

Modeling Learners' Cognitive and Affective States to Scaffold SRL in Open-Ended Learning Environments

Anabil Munshi
Institute for Software
Integrated Systems
Vanderbilt University
Nashville, TN, USA
anabil.munshi@vanderbilt.edu

Ramkumar Rajendran
Institute for Software
Integrated Systems
Vanderbilt University
Nashville, TN, USA
ramkumar.rajendran@vanderbilt.edu

Jaclyn Ocumpaugh
Penn Center for Learning Analytics
University of Pennsylvania
Philadelphia, PA, USA
jlocumpaugh@gmail.com

Gautam Biswas
Institute for Software
Integrated Systems
Vanderbilt University
Nashville, TN, USA
gautam.biswas@vanderbilt.edu

Ryan S. Baker
University of Pennsylvania
Philadelphia, PA, USA
ryanshaunbaker@gmail.com

Luc Paquette
University of Illinois at
Urbana-Champaign
Champaign, IL, USA
lpaq@illinois.edu

ABSTRACT

The relationship between learners' cognitive and affective states has become a topic of increased interest, especially because it is an important component of self-regulated learning (SRL) processes. This paper studies sixth grade students' SRL processes as they work in Betty's Brain, an agent-based open-ended learning environment (OELE). In this environment, students learn science topics by building causal models. Our analyses combine observational data on student affect with log files of students' interactions within the OELE. Preliminary analyses show that two relatively infrequent affective states, boredom and delight, show especially marked differences among high and low performing students. Further analysis shows that many of these differences occur after receiving feedback from the virtual agents in the Betty's Brain environment. We discuss the implications of these differences and how they can be used to construct adaptive personalized scaffolds.

ACM Reference format:

A. Munshi, R. Rajendran, J. Ocumpaugh, G. Biswas, R. S. Baker, and L. Paquette. 2018. Modeling Learners' Cognitive and Affective States to Scaffold SRL in Open-Ended Learning Environments. In *Proceedings of ACM UMAP conference, Singapore, July 2018 (UMAP'18)*, 8 pages. <https://doi.org/10.1145/3209219.3209241>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

UMAP '18, July 8–11, 2018, Singapore, Singapore

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5589-6/18/07...\$15.00

<https://doi.org/10.1145/3209219.3209241>

KEYWORDS

Cognitive and affective states, self-regulated learning, open-ended learning environments, affect recognition, adaptive scaffolding

1 INTRODUCTION

Researchers have highlighted the importance of modeling and scaffolding students' self-regulated learning (SRL) processes to improve their learning effectiveness in computer-based learning environments. SRL theories focus on how the regulation of cognitive, metacognitive, affective and motivational processes relate to student learning [1, 23, 28, 30-34]. SRL has been studied in a wide variety of contexts, including those where students are learning complex science phenomena [26]. Analyzing the interactions between students' cognitive and affective processes in open-ended learning environments (OELEs) can help us tailor these environments to provide personalized adaptive scaffolds [2] that make the learning more effective and engaging.

OELEs provide users with complex learning goals and a set of tools that support knowledge acquisition, solution construction, and solution checking [5]. Learners have choices in the way they use these tools to accomplish their learning and problem-solving goals. While these environments encourage exploration, strategic thinking, and developing monitoring skills, the open-ended nature of OELEs can make the process of tracking and interpreting learners' strategic and regulatory behaviors a challenging task. At the same time, it is essential to understand learners' behaviors to provide them with appropriate feedback when they encounter difficulties in achieving their learning goals in these environments. This paper examines students' cognitive and affective states as they work with Betty's Brain [16], an OELE where students learn about scientific phenomena by building causal maps of complex science topics. From our classroom studies, we derive empirical

4 METHODS

Our study included 93 sixth-grade students from four science classes in an urban public school in Nashville, TN, USA. The study was conducted over seven days. It included a pre-test on day 1, training on the system on day 2, working on Betty’s Brain to model the human causes and effects of climate change on days 3-6, and a post-test on day 7. The map that the students were expected to construct to teach Betty include a total of 25 causal links. The paper-based pre- and post-tests were identical and tested students’ knowledge on the science domain concepts and understanding of causal relations using a combination of multiple-choice (MC) and short-answer (SA) questions. Students could receive maximum scores of 7 points for the former and 9 points for the later, for a total maximum score of 16.

