

Can Adaptive Pedagogical Agents' Prompting Strategies Improve Students' Learning and Self-Regulation?

François Bouchet^{1(✉)}, Jason M. Harley², and Roger Azevedo³

¹ Sorbonne Universités, UPMC Univ Paris 06, CNRS,
LIP6 UMR 7606, Paris, France
francois.bouchet@lip6.fr

² Department of Educational Psychology,
University of Alberta, Edmonton, Canada
jharleyp@ualberta.ca

³ Department of Psychology, North Carolina State University, Raleigh, USA
razeved@ncsu.edu

Abstract. This study examines whether an ITS that fosters the use of metacognitive strategies can benefit from variations in its prompts based on learners' self-regulatory behaviors. We use log files and questionnaire data from 116 participants who interacted with MetaTutor, an advanced multi-agent learning environment that helps learners to develop their self-regulated learning (SRL) skills, in 3 conditions: one without adaptive prompting (NP), one with fading prompts based on learners' deployment SRL processes (FP), and one where prompts can also increase if learners fail to deploy SRL processes adequately (FQP). Results indicated that an initially more frequent but progressively fading prompting strategy is beneficial to learners' deployment of SRL processes once the scaffolding is faded, and has no negative impact on learners' perception of the system's usefulness. We also found that increasing the frequency of prompting was not sufficient to have a positive impact on the use of SRL processes, when compared to FP. These results provide insights on parameters relevant to prompting adaptation strategies to ensure transfer of metacognitive skills beyond the learning session.

Keywords: Adaptivity · Pedagogical agents · Self-regulated learning · Metacognition · User perception

1 Introduction

Designing intelligent tutoring systems (ITSs) that dynamically adapt to learners' emerging understanding of content and to their use of metacognitive processes has been a major objective for the past decade [1]. Specifically, intelligent systems should provide learners with individualized instruction, feedback and scaffolding during their learning session [2], in a way that fosters the transfer of metacognitive skills beyond that session [3]. It is even more challenging in non-linear open-ended learning environments (OELEs) where no optimal way to navigate through the learning material

exists and where learners' goals may vary [4, 5]. Many critical questions remain unanswered: how often should learners be prompted to perform actions known to foster effective learning? Should prompts vary over time? How can instances where scaffolding should fade be detected?

In this study, we investigated the effect of adaptive prompting on undergraduates' learning and their use of self-regulated learning (SRL) strategies in an OELE with embedded pedagogical agents (PAs). Specifically, we examined how adapting PA prompting impacted learners': (1) use of SRL processes, (2) learning gains, and (3) perception of the system's usefulness. Our associated hypotheses were that: (1) learners should deploy more SRL processes overall, particularly once the scaffolding fades; (2) more efficient SRL should lead to higher learning gains with adaptive prompts; (3) system adaptivity should have a positive effect on learners' evaluation, but the more frequent initial prompting could have a negative effect by making the learners feel overwhelmed.

2 Method

2.1 Participants and Experimental Conditions

One hundred and sixteen undergraduate students ($N = 116$, 17–31 years old, $M = 20.9$ years, $SD = 2.4$; 64.6 % female; 62.9 % Caucasian) from two North American Universities, studying different majors and with various levels of prior knowledge participated in this study. Each participant received \$50 upon completion of the study and was randomly assigned to one of three experimental conditions: (1) *non-adaptive prompt* (NP – $n = 29$), (2) *frequency-based adaptive prompt* (FP – $n = 29$) and (3) *frequency and quality-based adaptive prompt* (FQP – $n = 58$). Participants from adaptive conditions FP and FQP were grouped in some analyses, leading to two samples of identical sizes.

In the NP condition, learners received a moderate but constant amount of prompts from the PAs (on average, 1 per 10 min) to engage in various SRL processes. In the FP condition, learners received more prompts at the beginning of the session (on average, 3.5 per 10 min), but the probability of prompts being triggered decreased after each new prompt and after each self-initiated enactment of an SRL process. In the FQP condition, the same prompt decreasing rules as in FP apply, but the probability of prompts could also increase if: (1) the learner did not comply with a PA's prompt, or (2) a learner's metacognitive judgment was inaccurate (e.g., marked a page as relevant to their active sub-goal when it was not; cf. Table 1 for the list of conditions of success).

