On the Influence on Learning of Student Compliance with Prompts Fostering Self-Regulated Learning

Sébastien Lallé
University of British Columbia
2366 Main Mall
Vancouver, BC V6T1Z4, Canada
lalles@cs.ubc.ca

Cristina Conati
University of British Columbia
2366 Main Mall
Vancouver, BC V6T1Z4, Canada
conati@cs.ubc.ca

Roger Azevedo
North Carolina State University
106 Caldwell Hall
Raleigh, NC 27695-8101, USA
razeved@ncsu.edu

Nicholas Mudrick
North Carolina State University
106 Caldwell Hall
Raleigh, NC 27695-8101, USA
nvmudric@ncsu.edu

Michelle Taub
North Carolina State University
106 Caldwell Hall
Raleigh, NC 27695-8101, USA
mtaub@ncsu.edu

ABSTRACT

In this paper, we investigate the relationship between students' learning gains and their compliance with prompts fostering self-regulated learning (SRL) during interaction with MetaTutor, a hypermedia-based intelligent tutoring systems (ITS). When possible, we evaluate compliance from student explicit answers on whether they want to follow the prompts, When such answers are not available, *we mine* several student behaviors related to prompt compliance. These behaviors are derived from students' eye-tracking and interaction data (e.g., time spent on a learning page, number of gaze fixations on that page). Our results reveal that compliance with some, but not all SRL prompts provided by MetaTutor do influence learning. These results contribute to gain a better understanding of how students benefit from SRL prompts, and provides insights on how to further improve their effectiveness. For instance, prompts that do improve learning when followed could be the focus of adaptation designed to foster compliance for those students who would disregard them otherwise. Conversely, prompts that do not improve learning when followed could be improved based on further investigations to understand the reason for their lack of effectiveness

Keywords

Intelligent tutoring systems; Self-regulated learning; Scaffolding; Compliance with prompts; Learning gains; Eye tracking; Linear regression; Hypermedia

1. INTRODUCTION

There is extensive evidence that the effectiveness of Intelligent Tutoring Systems (ITS) is influenced by how well students can regulate their learning, e.g., [13, 22]. Current research has shown that scaffolding self-regulated learning (SRL) strategies such as setting learning goals or assessing progress through the learning content can improve learning outcomes with an ITS, e.g., [1, 10, 22]. In particular, one of the most common approaches to scaffold SRL is to deliver *prompts* designed to guide students in applying specific SRL strategies as needed [22]. Previous work has focused on assessing the general effectiveness of such SRL prompts, for instance by comparing learning outcomes of students working with versions of the same ITS with and without the prompts. (e.g., [1, 19, 21]). Other work has investigated the extent to which students comply with the overall set of prompts generated by an ITS [16, 21]. However, there has been no reported study on the relationship between compliance with *specific* SRL prompts and learning outcomes. In this paper, we aim to fill this gap. Specifically, we explore the impact of student compliance with SRL prompts on learning gains with MetaTutor, an ITS designed to scaffold student SRL processes while learning about topics of the human circulatory system [1].

Our results show that student learning is influenced by compliance with some, but not all, of the SRL prompts delivered by MetaTutor. Overall, we found a positive impact on learning for compliance with prompts fostering learning strategies (revising a summary, reviewing notes), or planning processes (setting new learning goals). On the other hand, we found no impact on learning with prompts related to metacognitive monitoring processes (e.g., prompts to stay on or move away from the current page depending on student performance on a quiz on that page). Having information on the efficacy of each specific prompt in a ITS is important to guide further research on how to improve prompts that do not seem to improve learning when students follow them. Furthermore, prompts that foster learning when followed can become the focus of adaptive interventions designed to improve compliance for those students who would disregard these prompts if left to their own device.

The paper also provides initial insights into prompts design issues that affect how easy it is to evaluate compliance. In MetaTutor, some prompts explicitly asked students whether they wanted to follow the prompt, and then provided suitable affordance to accommodate a positive reply. Compliance with these prompts is easy to assess, but the additional interactions that they require might not always be possible, or might even be intrusive for some students. Other prompts did not require any specific response from the students. Thus, such prompts are in less danger of being intrusive, and provide for a more open-ended interaction. On the other hand, assessing compliance with these prompts is not trivial, because there is no clear definition of what compliance means. For example, one of the MetaTutor prompts asks students to re-read the current MetaTutor content page, but there is no obvious way to map this rather generic suggestion to a specific desired behavior (e.g., spend a specific amount of time on the page, read a specific number of words). We addressed this problem by running linear models to correlate a variety of student behaviors related to prompt compliance with learning. The behaviours we mined are based on both action and eye-tracking data (e.g., time spent on that page, gaze fixations on the content of the page), and our