4.1 Affect Data

We analyse students’ logged actions (as described in Table 1) and their observed affective states while they work with Betty’s Brain. Data on affective states were collected using the Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP; [18]). In this protocol, trained observers (inter-rater reliability: *Cohen’s kappa* > 0.6) code students individually, in a pre-determined order using a momentary time sampling technique. Observers make holistic judgments, recording both affective states (i.e. *bored, confused, delight, engaged concentration, frustrated, other*) and behavioral categories (i.e. *on-task, off-task, on-task conversation, other*). Observers are trained to record the first thing they see, but have up to 20 seconds to make a decision, after which they are supposed to mark the observation as “other.” (For the purposes of this study, only affective categories are considered.)

Observations are recorded using an Android app [17], which enforces the protocol and synchronizes with an internet time server allowing each observation to be aligned with log files of student activities within the system. The app also helps to speed the recording process. In this study, an expert coder (the 3rd author) trained a second observer during the first two days. (None of the trainee’s data was used from this period.) After the trainee was certified, each observer coded independently. A total of 4233 affect observations was generated across the 2 classes over 4 days.

4.2 Learning Outcomes

This study considers both system internal learning measures (the final map score) and system external summative learning measures (pre- and post-test scores). The former was calculated as the difference between the number of correct and incorrect causal relations on each student’s concept map at the end of the study. The latter were used to generate a normalized learning gain score (This normalized learning gain, calculated as $\frac{Post-Pre}{Max-Pre}$, accounted for the effect of pre-test score on post-test score).

For our analyses, we divided students ($n = 93$) into “Hi” (*high performers*) and “Lo” (*low performers*) based on the median value of their final map scores (*median* = 11). Students with a map score greater than 11 were labeled “Hi” ($n = 47$) and those with a map score less than 10 were labeled “Lo” ($n = 42$). Data for students

at the median value ($n = 4$) was discarded to maintain the distinction between the two groups. There was a 16.5-point difference between the average Hi group (*mean* = 19.5, *sd* = 9.4), and Lo group (*mean* = 3, *sd* = 9.5) map scores. The difference in the normalized learning gains score was also substantial, with a mean of 0.36 (*sd* = 0.27) for students in the Hi group and 0.17 (*sd* = 0.27) for students in the Lo group.

4.3 Action Log Files

Students’ cognitive activities in Betty’s Brain were logged and classified into five primary action categories based on the Betty’s Brain task model [14]. As shown in Table 1, they are: *Read, Edit, Assess, Note* and *Conv (Conversation)*.

The combination of these actions illustrates the coordination of students’ learning and problem-solving activities with corresponding skills and cognitive processes.

Table 1: Cognitive Activities in Betty’s Brain

Primary Action	Description
Read	Learner reads resource pages or teacher’s guide (<i>information acquisition</i>)
Edit	Learner edits the concept map to teach Betty (<i>solution construction</i>)
Assess	Learner has Betty answer a causal question or take a quiz to evaluate the state of the map (<i>solution assessment</i>)
Note	Learner takes/edits notes used to track acquired information
Conv	Learner converses with one of the agents (Betty or Mr. Davis)

These cognitive activities (primary student actions) serve as the primary unit of analyses, but more specific actions (e.g., requesting/receiving certain kinds of hints, moving elements on the map, looking at quiz results, etc.), are also considered.

4.4 Analyses

We conducted several analyses. We studied performance by computing the significance of students’ normalized pre-to-post learning gains. We also computed correlations between student map scores and their normalized learning gains to establish the relationship between students’ performance in the system and their overall learnings.

Next, we compared the time spent by the Hi and Lo students on each of the 5 primary activities in the Betty’s Brain system. To compare the differences in the affective states between the two groups, we computed the relative number of times each of the affective states were recorded for each group and determined how the observed affective states were distributed across each of the 5 primary actions for the Hi versus Lo groups.

