selection rules taught/enforced in the tutor	Which version of <i>Lynnette</i> will be used
Condition 1: student control with gamification (interleaved practice)	R1 + R2	<i>Lynnette 3.0</i>
Condition 2: system control with interleaved practice	R1 + R2	<i>Lynnette 3.0</i> without gamification features
Condition 3: system control with blocked practice	R1 and sequentially blocked practice	<i>Lynnette 3.0</i> without gamification features

The two control conditions will not include any new designs of the gamification, but will both have the redesigned OLM just to show students where they are in the learning process, which is also common practice in ITSs. The design of the experiment helps answer my main research questions:

1. Whether we can teach the two rules for making problem selection decisions with the assistance of an OLM in an ITS through gamified support for student control over problem selection? [Condition 1]
2. Will the gamified support for student control over problem selection lead to better domain level learning outcomes and enhanced learning motivation, as compared to a system-controlled ITS? [Condition 1 vs. Condition 2; Condition 1 vs. Condition 3]
3. Whether the students will be able to transfer the two rules of problem selection to apply in a new tutoring environment that has an OLM but without the gamified support? [Condition 1]

4. Will the transferred problem selection skills with student control lead to better future domain level learning in the new tutoring environment, as compared to a system-controlled ITS? [Condition 1 vs. Condition 2; Condition 1 vs. Condition 3]

Additionally, whether the interleaved practice specified in Rule 2 will lead to superior learning outcomes than blocked practice? [Condition 2 vs. Condition 3]

3.2.2 Participants

The experiment will be conducted in real classrooms. Each condition should have at least 40 students, with a total of at least 120 students. The target group will be 6th – 8th grade students with a limited amount of prior knowledge on equation solving. In prior experiments with *Lynnette*, students generally only had learned about simple one-step equation before the start of the study.

3.2.3 Procedure

Table 8. Overview of the Procedure of the Experiment

Conditions	Pre-Test (10 – 15 minutes)	Learning Session – three types of equations (3 class periods)	Post-Test (15 – 20 minutes)	Transfer Learning Session – new three types of equations (3 class periods)	Post-Test-2 (10 – 15 minutes)
Condition 1 (Gamified Student Control – interleaved)	Items on equation solving abilities and self-efficacy	Learning with <i>Lynnette 3.0</i>	Items on equation solving abilities, self-efficacy and enjoyment	Learning with <i>Lynnette 3.0</i> that has the same student control but no gamified features	Items on equation solving abilities
Condition 2 (System control – interleaved)		Learning with system-controlled <i>Lynnette 3.0</i> that enforces interleaved practice		Learning with system-controlled <i>Lynnette 3.0</i> that enforces interleaved practice	
Condition 3 (System control – blocked)		Learning with system-controlled <i>Lynnette 3.0</i> that enforces blocked practice		Learning with system-controlled <i>Lynnette 3.0</i> that enforces blocked practice	

All three conditions will follow the same procedure. They will all complete a paper pre-test that measures their equation solving abilities and self-efficacy, and then start their first learning session with three problem types for three class periods. Next they will all take a paper post-tests that also measures their equation solving abilities and self-efficacy. A questionnaire for their enjoyment of using the systems will also be included on the post-test. After that, the students will start their second learning session, which will be the transfer learning session, where the two control conditions keep learning with the same systems on three new types of equations. However, Condition 1 (the gamified training condition) will learn with a version of *Lynnette 3.0* that allows them to select among the three new problem types but

without the scaffolding of the gamified features. The OLM will always be presented in all three conditions in both learning sessions. The transfer learning session will also last for three class periods. At the end, all conditions will take a second paper post-test, and it will measure their equation solving abilities. Table 8 shows an overview of the procedure of the experiment.

3.2.4 Measurements

Table 7 displays the target constructs and the corresponding measurements.