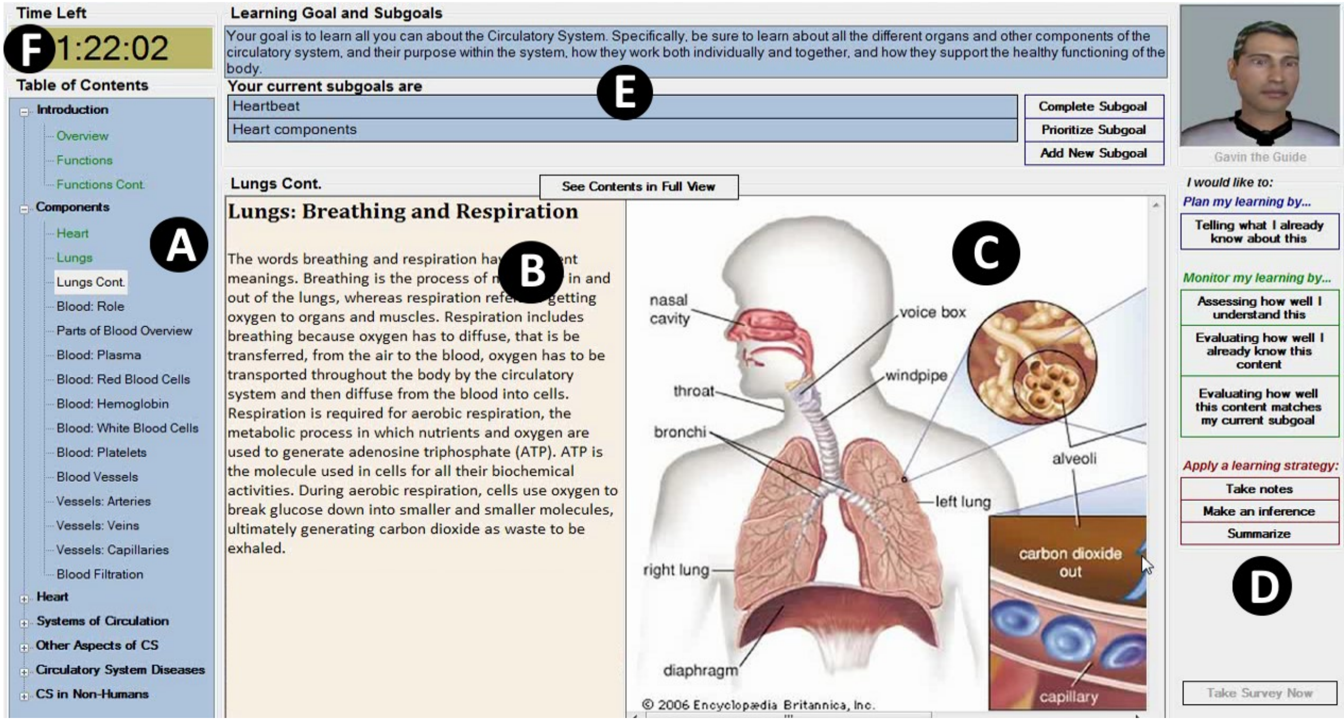


Figure. 1. Screenshot of MetaTutor.

results provide initial evidence that combining these two data sources can help to evaluate compliance. Thus, our findings represent a step toward research on how to evaluate compliance with prompts, both for the type of off line analysis presented in this paper, as well as for the real-time detection of compliance necessary if we want to have ITSs that adaptively help students follow prompts as needed.

The remainder of the paper starts with an overview of related work, followed by a description of MetaTutor and the study that generated the dataset we used for this research. Next, we illustrate how we mined data to evaluate compliance with MetaTutor's prompts, the statistical analysis we conducted, and our results.

2. RELATED WORK

There has been extensive work on assessing the effectiveness of scaffolding designed to support learning with ITSs. Scaffolding can include *prompts* or *hints* (i.e., interventions that guide the student in the right direction), *feedback* (evaluation of students answers, behavior or strategies), or *demonstration* (e.g., worked examples showing expert behavior) [22, 23]. Such scaffolding can be *domain-specific* to support the acquisition of domain-specific knowledge, or targeting domain-independent, meta-cognitive learning processes such as processes for self-regulated learning (SRL). There is extensive evidence that both domain-specific scaffolding (e.g., [3, 12, 18, 20]) and meta-cognitive scaffolding (e.g., [2, 10, 11, 21]) can improve the effectiveness of ITS. For example, domain-specific hints that explain how to solve the current problem step have been shown to improve skill acquisition in a variety of domains such as mathematics [20] and reading [3, 12]. At the meta-cognitive level, Roll et al. [21] tracked suboptimal help-seeking patterns (e.g., overuse of help) to deliver prompts and feedback on how to effectively use help. Prompts and feedback designed to help construct self-explanations during reading [10] or solving scientific problems [11] have been found to positively influence learning. Azevedo et al. [2] showed that SRL prompts and feedback effectively foster efficient use of SRL strategies while learning about biology.

Research has also examined student compliance with SRL prompts in ITS [5, 16]. Kardan and Conati [16] examined the benefit of providing a variety of prompts designed to help students progress within an interactive learning simulation. Overall they found that students largely complied with the prompts and that providing these prompts improved learning gains. However, they did not explore whether and how compliance with specific prompts influence learning outcomes, and which prompts are the most effective. Bouchet et al. [5] adapted the frequency of prompt delivery in MetaTutor based on whether students previously complied with prompts of the same type. However, their analysis uncovered no influence of such adaptive prompting strategy on learning gains. We extend the aforementioned work on prompt compliance by showing how learning gains are impacted by compliance with some, but not all SRL prompts in MetaTutor. Furthermore, whereas previous solely used interaction data to evaluate compliance, we also leverage eye-tracking data when compliance cannot be inferred directly from students' answers or actions (e.g., compliance with the prompts of reading a text further).