Finally, we used the system log files to study situations where Hi and Lo students showed marked differences in delight and boredom to understand the interactions between students’ cognitive

and affective states. We then compared students' performance after observations of four affective states (boredom, delight, confusion and frustration) to see how each affective state was associated with developments in student performance and learning.

5 RESULTS

5.1 Pre-Post Learning Gains and Performance in Betty's Brain

Learners' pre-test scores, post-tests scores and learning gains for multiple choice (MC) and short answer (SA) questions are reported in Table 2. A one-way ANOVA across all students indicates significant gains from pre- to post-test ($p < 0.00001$) with medium to large effect sizes. We also calculated the Pearson correlation coefficient between learning gains and learners' final map scores in Betty's Brain. The results show a moderate ($r = 0.2$) but statistically significant ($p < 0.05$, $t = 1.95$) correlation between learners' map scores and learning gains, indicating that the OELE provided moderate support to students' learning of the science content. (Some of the learning may be attributed to just reading the science resources).

Table 2: Pre-to-post learning gains - all students ($n = 93$)

Question Type	Pre-test score mean (sd)	Post-test score mean (sd)	Normalized Learning Gains $\frac{Post-Pre}{Max-Pre}$ mean (sd)	1-way ANOVA F -ratio (p -value)	Effect Size Cohen's d
MC (Max=7)	4.08 (1.42)	5.48 (1.31)	0.45 (0.52)	45.11 (< 0.0001)	1.02
SA (Max=9)	2.28 (1.79)	3.5 (2.29)	0.16 (0.31)	15.24 (< 0.00015)	0.59
Overall (Max=16)	6.35 (2.93)	8.98 (3.25)	0.26 (0.28)	30.64 (< 0.0001)	0.85

5.2 Distribution of Cognitive Activities

Students' cognitive activities were categorized into five primary actions based on the Betty's Brain task model [11]. Overall, each Hi performer showed a higher count of actions (437 more actions on an average) than Lo performers. We inferred the relative time spent doing each primary action from system logs. Our analysis (Table 3) shows minimal differences in the percentage of total time duration for each primary action across Hi and Lo performers. However, students in the Hi group spent a significantly higher percentage of their total time doing *Assess* actions (such as *evaluating the map*) compared to Lo groups (data tested for normality using Shapiro-Wilk test, 1-way ANOVA $p < 0.05$, corrected post-hoc using Benjamini-Hochberg procedure). This finding, that self-initiated assessment appears to be significantly higher for Hi performers, may relate to psychological memory research by Karpicke et al. [12], who demonstrate that repeated testing is more

effective than repeated studying for a later recall of learning content.

Table 3: Percentage of total duration spent in each primary action (Hi vs Lo)

Category	Read %	Edit %	Assess %	Note %	Conv %
Hi (n=47)	26.6	48.4	12.3	1.4	11.3
Lo (n=42)	29.6	49.5	6.5	2.8	11.6

5.3 Observations of Affective States

We analyzed the data collected on students' affective states by performance (Hi vs Lo). The total observations (for both Hi and Lo groups) for each affective state were 91 cases of *boredom*, 238 cases of *confusion*, 118 cases of *delight*, 3506 cases of *engaged concentration*, 207 cases of *frustration* and 73 cases of *other*. Table 4 presents the proportion of each affective state as a percentage of total affect observations for each group. Significant differences between high and low performers were observed for two affective states: (1) Delight was significantly higher for the Hi group; and (2) boredom was significantly higher for the Lo group ($p = 0.05$ for both delight and boredom, adjusted post-hoc for Type 1 error rate using Benjamini-Hochberg procedure).

Table 4: Affective states as percentage of total affect observations (Hi vs Lo)

Category	Boredom %	Confusion %	Delight %	Engaged %	Frustration %
Hi (n=47)	0.83	5.59	3.39	84.71	5.48
Lo (n=42)	4.31	5.91	1.97	83.63	4.18

5.4 Distribution of Observed Affective States Across Primary Actions

We next analyze how affective states are distributed across the five primary actions (Read, Edit, Assess, Note and Conversation) for the high versus low performers, to understand if these differences are more prominent in certain action contexts. We focus our attention specifically on the *boredom* and *delight* scenarios (since these two affective states showed significant differences between Hi and Lo groups). The results are presented in Table 5.