Table 7. Measurements of the Proposed Experiment

Construct/abilities to be measured	Assessments	Assessments on Transfer
Problem selection skills for students in Condition 1	<ul style="list-style-type: none"> Tutor log data: the sequences students select in the first learning session 	<ul style="list-style-type: none"> Tutor log data: the problem sequences they select in the transfer learning session
Equation solving abilities for all three conditions	<ul style="list-style-type: none"> Pre-Test Post-Test Post-Test-2 Tutor log data in both learning sessions (process measures, e.g. number of errors and hints per step) 	<ul style="list-style-type: none"> N/A
Self-efficacy for equation solving, algebra and math for all three conditions	<ul style="list-style-type: none"> Questionnaire on Pre-Test Questionnaire on Post-Test 	<ul style="list-style-type: none"> N/A
Enjoyment for all three conditions	<ul style="list-style-type: none"> Questionnaire on Post-Test 	<ul style="list-style-type: none"> N/A

Transfer Learning Session

The purpose of having this transfer learning session is to measure whether the students in the gamified student control condition (Condition 1) will be able to transfer the rules they learn to a new learning environment when the scaffolding is not in effect. It also allows for measuring if the transferred problem selection skills will lead to better learning outcomes in the transfer learning session with student control over problem selection, as compared to the two system-controlled conditions.

Test Items for Equation Solving

The Pre-Test will include the same types of items that the students will learn in the first learning session. Post-Test will include the same types of items as the Pre-Test, as well as items that the students will learn in the transfer learning session (i.e., the Post-Test also serves as a pre-test for Post-Test-2). Post-Test-2 will include items of the same types from the transfer learning session.

Questionnaires for Self-Efficacy and Enjoyment

I will adapt the commonly used questionnaires for self-efficacy and enjoyment to measure students' self-efficacy for equation solving, algebra, and math in general, as well as their enjoyment of using the tutors.

3.2.5 Hypotheses and Data Analyses

Table 9 shows the hypotheses and corresponding data analyses that address the hypotheses. In principle, I expect the gamified support for learning problem selection rules will teach students the transferrable skills of problem selection that they can apply not only when the gamified scaffolding is in effect, but also in a transfer learning environment. Also, I hypothesize that the integration of gamification with student

control will enhance students' domain level learning outcomes, enjoyment of using the system, and self-efficacy on equation solving, Algebra and math in general.

Table 9. Hypotheses and Data Analyses for the Proposed Experiment

Hypotheses	Tests of Hypotheses	Planned Data Analyses
H1: The gamified support for learning problem selection rules will teach students Rule 1 and Rule 2	<ul style="list-style-type: none"> If the tutor log data shows problem sequences that are aligned with Rule 1 and Rule 2 in the first learning session with students in Condition 1 	<ul style="list-style-type: none"> Check the problem sequences being practiced during the first learning session with students in Condition 1
H2: The students will be able to transfer the learned rules of problem selection in a new learning environment that has an OLM	<ul style="list-style-type: none"> If the tutor log data shows problem sequences that are aligned with Rule 1 and Rule 2 in the transfer learning session with students in Condition 1 	<ul style="list-style-type: none"> Check the problem sequences being practiced during the transfer learning session with students in Condition 1
H3: The gamification integrated with student control in <i>Lynnette 3.0</i> will lead to better domain level learning outcomes in the first learning session (and the interleaved practice will lead to better learning outcomes than the blocked practice)	<ul style="list-style-type: none"> If students in Condition 1 achieve the best learning outcomes on Post-Test on the items learned in the first learning session; expected test performance: Condition 1 > Condition 2 > Condition 3 	<ul style="list-style-type: none"> Test for significant difference among the three conditions on performance on the three items practiced in the first learning session on Post-Test, controlling for performance on Pre-Test
H4: The student control over problem selection will lead to better learning outcomes when the gamified scaffolding is not in effect in the transfer learning session (and the interleaved practice will lead to better learning outcomes than the blocked practice)	<ul style="list-style-type: none"> If students in Condition 1 achieve the best learning outcomes on Post-Test-2 on the items learned in the transfer learning session; expected test performance: Condition 1 > Condition 2 > Condition 3 	<ul style="list-style-type: none"> Test for significant difference among the three conditions on performance on the three items practiced in the transfer learning session on Post-Test, controlling for performance on Post-Test for the same problem types
H5: The gamification integrated with student control in <i>Lynnette 3.0</i> will lead to higher enjoyment of using the system during the first learning session	<ul style="list-style-type: none"> If students in Condition 1 rate highest enjoyment on the Post-Test; expected enjoyment ratings: Condition 1 > Condition 2; Condition 1 > Condition 3 	<ul style="list-style-type: none"> Test for significant differences between Condition 1 and Condition 2, as well as Condition 1 and Condition 3 on enjoyment on the Post-Test
H6: The gamification integrated with student control in <i>Lynnette 3.0</i> will lead to improved self-efficacy after the first learning session	<ul style="list-style-type: none"> If students in Condition 1 have increased self-efficacy from Pre-Test to Post-Test If students in Condition 1 rate highest self-efficacy on the Post-Test; expected self-efficacy ratings: Condition 1 > Condition 2; Condition 1 > Condition 3 	<ul style="list-style-type: none"> Test for significant improvement on self-efficacy from Pre-Test to Post-Test for students in Condition 1 Test for significant differences between Condition 1 and Condition 2, as well as Condition 1 and Condition 3 on self-efficacy on the Post-Test