Eye-tracking has been used in ITS to model a variety of students traits and behavior, e.g., emotions [14], learning outcomes [15], metacognitive behavior [7], or mind wandering [4]. Eye tracking has also been used to capture students attention to prompts [6, 8] and to pedagogical agents [17]. Conati et al. [6] leveraged gaze data to detect whether students processed domain-specific textual prompts in an educational game for math, and found that reading the prompts more extensively improved game performance. Lallé et al. [17] used gaze data to capture student visual attention to pedagogical agents in MetaTutor, and found that student learning gains are significantly influenced by specific metrics for visual attention (fixation rate, longest fixation). Eye-tracking has also been used to add real-time adaptive prompts to Guru, an agent-based ITS for learning biology [9]. In that work, audible prompts designed to reorient student attention towards the screen were triggered if a student had not looked at the screen for more than 5s while Guru was providing scaffolding. This research showed that this gaze-reactive feedback can improve learning with Guru. In our work, we mine eye-tracking data to evaluate compliance with specific SRL prompts, and examine whether and how compliance with such SRL prompts influences learning gains.

3. METATUTOR

MetaTutor [1] is a hypermedia-based ITS containing multiple pages of content about the circulatory system, as well as mechanisms to help students self-regulating their learning with the assistance of multiple speaking pedagogical agents (PAs). When working with MetaTutor, students are given the overall goal of learning as much as they can about the human circulatory system. The main interface of MetaTutor (see Fig. 1) includes a table of contents (Fig. 1A), the text of the current content page (Fig.1B), a miniature image allowing the student to display a diagram along with the text (Fig. 1C), the current goals and subgoals to learn about (Fig. 1E), a timer indicating how much time remains in the learning session (Fig. 1F), and an SRL palette (Fig. 1D). This palette is designed to scaffold students self-regulatory processes by providing buttons they can select to initiate specific SRL activities (e.g., making a summary, taking a quiz, setting subgoals). Further SRL scaffolding is provided by three PAs in the form of *feedback* on student performance on these SRL activities (e.g., performance on quiz or on the quality of their summaries), as well as *prompts* designed to guide these activities as needed. The PAs deliver these prompts based on student behavior (e.g., time spent on page, number of pages visited).

Specifically, *Pam the Planner* prompts planning processes primarily at the beginning of the learning session by suggesting to add a new subgoal and, if needed, which one to choose (e.g., path of blood flow, heart components). *Mary the Monitor* scaffolds students' metacognitive monitoring processes by making them take quizzes on the target material when they appear to be ready for them. Based on quiz outcomes, Mary prompts students to evaluate the relevance of the current content and subgoal to their knowledge, and suggests how to move through the available material and sub goals accordingly. *Sam the Strategizer* prompts students to apply the learning strategies consisting of summarizing the content studied so far or reviewing notes they have taken on the content[1].

All PAs provide audible assistance through the use of a text-to-speech engine (Nuance). The PAs are visually rendered using Haptek virtual characters, which generate idle movements when the PAs are not speaking (subtle, gradual head and eye movements), as well as lip movements during speech.

4. USER STUDY

The data used for the analysis presented in this paper were collected via a user study designed to gain a general understanding of how students learn with MetaTutor [1]. The study included the collection of a variety of multi-channel trace data (e.g., eye tracking, log files, physiological sensors). In this paper, we focus on using interaction and eye-tracking data to track compliance with the SRL prompts provided by MetaTutor, and study the relationship among compliance with the prompts and learning gains.

Twenty-eight college students participated in the study, which consisted of two sessions conducted on separate days. During the first session, lasting approximately 30-60 minutes, students were administered several questionnaires, including a 30-item pretest to assess their knowledge of the circulatory system. During the second session lasting approximately three hours, students first underwent a calibration phase with the eye tracker (SMI RED 250) as well as a training session on MetaTutor. Each student was then given 90 minutes to interact with the system. Finally, students completed a posttest analogous to the pretest, followed by a series of questionnaires about their experience with MetaTutor.

5. DATA ANALYSIS

5.1 Evaluating Compliance with Prompts

In our analysis we categorize prompts into two types based on how compliance can be evaluated. The first type includes prompts for which compliance can be explicitly assessed from students subsequent responses (*explicit compliance prompts*); the second type includes prompts for which compliance needs to be inferred by mining a variety of behaviors (*inferred compliance prompts*).

Explicit compliance prompts are those that:

- Require students to answer "yes" or "no" (using a dialogue panel that becomes active at the bottom of the display). If students answers yes, the only action they can perform in the MetaTutor interface is the one they agreed upon (e.g., adding a specific subgoal suggested by the agent, making or revising a summary, moving to a previously added subgoal or staying on the current one)[2].
- Require students to take a specific action within a specific time frame (i.e., open the diagram while they are on the current page, and review notes by the end of the learning session).

Table 1 lists the explicit compliance prompts considered in this analysis.