2.2 The Testbed System, Experimental Procedure and Data Used

System overview. MetaTutor [6] is an intelligent, hypermedia learning environment in which four embedded PAs help the student learn by prompting them to engage in SRL processes (cf. Table 1). A table of contents gives access to 38 pages (with text and images) on the human circulatory system. The overall learning goal is always visible,

Table 1. Condition of successes associated to the different type of SRL prompts.

Type	Type of PA's prompt	Condition of success
Monitoring	Judgment of Learning (JOL)	Accurate evaluation of what has been learnt
	Feeling of Knowing (FOK)	Accurate evaluation of what is already known
	Content Evaluation (CE)	Accurate evaluation of the relevance of the content relative to the active sub-goal
	Management of Progress Toward Goal (MPTG)	Learner validates their sub-goal in the next 45s
Strategy	Summarization (SUMM)	If learner delays, must be performed later on
	Coordination of Information Sources (COIS)	Image is opened in the next 45s
	Draw image already opened	Digital notepad in the next 45s
	Draw image not opened yet	Learner accepts to open the image

as well as two progress bars associated to the sub-goals chosen at the beginning of the session. A timer displays the time remaining in the learning session. One of the four PAs is always visible. Each PA has a specific role: *Pam the Planner* helps the student to plan their learning sub-goals, *Mary the Monitor* helps in monitoring the learning, *Sam the Strategizer* assists with the deployment of learning strategies and *Gavin the Guide* introduces the system and its questionnaires. The frequency and circumstances under which PAs' prompts are triggered depends on parameters such as the time spent on a page or the relevance of the page to students' current sub-goal. Below the PA, a palette of buttons allows students to self-initiate SRL processes (cf. Table 1), leading to a set of steps very similar to when the prompt comes from a PA: an invitation to perform the process followed by a feedback on its validity (e.g. agreeing the page is relevant to the current learning sub-goal).

Experimental procedure. The experiment involved two different sessions separated by one hour to three days. During the first one (30 to 40 min. long), participants filled and signed a consent form and completed several computer-based self-report questionnaires, a demographics survey and a pre-test on the circulatory system. During the second session (90 min. long), participants used MetaTutor to learn about the circulatory system. Participants had exactly 60 min to interact with the content during which they could initiate SRL processes or do so after a PA's prompt. MetaTutor was paused when participants were watching a video, taking a survey, and during an optional 5 min break half-way through the session. At the end of the session, participants were given a post-test and filled a questionnaire, the Agent Response Inventory (ARI) [7], which included questions on their perception of the quality of PAs' prompts. All participants completed their sessions individually on a desktop computer.

Data coding and scoring. Six variables were extracted from the pre-test and post-test questionnaires (two equivalent 25-item multiple choice tests on the human circulatory system), the ARI questionnaire, as well as from the system log files (cf. Table 2).

Table 2. List of the six variables used for analyses.

Variable name	Description
PropLearnGain	Proportional learning gains (between 0 and 1) using the standard formula: (posttest-pretest)/(1-pretest), for questions relevant to the 2 initial sub-goals and treating negative values as 0
UserAllProc_ [Session first30 last30]	Ratio (per period of 10 min) of all SRL processes initiated by the user during: the whole learning session/the first 30 min of the session/the last 30 min of the session
FBQuality[Mary Sam]	Learner's evaluation of the quality of the PA's feedback (1 to 7)

3 Results

In all of the following statistical analyses, an outlier screening was performed beforehand and outlying scores were replaced by the next most extreme score.