4 Timeline

October 1 st , 2014	Thesis proposal presentation
October	Design and build the lo-fi prototypes Run the first round of user testing
November to December	Build the high-fi prototypes based on first round of user testing Conduct the second round of user testing Finalize the designs based on the second round of user testing
January to February	Implement the new tutors with the CTAT team Schedule time with the schools for the classroom experiment
March	Conduct a Test-a-thon of the new tutors Finish the implementation work based on results from the Test-a-thon Create the test instruments of the experiment
April to Early-May	Run the classroom experiment
Mid-May to Early-June	Data Analyses
Mid-June to Early-August	Write the thesis
Mid-August	Defend

5 Conclusions and Contributions

My proposed work focuses on designing and integrating gamified features in an ITS to teach students transferrable skills of problem selection with the aid of an OLM. The students are expected to transfer the learned skills to apply in other learning environments that offer learning analytics. My work should contribute to different strands of research.

Firstly, my work contributes to the research of supporting Self-Regulated Learning (SRL) in ITS. Few studies with ITSs have successfully produced transfer effects of scaffolding SRL skills in new learning environments when the scaffolding is not in effect. My work focuses on learning of an important SRL skill, that is, making effective problem selection decisions, and aims at fostering learning of the skills not only when the scaffolding is in effect, but also when the students are learning in a new environment.

My work also contributes to the research of integrating gamified features to ITSs. Prior work in this area has found mixed results with respect to the effects of gamification on students' learning and motivation, and more work is demanded to investigate what gamified features are most effective for different domains of learning. Results from my design and classroom experiment will highlight the effective features that can be helpful for supporting SRL skill learning and can be generalized to the design of learning programs for other SRL skills. The results will also help establish the beneficial role of using gamification to support SRL in ITSs, which is a relatively new approach for scaffolding SRL skills.

My contribution to HCI includes a set of design recommendations for gamified learning environments for middle school students that can be produced throughout the design process and classroom experiment. For example, my work will shed light on how to design explanatory feedback in gamified learning environments to effectively encourage desirable behaviors without diminishing the fun of the systems. Also, design implications on how to use the incentives to accompany feedback without distracting students from the main learning tasks can inform the design of learning technologies that include incentives.

Lastly, my work contributes to math education by producing a learning technology that effectively supports the learning of an important Algebra topic, solving linear equations, for middle school students.

The comparison between the interleaved practice specified in Rule 2 and the blocked practice may suggest a specific way of effectively interleaving the problem types for learning equation solving. Data mining work on student models of the ITS can also help reveal if the interleaved practice leads to more generalized knowledge for equation solving. These results should be able to inform instructional design for equation solving and other similar math topics where the different problem types share some common skills (e.g. solving a one-step equation requires a subset of skills for solving a two-step equation).