Inferred compliance prompts are those for which the PAs do not force students to provide an explicit answer. Specifically, after the agent utters one of these prompts, the student simply clicks on "continue" in the same dialogue panel, and can either ignore the prompted action, or comply at some point. These prompts (listed in Table 2) include all prompts related to staying on or moving away from the current page, as well as initiating the action of adding a new subgoal.

5.2 Statistical Analysis

Our analysis aims to investigate if and how compliance with MetaTutor's SRL prompts influence learning. The variable we

[1] More details about the design of the agents can be found in [1].

[2] For the "stay on current subgoal" prompt, students are not forced to comply after answering "yes", but we have listed it in this category because student are still required to explicitly answer "yes" or "no" to the PAs as for whether they want to follow the prompt or not.

Table 1. List of explicit compliance prompts provided in MetaTutor (grouped by type of prompted SRL processes).

Prompt label	Description	Prompts for
Suggest subgoal	Recommend possible subgoals to learn about while the students is adding new subgoal.	Planning processes
Moving to next subgoal	Recommend moving on to another subgoal when the student did well on a quiz related to the current subgoal.	Metacognitive monitoring processes
Stay on subgoal	Recommend to learn more about the current subgoal when the student did not do well enough on a quiz related to that subgoal.	
Open diagram	Recommend opening the diagram when it is relevant to the current subgoal.	
Summarize	Recommend making a summary of the current page when the student has spent enough time on that page.	Learning strategies
Revise summary	Recommend revising the summary submitted by the student when there are issues with the summary (e.g., the summary is too long or too short).	
Review notes	Recommend reviewing notes taken on the learning content when approaching from the end of the session.	

Table 2. List of inferred compliance prompts provided in MetaTutor (grouped by type of prompted SRL processes).

Prompt label	Description	Prompts for
Add subgoal	Recommend adding a new subgoal to learn about when a student has no active subgoal.	Planning processes
Move to next page	Recommend moving on to another page when the student did well on a quiz related to the current page.	Metacognitive monitoring processes
Stay on page	Recommend staying on the current page when the student did not well enough on a quiz related to that page.	

adopted to measure learning in our analysis is *proportional learning gain*, defined as:

$$\frac{posttest\ score\ ratio - pretest\ score\ ratio}{1 - pretestscore\ ratio}$$

Table 3 reports statistics for pre- and post-test scores, as well as for the corresponding learning gains.[3]

Table 3. Descriptive statistics for pretest, posttest, and learning gain.

Measures of learning	M	SD	Median
Pretest	18.6	4.2	19
Posttest	21.4	4	21
Proportional learning gain	15.3	50.2	20

We conducted two separate analyses for explicit and inferred compliance prompts, described next.

Explicit compliance prompts. Since compliance is directly observed in the data for explicit compliance prompts (listed in Table 2), we computed a *compliance rate* for each of these prompts as follow:

$$\frac{Number\ of\ prompts\ followed}{Total\ number\ of\ prompts\ delivered}$$

Table 4 shows the compliance rate averaged across students for each of the seven explicit compliance prompts in MetaTutor, and the number of prompts delivered.

Table 4. Descriptive statistics of the number of explicit compliance prompts delivered, as well as on compliance rate.

Prompt	Total number of prompts delivered	Compliance rate *Mean (SD)*
Suggest subgoal	60	.90 (.25)
Move next subgoal	25	.85 (.34)
Stay on subgoal	44	.27 (.37)
Open diagram	77	.21 (.32)
Summarize	105	.32 (.41)
Revise summary	59	.76 (.37)
Review notes	28	.46 (.51)

To investigate the impact of compliance with explicit compliance prompts on learning, we ran a multiple linear regression model with *proportional learning gain* as the dependent variable, as well as the *compliance rate* for each of the seven explicit compliance prompts, and the *total number of prompts received* as the factors. For post-hoc analysis we ran pairwise *t*-test comparisons, and *p*-values were adjusted with the Holm-Bonferroni approach to account for multiple comparisons.

Inferred compliance prompts. As stated above, for inferred compliance prompts (listed in Table 5), students are not forced to explicitly accept or ignore the prompt. This means that compliance with those prompts has to be assessed from student behaviors following the prompts. One approach we considered was to make this assessment binary, as we did for explicit compliance prompts, by establishing thresholds for relevant behaviors. For instance, compliance with the prompt to re-read the current page could be assessed to be true if the student stays on the page for a fixed number of seconds after receiving this prompts. However, it

[3] The increase from pretest to post-test is statistically significant indicating that MetaTutor is overall effective at fostering learning, as further discussed in [1].

is difficult to fix these thresholds in an informed manner, as they may depend on the student (e.g., on a student's readings speed, existing understanding of the page, etc.), and on the object of the prompt (e.g., on the length or difficulty of the page to be re-read). It is also difficult to decide which specific behaviors should be considered for compliance, as several might be relevant (e.g., time spent on a page, specific attention patterns on a page).

Thus, for the subsequent analysis, we avoided committing to specific thresholds and behaviors, and we opted instead for performing regression analyses to try to relate multiple relevant compliance behaviors to learning.

We started by building *data windows* that capture student data from the delivery of each inferred compliance prompt in Table 2, to the following actions:

- "Moving to another page" for the *move to next page* and *stay on page* prompts;
- "Adding a new subgoal" for the *add new subgoal* prompt.