Table 5: Distribution of affective states across the 5 primary actions (Hi vs Lo)

Values denote affect instances observed during each action

Action	Category	Boredom	Confusion	Delight	Engagement	Frustration
Read	Hi	7	27	10	329	26
	Lo	10	23	4	200	11
Edit	Hi	6	52	20	888	42
	Lo	24	33	15	618	27
Assess	Hi	2	33	25	345	31
	Lo	6	14	3	156	11
Note	Hi	1	1	2	33	2
	Lo	1	3	3	35	2
Conv	Hi	5	29	29	553	38
	Lo	29	23	7	342	17

Looking at *boredom* and *delight*, we observe 5 cases (shaded in Table 5) where large differences were observed between Hi and Lo groups for a specific action. Boredom was substantially more likely among Lo learners during Edit actions (24 vs. 6 instances) and Conversation actions (29 vs. 5 instances). Conversely, delight was substantially more common for the Hi group than for the Lo group during Read actions (10 vs. 4 instances) Assess actions (25 vs. 3 instances), and Conversation actions (29 vs. 7 instances).

5.5 Temporal Relations between Students' Cognition and Affect States

To better understand the interactions between affective states and cognitive activities, we considered more specific actions rather than just the five primary ones discussed above. We examined the temporal ordering of these actions, to determine the common antecedents, if any, to boredom and delight. The results of this case-by-case analysis are presented in Table 6, which shows differences between the Hi and Low groups.

Table 6: Temporal antecedents for substantial differences in boredom and delight

Case 1: Instances of Boredom during Conv		
Antecedent	Hi (n=5)	Lo (n=29)
Post-quiz <i>indirect</i> hint	0	5
Agent progress report (bad)	1	8
Agent response to query	2	4
Long MapView or Read	2	11
Off-task	0	1
Case 2: Instances of Boredom during Edit		
Antecedent	Hi (n=6)	Lo (n=24)
Hint→MapView	0	7
Long period of MapView & MapElementsMove	4	8
Agent progress report (bad)	1	3

Read multipletimes→MapView	1	1
ConceptEdit multiple times	0	1
QuizProvenWrong→MapView	0	3
Query agent → MapView	0	1

Case 3: Instances of Delight during Assess		
Antecedent	Hi (n=25)	Lo (n=3)
Quiz results correct	23	2
Quiz explanations correct	2	1

Case 4: Instances of Delight during Conv		
Antecedent	Hi (n=29)	Lo (n=7)
Post-quiz hint	6	2
Encouragement prompt	1	1
Agent progress report (good)	5	1
Agent response to query	2	0
MapView/Read→Quiz	14	1
Off-task	1	2

Case 5: Instances of Delight during Read		
Antecedent	Hi (n=10)	Lo (n=4)
MapView → Read guide	1	1
Hint → Read	4	1
Off-task	1	2
Agent progress report (good)→Read	2	0
MapEdit→Read	2	0

As shown in Table 6, several of the shaded antecedents seem to trigger major differences in the number of instances observed in Hi (*high performers*) versus Lo (*low performers*). Therefore, we study these cases in greater detail below.

1. Case 1 - *Boredom during Conversation*: Several antecedents of boredom observed during agent conversations are noted in Table 6. Again, we explore the ones showing large Hi vs Lo differences below:

(i) Post-quiz *indirect* hint: Lo students were bored in 5 instances after receiving *indirect* post-quiz hints from the mentor agent. No such instances of boredom were observed in Hi students. In Scenario 2, we found that Hi students showed more delight instances compared to Lo students when they got post-quiz hints from the mentor. In contrast, the Lo students showed higher instances of boredom (*Hi students show none*) after receiving *indirect* hints, implying these students did not find the hints to be useful, or they could not use them to overcome their difficulties.