6 References

- Aleven, V., Roll, I., & Koedinger, K. R. (2012). Progress in assessment and tutoring of lifelong learning skills: An intelligent tutor agent that helps students become better help seekers. In P. J. Durlach & A. M. Lesgold (Eds.), *Adaptive technologies for training and education* (pp. 69-95). New York: Cambridge University Press
- Atkinson, R. C. (1972). Optimizing the learning of a second-language vocabulary. *Journal of Experimental Psychology*, 96(1), 124–129. doi:10.1037/h0033475
- Boyce, A., & Barnes, T. (2010). BeadLoom Game : Using Game Elements to Increase Motivation and Learning. *Foundations of Digital Games (FDG) Conference*, 25–31.
- Brusilovsky, P., Sosnovsky, S., and Shcherbinina, O. QuizGuide: Increasing the Educational Value of Individualized Self-Assessment Quizzes with Adaptive Navigation Support. In: J. Nall and R. Robson (eds.) *Proceedings of World Conference on E-Learning*, E-Learn 2004, AACE, 1806-1813, (2004)
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354e380. <http://dx.doi.org/10.1037/0033-2909.132.3.354>.
- Clark, C. R., Mayer, E. R., (2011) *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*. San Francisco: Jossey-Bass.
- Cleary, T., & Zimmerman, B.J. (2000). Self-regulation differences during athletic practice by experts, nonexperts, and novices. *Journal of Applied Sport Psychology*, 13, 61-82.
- Corbalan, G., Kester, L., Van Merriënboer, J.J.G. Selecting learning tasks: Effects of adaptation and shared control on efficiency and task involvement. *Contemporary Educational Psychology*, 33(4), 733-756, (2008)
- Corbett, A. (2000). Cognitive Mastery Learning in the ACT Programming Tutor. AAI Technical Report SS-00-01
- De Croock, M. B. M., Van Merriënboer, J. J. G., & Paas, F. (1998). High versus low contextual interference in simulation-based training of troubleshooting skills: effects on transfer performance and invested mental effort. *Computers in Human Behavior*, 14(2), 249e267. [http://dx.doi.org/10.1016/S0747-5632\(98\)00005-3](http://dx.doi.org/10.1016/S0747-5632(98)00005-3).

- Dignath, C., & Buettner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3, 231–264
- Hebert, E. P., Landin, D., & Solmon, M. A. (1996). Practice schedule effects on the performance and learning of low- and high-skilled students: an applied study. *Research Quarterly for Exercise and Sport*, 67(1), 52-58
- Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, 105(4), 1036–1049. doi:10.1037/a0032580
- Koedinger, K. R., & Corbett, A. T. (2010). The Knowledge-Learning-Instruction (KLI) framework: Toward bridging the science-practice chasm to enhance robust student learning (CMU-HCII Tech Rep. No. 10-102). Available from <http://reportsarchive.adm.cs.cmu.edu/hcii.html>
- Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction*, 22(2), 121–132. doi:10.1016/j.learninstruc.2011.08.004
- Kulik, C. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of Mastery Learning Programs: A meta-analysis. *Review of Educational Research*, 60, 265-299
- Long, Y., & Aleven, V. (2013a). Active Learners : Redesigning an Intelligent Tutoring System to Support Self-regulated Learning. In *Scaling up Learning for Sustained Impact* (pp. 490–495). Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-40814-4_44
- Long, Y., & Aleven, V. (2013b). Skill Diaries: Improve Student Learning in an Intelligent Tutoring System with Periodic Self-Assessment. In *Artificial Intelligence in Education* (pp. 249–258). Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-39112-5_26
- Long, Y., & Aleven, V. (2013c). Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In *Artificial Intelligence in Education* (pp. 219–228). Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-39112-5_23
- Long, Y., Aleven, V. (2014). Gamification of Joint Student/System Control Over Problem Selection in a Linear Equation Tutor. In S. Trausan-Matu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), *Proceedings of the 12th International Conference on Intelligent Tutoring Systems, ITS 2014* (pp. 378-387). New York: Springer. doi:10.1007/978-3-319-07221-0_47
- McNamara, D. S., Jackson, G. T., & Graesser, A. C. (2010). Intelligent tutoring and games (ITaG). In Y. K. Baek (Ed.), *Gaming for classroom-based learning: Digital role-playing as a motivator of study* (pp. 44– 65). Hershey, PA: IGI Global. doi:10.4018/978-1-61520-713-8.ch003
- Metcalf, J. (2009). Metacognitive Judgments and Control of Study. *Current Directions in Psychological Science*, 18(3), 159–163. doi:10.1111/j.1467-8721.2009.01628.x
- Metcalf, J., & Kornell, N. (2003). The dynamics of learning and allocation of study time to a region of proximal learning. *Journal of Experimental Psychology. General*, 132(4), 530–42. doi:10.1037/0096-3445.132.4.530