3.1 Effects of Adaptive Prompting on the Use of SRL Processes

Effect on learner-initiated SRL, overall. A one-way ANOVA with prompt condition as the 3-level independent variable and UserAllProc_Session as the dependent variable revealed a significant main effect of condition on learners' self-initiated SRL behaviors, $F(2,113) = 10.17$, $p < .001$, $\eta_p^2 = 0.15$. The application of a more stringent alpha ($p < .01$) and the general robustness of ANOVAs to violations of assumptions supports the legitimacy of this finding and rendered a transformation unnecessary, despite equality of variances not being met (Levene's test). Follow-up post hoc comparisons using a Bonferroni correction revealed that the quantity of SRL behaviors that learners self-initiated were significantly different between the NP ($M = 1.00$; $SD = 0.89$) and FP ($M = 2.04$; $SD = 1.57$), and NP and FQP ($M = 2.02$; $SD = 1.42$) conditions, but not between FP and FQP conditions.

Effect on learner-initiated SRL, over time. A repeated measures ANOVA with prompt condition as the 3-level independent variable, time as an independent 2-level within-subjects variable (first and last 30 min) and learners' self-initiated SRL processes as the dependent variable (i.e. UserAllProc_first30 and UserAllProc_last30) revealed a significant main effect of time on learners' self-initiated behaviors $F(1,113) = 43.95$, $p < .001$, $\eta_p^2 = 0.27$ as well as a significant interaction effect of time and condition on learners' self-initiated behaviors $F(2,113) = 6.65$, $p < .001$, $\eta_p^2 = 0.11$; both results remained significant after the application of a stricter alpha (related to results of Box's Test of Equality of Covariance Matrices). A significant main effect of condition on learners' use of SRL behaviors was found, $F(2,113) = 7.61$, $p < .001$, $\eta_p^2 = 0.12$ (even with a more stringent alpha). An examination of Table 3 reveals that participants consistently engaged in more self-initiated SRL behavior during the second thirty minutes than the first, the most striking changes occurring in FP and FQP.

Table 3. Learner-initiated SRL processes by time and condition.

Variable	NP (<i>n</i> = 58)		FP (<i>n</i> = 29)		FQP (<i>n</i> = 29)		All (<i>n</i> = 116)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
	UserAllProc_first30	0.86	0.91	1.16	1.09	1.62	1.17	1.14
UserAllProc_last30	1.09	1.13	2.12	1.74	2.35	1.84	1.66	1.60

3.2 Effects of Adaptive Prompting on Learning Gains

Table 4 reveals no difference on average between conditions NP and FP&FQP, counter to our hypothesis that adaptive prompting would help with learning. However, when learning gains from NP and FP are compared, it appears that learners in the FP condition had a small benefit over those in the NP, and that FQP did not help.

3.3 Effects of Adaptive Prompting on Perceived System's Usefulness

Two one-way ANOVAs with prompt condition as the 3-level independent variable and FBQualitySam (resp. FBQualityMary) as the dependent variable failed to reveal a significant main effect of condition on learners' self-initiated satisfaction regarding the PAs. Descriptive statistics revealed that participants were most satisfied with Sam in the NP condition ($M = 3.77$, $SD = 1.63$) in comparison to Sam in the FP ($M = 3.13$, $SD = 1.77$) and FQP condition ($M = 3.31$, $SD = 1.79$). In contrast, participants were least satisfied with Mary in the FQP condition ($M = 4.41$, $SD = 1.74$) in comparison to Mary in the NP ($M = 5.00$, $SD = 1.95$) and FP condition ($M = 4.95$, $SD = 1.66$).

4 General Discussion

Adaptive prompting helps learners to self-initiate SRL processes. Learners in (pooled) condition FP&FQP deployed more SRL processes than those in condition NP, as they received more frequent prompting from the system. The number of learner-initiated processes increased over time despite the decrease of agent-initiated prompts, which can be interpreted as a residual and impactful effect of prompting. Our hypothesis was therefore verified. However, taking into account the quality of SRL processes to reduce PAs' prompts did not help: it may be because inefficient self-regulated learners need more than mere (potentially frustrating) reminders to self-regulate.