(ii) Bad progress report by agent: A second instance, of large differences between Hi and Lo students, is when the mentor agent informs them that they are not making sufficient progress toward completing their map. This occurs more for low performers, and lack of progress results in boredom for these students. This is a situation, where we have to rethink the type of feedback provided by the mentor agent, and it may differ for Lo versus Hi performers.

(iii) Long MapView or Read→*boredom*, Conv: In this case boredom started with a previous action i.e., long periods of reading the resources or viewing the causal map. This boredom was still present when the conversation action occurred. For high performers, there were two such instances (both due to long reads), and for

low performers there were 11 such instances (4 due to long reads and 7 due to long map views). This result also suggests that scaffolds should be designed to help low-performing students regulate their affect with more helpful hints, especially when they spend too much time on a resource page or just viewing the causal map without making changes.

2. Case 2 - *Boredom during Edit actions*: The three notable antecedents of boredom in this case are summarized below:

(i) Hint → MapView is an antecedent of boredom during map edit. This case is only seen among Lo performers, with 7 instances in Lo performers. The student received a hint from the mentor agent and then started looking at the causal map, and then got bored. This result appears to follow from [19] in that the gap between task difficulty and learner ability leads to boredom. Scaffolding Lo students in such a situation can ensure that they are able to utilize the mentor's hint effectively and do not get disengaged from the task at hand.

(ii) Long period of viewing map or moving map elements (no read, quiz or link-edit): These are several instances (4 for Hi and 8 for Lo) where long periods of viewing the map and moving map elements on the screen (idle period with no quiz taking or reading or link editing in between) were followed by an observed bored state. This implies that it may be important to detect unproductive phases and to provide adequate scaffolds to help the students become more productive.

(iii) QuizProvenWrong→MapView: Students got bored when viewing their causal map following a quiz result that was graded as incorrect. This is more frequent for Lo, most likely because Lo students generally have worse quiz results, and not much of an idea of how to fix the errors in their map.

3. Case 3 - *Delight during Assess*. For both Hi and Lo learners, there are only two antecedents to delight during Assess actions: (i) when they looked at their quiz results and found that all the answers were correct; and (ii) when the causal explanations to quiz answers were correct. The first antecedent shows large differences between Hi and Lo groups. It is not surprising that the count of such instances was much higher for Hi group, since they were more successful in their map building efforts and hence obtained higher quiz scores.

4. Case 4 - *Delight during Conversation*. The antecedents of delight during conversation, as observed from the action logs, can be subdivided into different cases. Of these, the cases where large differences are noted between instances observed in Hi and Lo groups are discussed below:

(i) Post-quiz hint: this is the situation when the mentor agent gave a post-quiz hint. These hints may be *direct* or *indirect* (discussed in Section 3.2).

It was found that Hi students got delighted more often after receiving hints from the mentor agent. (We later observe, in Scenario 3, that Lo students show higher instances of boredom following *indirect* hints.)

(ii) Good progress report by agent: Delight also occurs when the mentor agent analyzes the state of the students' maps and praises their progress toward generating the correct map from the last

time the mentor performed this check. Unsurprisingly, these situations occur more often for the Hi group

(iii) MapView or Read→Quiz: In this situation, the student starts showing delight much before they have even begun a conversation with an agent. In all instances of this type, the student is delighted when they follow up a map view or read action by taking a quiz action i.e., looking at the map or reading resource pages and then asking Betty to take a quiz. This delight continues through the conversation with her, presumably because the student does well in the quiz. In other words, students had figured out how to improve their map, and this resulted in delight. The delight continued through the quiz taking action. This case occurred 14 times for the Hi performers and only once for the Lo performers.

5. Case 5 - *Delight during Read*: 5 antecedents were identified from student logs for the scenario where student got delighted during reading a resource or the teacher's guide page in Betty's Brain. Hint → Read is the only case where a large difference between Hi and Lo instances was observed.

Hint → Read: This is when students got delighted upon reading resource pages after the mentor agent gave a hint. Hi students have more instances of this type, possibly since they were better able to connect the hint to causal relations discussed in resource pages that the hint mentioned. This finding, along with the Hint→MapView antecedent of boredom discussed in Scenario 4, give us a very fine-grained insight into student cognition and affect, and call for the design of scaffolds to regulate cognitive and affective states in these situations.