- Metcalfe, J., & Kornell, N. (2005). A Region of Proximal Learning model of study time allocation. *Journal of Memory and Language*, 52(4), 463–477. doi:10.1016/j.jml.2004.12.001
- Mitrovic, A., & Martin, B. Scaffolding and Fading Problem Selection in SQL-Tutor. In U. Hoppe, F. Verdejo & J. Kay (Eds.) *Proceedings of the 11th International Conference on Artificial Intelligence in Education*, Berlin: Springer Verlag , 479-481, (2003)
- Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science*, 32, 99–113.
- Pintrich, P.R., De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33–40
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16, 385–407.
- Rau, M. a., Aleven, V., & Rummel, N. (2013). Interleaved practice in multi-dimensional learning tasks: Which dimension should we interleave? *Learning and Instruction*, 23, 98–114. doi:10.1016/j.learninstruc.2012.07.003
- Rohrer, D., Dedrick, R. F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-014-0588-3
- Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, 35(6), 481–498. doi:10.1007/s11251-007-9015-8
- Roll, I., Wiese, E., Long, Y., Aleven, V., & Koedinger, K. R. (2014). Tutoring self- and co-regulation with intelligent tutoring systems to help students acquire better learning skills. In R. Sottolare, A. Graesser, X. Hu, & B. Goldberg (Eds.), *Design Recommendations for Adaptive Intelligent Tutoring Systems: Volume 2 - Adaptive Instructional Strategies*. Orlando, FL: U.S. Army Research Laboratory.
- Rowe, J., Shores, L., Mott, B., & Lester, J. C. (2010). Integrating Learning and Engagement in Narrative-Centered Learning Environments. *Proceedings of the International Conference on Intelligent Tutoring Systems*, 166–177. Retrieved from <http://www.springerlink.com/index/XJ5041R67060V8Q1.pdf>
- Schneider, W.; Lockl, K. The development of metacognitive knowledge in children and adolescents. In: Perfect, T.J.; Schwartz, B.L., (Eds.). *Applied metacognition*. Cambridge, UK: Cambridge University Press; 2002. p. 224-257.
- Schunk, D.H., & Zimmerman, B.J. (Eds.). (1998). *Self regulated learning: From teaching to self-reflective practice*. New York: Guilford Press.
- Taylor, K., & Rohrer, D. (2010). The effects of interleaving practice. *Applied Cognitive Psychology*, 24, 837-848.

- Thiede, K.W., Anderson, M.C.M., Theriault, D. (2003). Accuracy of Metacognitive Monitoring Affects Learning of Texts. *Journal of Educational Psychology*, 95, 66–73
- VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, 46(4), 197–221.
doi:10.1080/00461520.2011.611369
- Winne, P. H., & Hadwin, A. E. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 277-304). Mahwah, NJ: Lawrence Erlbaum Associates.
- Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, 11, 307-313.
- Zimmerman, B.J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing students' use of self-regulated learning strategies. *American Educational Research Journal*, 23, 614-628.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation*, pp. 1–39,. San Diego, CA: Academic Press