We used these data windows to derive three behavioral measures related to compliance:

- *Window length*, capturing how long students spent before moving on to another page or adding a new subgoal;
- *Number of fixations*[4] made on MetaTutor's learning content (text and diagram), as captured by eye tracking. We use this measure to understand whether students read the page and/or processed the diagram;
- *Number of SRL strategies* initiated by the student by pressing the corresponding buttons in the SRL palette (see Fig. 1 D).

Higher values of these measures (i.e., long windows, high number of fixations on the page and high number of SRL strategies used) are possible indicators that the student is processing the current page, e.g., the student is thinking about or reading the content (as captured by the length of the data window and number of fixations on the page), or using SRL strategies on the current page. Thus, we hypothesized that higher values of these measures could reveal compliance with *stay on page* prompts, whereas lower values could reveal compliance with prompts instructing students to *move on*. Similarly, because prompts to *add a subgoal* requires moving on from the learning content to actually add a subgoal, we expected a short window, a small number of fixations on the page, and a small number of SRL strategies to indicate compliance.

It should be noted that we could have generated other eye-tracking measures, such as fixation duration on the text or the number of transitions from the text to other components of the MetaTutor's interface. However, because valid eye-tracking data were collected for only 16 students out of the 28 who participated in the study, resulting in a rather small dataset, we focused on the most promising behavioral measures that could be related to compliance, as a proof of concept. Table 5 shows the amount of inferred compliance prompts delivered to those 16 students.

[4] Fixation is defined as gaze maintained at one point on the screen for at least 80ms.

Table 5. Number of inferred prompts delivered.

Prompt	Total number of prompts delivered
Add a subgoal	34
Stay on page	117
Move to next page	326

We leveraged the three aforementioned measures of student behavior to investigate if complying with inferred compliance prompts influences learning, and if so, how. Specifically, for each of the three inferred compliance prompts, we ran a multiple linear regression model with *proportional learning gain* as the dependent variable, as well as the *window length*, *number of SRL strategies performed*, and *number of fixations on the learning content* as the factors. As done for explicit compliance prompts, we used pairwise *t*-test comparisons for post-hoc analysis, and all *p*-values were adjusted with the Holm-Bonferroni approach.

6. RESULTS

We describe below the significant[5] effects found in our analysis, first for explicit compliance prompts, and second for inferred compliance prompts.

6.1 Effects for Explicit Compliance Prompts

Our statistical analysis uncovered significant main effects of *compliance rate* for three explicit compliance prompts:

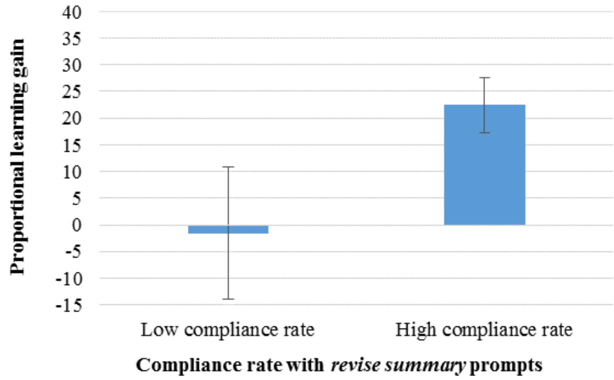
- *Revise summary* ($F_{1,20} = 6.17$, $p=.02$, $\eta_p^2 = .15$), shown Fig. 2a.
- *Review notes* ($F_{1,20} = 7.43$, $p=.013$, $\eta_p^2 = .16$), shown Fig. 2b.
- *Suggest subgoal* ($F_{1,20} = 11.4$, $p=.003$, $\eta_p^2=.27$), shown Fig. 2c.

These three main effects and related pairwise comparisons all reveal that students learned more when they complied more with these prompts than when complying less.

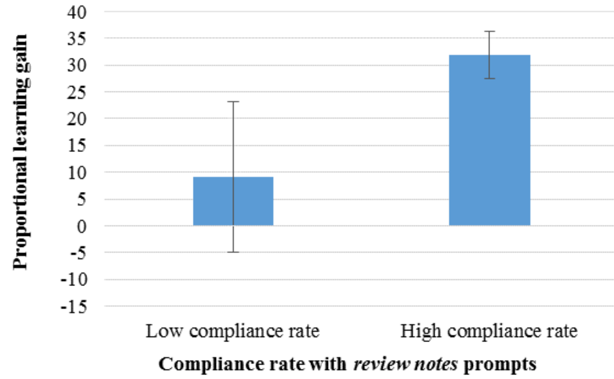
These results for *revise summary* and r*eview notes* are consistent with previous findings showing these learning strategies can be beneficial for learning [17, 22, 24], and extend them by showing that prompting these strategies is effective when students comply with the prompts. Notably, we found a significant effect for prompts to *revise summary*, but not for prompts to *summarize*. This indicates that solely prompting to summarize is not enough to improve learning, and that guiding the students through the process of making a good summary is necessary. Results for *suggest subgoal* indicate that recommending a particular learning subgoal is useful, possibly because it is difficult for students to choose good subgoals by themselves.