Tracing the temporal activity sequences in the log files that occurred around affect events that were recorded by researchers using the BROMP protocol provides us with insights on the interactions between students' activities, performance, and affect. Also, one of the most important findings in this section is that an incongruity between the students' capabilities and the feedback provided to them by the agents in the Betty's Brain environment leads to unproductive affect states. This happens disproportionately for low performers, who are the ones likely to need the most encouragement and hints that help them get back to more productive behaviors. We discuss this further in the discussion section.

5.6 Performance Consequents of Affective States

Finally, we explore the immediate influence of affective experiences on student performance in Betty's Brain by comparing how Hi and Lo students edit their causal model in the 3 min interval after each affect occurs. Specifically, we compare the number of *effective* (correct) and *ineffective* (incorrect) link edits in the causal model. We chose a 3 min interval for this purpose since two consecutive affect observations for a given student had a mean duration of 3 min.

Link edit effectiveness, the measure of performance used here, is determined by comparing the current state of the student map after each causal link edit with an expert map embedded in the system that is hidden from the student's view.

The results of this analysis, given in Table 7, show how boredom is associated with poorer future performance among both Hi and Lo students. After experiencing delight, Hi students make 48% more effective than ineffective link edits. But Lo students appear to make more mistakes in map building after getting delighted and show 70% more ineffective than effective link edits, possibly due to careless errors from over-confidence [7, 25]. It appears that Hi students can regulate their frustration more productively than Lo students, as they do 22% more effective edits than ineffective edits after frustration. But frustration is associated with poorer performance for Lo students who show 32% more ineffective than effective links. As for confusion, it appears to be productive for both, it is more so for Hi students. This finding is supported by prior research on the positive influence of confusion on learning [8]. The precedence of confusion to positive performance would also harmonize with empirical studies around the co-evolution framework, which predict and find cognitive conflicts to make people accommodate their shared knowledge representations in social web environments [13].

Table 7: Comparing student performance in Hi vs Lo after each affective state

	Preceding Affect	Effective Links		Ineffective Links		Effective – Ineffective Links
		N	%	N	%	
HI	Confusion	244	61%	15 6	39%	22%
	Frustration	258	56%	20 1	44%	12%
	Delight	130	66%	67	34%	32%
	Boredom	40	42%	55	58%	-16%
LO	Confusion	144	57%	10 7	43%	15%
	Frustration	52	40%	77	60%	-19%
	Delight	9	23%	31	78%	-55%
	Boredom	74	45%	90	55%	-10%

6 DISCUSSION AND FUTURE WORK

In this paper, we analyzed cognitive and affective experiences of sixth grade students modeling a complex science topic in the Betty’s Brain OELE. Our analyses demonstrate that a combined study of students’ activities in an open-ended learning environment, their resulting learning performance and their recorded affective states can allow for a systematic and detailed investigation of the interactions between cognitive and affective processes, both important components of their self-regulated learning behaviors, within different task contexts in the OELE.

We found that learners’ cognitive and affective states within an agent-based OELE are contextualized by their interactions with the environment as well as their performance within the system. Our initial results revealed significant differences in the levels of *delight* and *boredom* between high and low performers. So, we

identified five specific task contexts in which these affect differences were most pronounced. Then we determined the cognitive processes, which were temporal antecedents to these contextualized affect cases in high versus low performers. However, our findings on temporal relationships between cognition and affect states do not claim any possible affect-cognition causality.

One particularly fascinating finding from our temporal analyses is how the virtual “mentor” agent’s feedback (*in the form of hints or reports of progress*) appeared to lead to different cognitive and affective experiences in high versus low performers. The different learner reactions to feedback recalls previous research on learner’s beliefs about the nature of learning [9].

These findings suggest that it may be valuable to scaffold students, especially low performers, based on their cognitive and affective interactions in specific learning situations to help regulate their SRL processes and create a more productive learning experience. The findings in this paper show the value of aligning affect with students’ performance and their actions, and then developing the ability to track them online, so that personalized hints and positive encouragements can be provided to help students get back to productive learning behaviors.