Adaptive prompting may not directly help to improve learning. We observed no significant differences in learning between conditions NP and FP&FQP, but the expected trend was there when comparing NP and FP. Therefore, it appears that the adaptiveness in FP was going in the right direction, contrary to the one in FQP. Hence our hypothesis was not supported, which could be partially explained by the fact learners might not have been left without scaffolding for long enough for a difference to appear.

Table 4. Learning-related variables in the 3 conditions considered.

Variable	NP		FP		FQP		FP&FQP	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pre-test score (/1)	0.69	0.24	0.73	0.19	0.71	0.21	0.72	0.20
Post-test score (/1)	0.83	0.15	0.85	0.12	0.79	0.19	0.82	0.16
PropLearningGains	0.400	0.378	0.448	0.375	0.339	0.299	0.394	0.341

Initially frequent but fading prompting doesn't degrade perceived system's usefulness. We observed that PAs in FP and FQP were not perceived as less helpful than in NP, despite more frequent prompting at the beginning of the session, which could have been detrimental to learners' willingness to follow PAs' recommendations. Conversely, learners who appreciated PAs' interventions could have found them less useful overall as they were less present towards the end.

Limitations and future work. Although this study benefited from a significantly larger sample size than [8], a larger sample size (with as many participants in FP as in NP) may have led to more significant results. The limited duration of the learning session (1 h) might also have prevented observing internalization and integration of the use of SRL processes by learners once agents' scaffolding was fully gone [9]. Another limitation is the lack of evaluation of the importance of the progressiveness in the scaffolding reduction: another condition with frequent prompting for half a session and no prompting for the second half would be necessary to do so. Finally, we have seen that the adaptation exclusively in terms of frequency of prompting might have been detrimental to learners in condition FQP, and that the quality of the feedback should also be adjusted—confirming its importance [10]. The next steps are to test this approach on other systems, on longer periods of time and to have a finer-grained adaptation.

References

1. Azevedo, R., Aleven, V. (eds.): *International Handbook of Metacognition and Learning Technologies*. Springer, New York (2013)
2. Bannert, M., Mengelkamp, C.: Scaffolding hypermedia learning through metacognitive prompts. In: Azevedo, R., Aleven, V. (eds.) *International Handbook of Metacognition and Learning Technologies*, pp. 171–186. Springer, New York (2013)
3. Roll, I., Aleven, V., McLaren, B.M., Koedinger, K.R.: Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learn. Instr.* **21**, 267–280 (2011)
4. Land, S.M.: Cognitive requirements for learning with open-ended learning environments. *Educ. Technol. Res. Dev.* **48**, 61–78 (2000)
5. Kinnebrew, J.S., Gauch, B.C., Segedy, J.R., Biswas, G.: Studying student use of self-regulated learning tools in an open-ended learning environment. In: Conati, C., Heffernan, N., Mitrovic, A., Verdejo, M. (eds.) *AIED 2015. LNCS*, vol. 9112, pp. 185–194. Springer, Heidelberg (2015)

6. Azevedo, R., et al.: The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 212–221. Springer, Heidelberg (2012)
7. Harley, J.M., Carter, C.K., Papaioannou, N., Bouchet, F., Landis, R.S., Azevedo, R., Karabachian, L.: Examining the potential of personality traits and trait emotions to create emotionally-adaptive intelligent tutoring systems. *User Model. User-Adapt. Interact.* **26**, 1–43 (2016)
8. Bouchet, F., Harley, J.M., Azevedo, R.: Impact of different pedagogical agents' adaptive self-regulated prompting strategies on learning with MetaTutor. In: Lane, H., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS, vol. 7926, pp. 815–819. Springer, Heidelberg (2013)
9. Azevedo, R.: Issues in dealing with sequential and temporal characteristics of self- and socially-regulated learning. *Metacogn. Learn.* **9**, 217–228 (2014)
10. VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., Rosé, C.P.: When are tutorial dialogues more effective than reading? *Cogn. Sci.* **31**, 3–62 (2007)