These results suggest to examine ways to improve compliance with prompts to *revise summary, review notes* and *suggest subgoal*, since our analysis reveals that not complying with them hinders learning. For instance, MetaTutor could foster compliance with these prompts by explaining how they can help the students, or conversely force the students to follow these prompts.

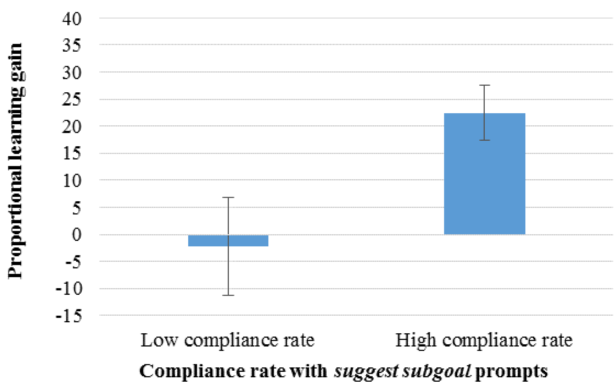
[5] We report statistical significance at the 0.05 level throughout this paper, and effect sizes as small for $\eta_p^2 \geq 0.02$, medium for $\eta_p^2 \geq 0.13$, and large for $\eta_p^2 \geq 0.26$.



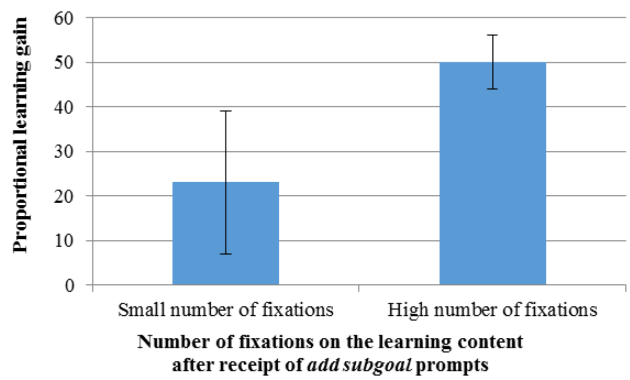
a. Main effect of compliance rate with "revise summary".



b. Main effect of compliance rate with "review notes".



c. Main effect of compliance rate with "suggest subgoal".



d. Main effect of fixation on page after reception of "add subgoal".

Figure 2. Main effects found in this analysis, for explicit compliance prompts (charts a, b, c) and inferred conpliance prompts (chart d). Error bars show 95% confidence interval.

We found no significant effects and small effect sizes (see Appendix A) for the four remaining prompts, namely *summarize*, *stay on subgoal* or *move to next subgoal*, and *open the diagram*. These results indicate it is important to study the effectiveness of SRL prompts individually, to identify those for which compliance does not improve learning. Based on these findings, it is justified to further investigate why complying with these prompts is not beneficial for learning in MetaTutor, and revise the prompts accordingly. For example, it might be due to the nature of the prompts, their timing, their frequency, their wording, and so forth.

6.2 Effects for Inferred Compliance Prompts

We found a main effect of *fixation on learning content* for the "add subgoal" prompts ($F_{1,3} = 13$, $p = .03$, $\eta_p^2 = .29$), shown in Fig. 2d. This effect and related pairwise comparisons reveal that students learned more when they fixate more on the current page than when fixating less. Since students were instructed to add a new subgoal rather than process the current page, this finding suggests that complying with this prompt might not be effective for learning with MetaTutor, possibly because of the timing of this prompt, its frequency or its wording. Although only seven students with valid gaze data received this prompt, the effect size is large, suggesting it is worth conducting further analysis to ascertain whether and why complying with this prompt is not beneficial for learning.

We found no effects and small effect sizes (see Appendix B) for the other inferred compliance prompts, namely *stay on page* and *move to next page*, two prompts related to metacognitive monitoring processes. We cannot make final conclusions on the pedagogical effectiveness on these prompts based on these results, because the dataset is not large and for this reason we did not include in the analysis other features that could indicate compliance (for example other eye-tracking measures such as fixation duration on text or gaze transitions from the text to other components of MetaTutor). However, it should be noted that we also found no effect for the explicit compliance prompts that foster metacognitive monitoring processes (*stay on subgoal*, *move to next subgoal*, and *open the diagram*, see previous section). This lack of effect for all prompts fostering metacognitive monitoring, even when compliance is explicitly assessed, suggests that these prompts are not beneficial for learning with MetaTutor. This could be due to the way these prompts are currently implemented in MetaTutor (e.g., their wording, timing delivery or frequency), or to the nature or the prompts itself. Our results nonetheless justify to run further analysis to ascertain whether (and why) prompts fostering metacognitive monitoring are not effective, and revise them as needed.

7. CONCLUSION

In this research we investigated the relationship between compliance with prompts designed to support the use of self-regulated learning (SRL) processes and learning gains while learning about the human circulatory system with MetaTutor. We identified two approaches to evaluate compliance to MetaTutor's prompts:

(i) Assess compliance from students' subsequent response to the prompts when students are forced to express compliance (e.g., by answering "yes" or "no" to a prompt);

(ii) Run linear models to examine the influence on learning of a variety of student behaviors related to prompt compliance, when compliance is not elicited by MetaTutor. The behaviors we mined are based on both interface and eye-tracking data (e.g., time spent on that page, gaze fixations on the content of the page).