A primary limitation of this work was that students’ affect was recorded at discrete intervals, therefore, it was not possible to track transitions and evolution of their affective states in a fully fine-grained manner. On the other hand, recording of students’ activities in context, and their performance over time, though discrete and recorded as events in log files, was available in a more complete form.

In future, to enable better alignment between affect, learning activities and performance, we plan to develop and deploy affect detectors that will be capable of recording students’ affective states at finer-grained intervals. Once this is established we hope to develop temporal models of the interactions between learning activity and behaviors, affective states and performance, their transitions, and their evolutions over time. This will provide a better framework for developing rich learner models of SRL processes, which, in turn, will lead to richer frameworks for scaffolding and feedback in OELEs.

ACKNOWLEDGMENTS

This research was supported by NSF ECR Award #1561676.

REFERENCES

- [1] R Azevedo, R Behnagh, M Duffy, J Harley and G Trevors. 2012. Metacognition and self-regulated learning in student-centered learning environments. *Theoretical foundations of student-centered learning* 171-197.
- [2] Roger Azevedo and Allyson Hadwin. 2005. Scaffolding Self-Regulated Learning and Metacognition—Implications for the Design of Computer-Based Scaffolds. *Instructional Science*, Vol 33. 367-379. 10.1007/s11251-005-1272-9.
- [3] Ryan Baker and Jaclyn Ocumpaugh. 2015. Interaction-Based Affect Detection in Educational Software. *In the Oxford Handbook of Affective Computing: Oxford University Press*.
- [4] L F Barrett. 2009. Variety is the spice of life: a psychological construction approach to understanding variability in emotion. *Cognition & Emotion*, Vol 23(7), 1284-1306.
- [5] Gautam Biswas, James R Segedy and Kritya Bunchongchit. 2016. From Design to Implementation to Practice a Learning by Teaching System: Betty’s Brain. *Int J Artif Intell Educ*, Vol 26: 350. <https://doi.org/10.1007/s40593-015-0057-9>