Our results revealed that student learning gains are influenced by compliance with some, but not all SRL prompts provided by MetaTutor. Specifically, we found a positive influence on learning for prompts that foster learning strategies (*revise a summary* and *review notes*) as well as prompts that recommend setting a specific learning subgoal. Based on these findings, it is worth exploring ways to improve compliance with these prompts. In particular, in future research we plan to examine whether forcing students to comply with these prompts or providing detailed explanations on how the prompted SRL strategies can be useful can improve learning.

We found that compliance with the other MetaTutor's prompts studied in this analysis does not improve learning. This finding reveals that assessing compliance to SRL prompts individually is useful to identify prompts that may not be effective at supporting learning. In particular, we found no results for all prompts related to metacognitive monitoring processes (e.g., staying on/moving away from the current page), suggesting to examine further why complying with these prompts do not influence learning with MetaTutor. For example, it could be due to their timing and frequency, their wording, their nature, and so forth.

In this paper we also addressed the challenge of evaluating compliance with rather open-ended prompts for which there is no clear definition of compliance. Specifically we ran a linear regression analysis to relate relevant compliance behaviors to learning. Such behaviors were derived from a combination of student interaction and eye-tracking data after receipt of a prompt (e.g., time spent and amount of gaze fixations on a page can reveal compliance with prompt to read that page). Preliminary results show that such interaction-based and eye-tracking-based measures can help evaluate compliance. In future research, we plan to investigate further behavioral measures relevant to assessing compliance, such as tracking eye gaze patterns on the different components of MetaTutor as well as transitions between those components.

Lastly, we plan to investigate the possibility of detecting in real time compliance with SRL prompts for which we found a positive effect on learning, using eye-tracking and interaction data. Such real-time detection could inform the design of adaptive prompts to foster compliance for those students who might otherwise disregard these prompts. For instance, adaptive prompts could force students to follow them or explain how the prompted SRL processes can improve learning. Evaluating such adaptive prompts fostering SRL processes would provide further insights on how students comply with and benefit from SRL prompts.

8. ACKNOWLEDGMENTS

This publication is based upon work supported by the National Science Foundation under Grant No. DRL-1431552 and the Social Sciences and Humanities Research Council of Canada. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the Social Sciences and Humanities Research Council of Canada.

9. REFERENCES

[1] Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F. and Landis, R. 2013. Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. *International handbook of metacognition and learning technologies*. Springer, 427–449.

[2] Azevedo, R., Martin, S.A., Taub, M., Mudrick, N.V., Millar, G.C. and Grafsgaard, J.F. 2016. Are Pedagogical Agents' External Regulation Effective in Fostering Learning with Intelligent Tutoring Systems? *Proceedings of the 13th International Conference on Intelligent Tutoring Systems* (Zagreb, Croatia, 2016). Springer, 197–207.

[3] Beck, J., Chang, K., Mostow, J. and Corbett, A. 2008. Does help help? Introducing the Bayesian Evaluation and Assessment methodology. *Proceedings on the 9th International Conference on Intelligent Tutoring Systems* (Montréal, QC, Canada, 2008). Springer, 383–394.

[4] Bixler, R. and D'Mello, S. 2015. Automatic Gaze-Based Detection of Mind Wandering with Metacognitive Awareness. *Proceedings of the 23rd International Conference on User Modeling, Adaptation and Personalization* (Dublin, Ireland, 2015). Springer, 31–43.

[5] Bouchet, F., Harley, J.M. and Azevedo, R. 2016. Can Adaptive Pedagogical Agents' Prompting Strategies Improve Students' Learning and Self-Regulation? *Proceedings of the 13th International Conference on Intelligent Tutoring Systems* (Zagreb, Croatia, 2016). Springer, 368–374.

[6] Conati, C., Jaques, N. and Muir, M. 2013. Understanding attention to adaptive hints in educational games: an eye-tracking study. *International Journal of Artificial Intelligence in Education*. 23, 1–4 (2013), 136–161.

[7] Conati, C. and Merten, C. 2007. Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation. *Know.-Based Syst.* 20, 6 (2007), 557–574.

[8] D'Mello, S., Olney, A., Williams, C. and Hays, P. 2012. Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*. 70, 5 (2012), 377–398.

[9] D'Mello, S., Olney, A., Williams, C. and Hays, P. 2012. Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*. 70, 5 (2012), 377–398.

[10] Graesser, A. and McNamara, D. 2010. Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*. 45, 4 (2010), 234–244.

[11] Hausmann, R.G. and Vanlehn, K. 2007. Explaining self-explaining: A contrast between content and generation. *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (Los Angeles, CA, USA, 2007). Springer, 417–424.

[12] Heiner, C., Beck, J. and Mostow, J. 2004. Improving the help selection policy in a Reading Tutor that listens. *Proceedings of the InSTIL/ICALL Symposium on NLP and Speech Technologies in Advanced Language Learning Systems* (Venice, Italy, 2004), 195–198.

[13] Jacobson, M.J. 2008. A design framework for educational hypermedia systems: Theory, research, and learning emerging scientific conceptual perspectives. *Educational technology research and development*. 56, 1 (2008), 5–28.