- [6] M Boekaerts, M and L Corno. 2005. Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, Vol 54(2), 199–231
- [7] K Clements. 1982. Careless errors made by sixth-grade children on written mathematical tasks. *Journal for Research in Mathematics Education*, Vol 13, 136–144.
- [8] S D’Mello, B Lehman, B, R Pekrun and A Graesser. 2014. Confusion can be beneficial for learning. *Learning and Instruction*, Vol 29: 153-170.
- [9] C S Dweck. 2002. Messages that motivate: How praise molds students’ beliefs, motivation, and performance (In surprising ways). In J. Aronson (Ed.), *Improving academic achievement*. New York: Academic Press.
- [10] A Efklides. 2011. Interactions of metacognition with motivation and affect in self-regulated learning: the MASRL model. *Educ. Psychol.* Vol 46, 6–25. doi: 10.1080/00461520.2011.538645
- [11] Jeffrey Alan Greene and Roger Azevedo. 2007. A Theoretical Review of Winne and Hadwin’s Model of Self-Regulated Learning: New Perspectives and Directions. *Review of Educational Research* Vol 77, Issue 3, 334 – 372.
- [12] Jeffrey D Karpicke and Henry L Roedinger. 2008. The Critical Importance of Retrieval for Learning. *Science*, Vol 319, 5865, 966-968.
- [13] Joachim Kimmerle, Ulrike Cress, Christoph Held and Johannes Moskaliuk. 2010. Social software and knowledge building: supporting co-evolution of individual and collective knowledge. *Proceedings of the 9th International Conference of the Learning Sciences*, Vol 1, 9-16.
- [14] John Kinnebrew, James R Segedy and Gautam Biswas. 2017. Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments. *IEEE Trans. Learn. Technol.* doi: 10.1109/TLT.2015.2513387
- [15] Reed W Larson and Maryse H Richards. 1991. Boredom in the Middle School Years: Blaming Schools versus Blaming Students. *American Journal of Education*, Vol 99, 4, 418-443.
- [16] K Leelawong and G Biswas. 2008. Designing learning by teaching agents: The Betty’s Brain system. *International Journal of Artificial Intelligence in Education*, Vol 18(3), 181–208. IOS Press.
- [17] Jaclyn Ocumpaugh, Ryan Shaun Joazeiro de Baker, Ma Mercedes T Rodrigo, Aatish Salvi, Martin Van Velsen, Ani Aghababayan and Taylor Martin. 2015. HART: the human affect recording tool. *SIGDOC 2015*.
- [18] J Ocumpaugh, R S Baker and M M T Rodrigo. 2015. Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) 2.0 technical and training manual. *Technical Report*. New York, NY: Teachers College, Columbia University. Manila, Philippines: Ateneo Laboratory for the Learning Sciences.
- [19] J Ocumpaugh et al. 2017. Affect Dynamics in Military Trainees Using vMedic: From Engaged Concentration to Boredom to Confusion. In: *André E., Baker R., Hu X., Rodrigo M., du Boulay B. (eds) Artificial Intelligence in Education. AIED 2017. Lecture Notes in Computer Science*, Vol 10331. Springer
- [20] Ernesto Panadero, Julia Klug and Sanna Järvelä. 2015. Third wave of measurement in the self-regulated learning field: when measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research*, DOI: 10.1080/00313831.2015.1066436
- [21] R Pekrun. 2010. Academic emotions. In T. *Urduan (Ed.), APA educational psychology handbook*, Vol. 2. Washington, DC: American Psychological Association
- [22] R Pekrun. 2006. *Educ Psychol Rev.* 18: 315. <https://doi.org/10.1007/s10648-006-9029-9>
- [23] P R Pintrich. 2000. The role of goal orientation in self-regulated learning. In *Handbook of Self-Regulation*, eds M. Boekaerts, P. R. Pintrich, and M. Zeidner (San Diego, CA: Academic Press), 452–502.
- [24] Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110, 145e172.
- [25] M O Z San Pedro, R S J Baker and M M T Rodrigo. 2014. *Int J Artif Intell Educ*, Vol 24, 189. <https://doi.org/10.1007/s40593-014-0015-y>
- [26] G Schraw, K J Crippen and K Hartley. 2006. Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in science education*, Vol 36(1-2), 111-139.
- [27] Allan Wigfield, Jacquelynne S Eccles. 2000. Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, Vol 25, 1, 68-81, ISSN 0361-476X, <https://doi.org/10.1006/ceps.1999.1015>
- [28] P H Winne. 2001. Self-regulated learning viewed from models of information processing. In *Self-Regulated Learning and Academic Achievement*, eds B. J. Zimmerman and D. H. Schunk (New York, NY: Lawrence Erlbaum Associates), 153–190.
- [29] P H Winne. 2011. A cognitive and metacognitive analysis of self-regulated learning. In *Handbook of Self-Regulation of Learning and Performance*, eds B. J. Zimmerman and D. H. Schunk (New York, NY: Routledge), 15–32.
- [30] P H Winne, A F Hadwin. 1998. Studying as self-regulated engagement in learning. In *Metacognition in Educational Theory and Practice*, eds D. Hacker, J. Dunlosky, and A. Graesser (Hillsdale, NJ: Erlbaum), 277–304.
- [31] P H Winne and A F Hadwin. 2008. The weave of motivation and self-regulated learning. In *Motivation and Self-Regulated Learning: Theory, Research and Applications*, eds D. H. Schunk and B. J. Zimmerman (New York, NY: Lawrence Erlbaum Associates), 297–314.
- [32] B J Zimmerman. 2000. Attaining self-regulation: a social cognitive perspective. In *Handbook of Self-Regulation*, eds M. Boekaerts, P. R. Pintrich, and M. Zeidner (San Diego, CA: Academic Press), 13–40. doi: 10.1016/b978-012109890-2/50031-7
- [33] Barry J Zimmerman. 2008. Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal*, Vol 45, 1, 166 - 183.
- [34] B J Zimmerman, D H Schunk. 2011. *Handbook of Self-Regulation of Learning and Performance*. New York, NY: Routledge.