[14] Jaques, N., Conati, C., Harley, J.M. and Azevedo, R. 2014. Predicting Affect from Gaze Data during Interaction with an Intelligent Tutoring System. *Proceedings of the 12th International Conference on Intelligent Tutoring Systems* (Honolulu, HI, USA, 2014). Springer, 29–38.

[15] Kardan, S. and Conati, C. 2012. Exploring gaze data for determining user learning with an interactive simulation. *Proceedings of the 20th international conference on User Modeling, Adaptation, and Personalization* (Montréal, QC, Canada, 2012). Springer, 126–138.

[16] Kardan, S. and Conati, C. 2015. Providing adaptive support in an interactive simulation for learning: An experimental evaluation. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, South Korea, 2015). ACM, 3671–3680.

[17] Lallé, S., Taub, M., Mudrick, N.V., Conati, C. and Azevedo, R. 2017. The Impact of Student Individual Differences and Visual Attention to Pedagogical Agents during Learning with MetaTutor. *Proceedings of the 18th International Conference on Artificial Intelligence in Education* (Wuhan, China, 2017). Springer (to appear).

[18] McNamara, D.S., Boonthum, C., Levinstein, I.B. and Millis, K. 2007. Evaluating self-explanations in iSTART: Comparing word-based and LSA algorithms. *Handbook of latent semantic analysis*. Psychology Press. 227–241.

[19] Najar, A.S., Mitrovic, A. and McLaren, B.M. 2014. Adaptive Support versus Alternating Worked Examples and Tutored Problems: Which Leads to Better Learning? *Proceedings of the 22nd International Conference on User Modeling, Adaptation, and Personalization* (Aalborg, Denmark, 2014). Springer, 171–182.

[20] Poitras, E.G. and Lajoie, S.P. 2014. Developing an agent-based adaptive system for scaffolding self-regulated inquiry learning in history education. *Educational Technology Research and Development*. 62, 3 (2014), 335–366.

[21] Ritter, S., Anderson, J.R., Koedinger, K.R. and Corbett, A. 2007. Cognitive Tutor: Applied research in mathematics education. *Psychonomic bulletin & review*. 14, 2 (2007), 249–255.

[22] Roll, I., Aleven, V., McLaren, B.M. and Koedinger, K.R. 2011. Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*. 21, 2 (2011), 267–280.

[23] Roll, I., Wiese, E.S., Long, Y., Aleven, V. and Koedinger, K.R. 2014. Tutoring self-and co-regulation with intelligent tutoring systems to help students acquire better learning skills. *Design Recommendations for Intelligent Tutoring Systems, Volume 2*. U.S. Army Research Laboratory. 169–182.

[24] Shute, V.J. 2008. Focus on formative feedback. *Review of educational research*. 78, 1 (2008), 153–189.

[25] Trevors, G., Duffy, M. and Azevedo, R. 2014. Note-taking within MetaTutor: interactions between an intelligent tutoring system and prior knowledge on note-taking and learning. *Educational Technology Research and Development*. 62, 5 (2014), 507–528.

APPENDIX A

All statistical results for explicit compliance prompts (discussed in Section 6.1). Bold indicates a significant effect.

Prompt	F value	*p*-value	Effect size
Suggest subgoal	$\mathbf{F_{1,20} = 11.4}$	***p*=.003**	$\boldsymbol{\eta_p^2=.27}$
Review notes	$\mathbf{F_{1,20} = 7.43}$	***p*=.013**	$\boldsymbol{\eta_p^2=.16}$
Revise summary	$\mathbf{F_{1,20} = 6.17}$	***p*=.02**	$\boldsymbol{\eta_p^2=.15}$
Summarizing	$F_{1,20} = 1.76$	*p*=.20	$\eta_p^2=.06$
Move on subgoal	$F_{1,20} = 0.92$	*p*=.35	$\eta_p^2=.02$
Stay on subgoal	$F_{1,20} = 1.47$	*p*=.24	$\eta_p^2=.01$
Open diagram	$F_{1,20} = 0.71$	*p*=.41	$\eta_p^2=.08$

APPENDIX B

All statistical results for explicit compliance prompts (discussed in Section 6.2). Bold indicates a significant effect.

Prompt	Measure	F value	*p*-value	Effect size
Add subgoal	Window length	$F_{1,3} = .91$	*p* = .41	$\eta_p^2 = .04$
	#fixations on page	$\mathbf{F_{1,3} = 13}$	***p* = .03**	$\boldsymbol{\eta_p^2 = .29}$
	#SRL strategies	$F_{1,3} = .02$	*p* = .90	$\eta_p^2 = .01$
Move on page	Window length	$F_{1,10} = .00$	*p* = .98	$\eta_p^2 = .00$
	#fixations on page	$F_{1,10} = .03$	*p* = .86	$\eta_p^2 = .00$
	#SRL strategies	$F_{1,10} = .40$	*p* = .54	$\eta_p^2 = .01$
Stay on page	Window length	$F_{1,10} = .34$	*p* = .57	$\eta_p^2 = .01$
	#fixations on page	$F_{1,10} = .07$	*p* = .79	$\eta_p^2 = .03$
	#SRL strategies	$F_{1,10} = .004$	*p* = .95	$\eta_p^2 = .02$