How do Emotions Change during Learning with an Intelligent Tutoring System? Metacognitive Monitoring and Performance with MetaTutor

Elizabeth B. Cloude[1] (elizabeth.cloude@knights.ucf.edu)
Franz Wortha[2,3] (franz.wortha@uni-tuebingen.de)
Daryn A. Dever[1] (ddever@knights.ucf.edu)
Roger Azevedo[1] (roger.azevedo@ucf.edu)

[1]Department of Learning Sciences and Educational Research, 4364 Scorpius St, Orlando, FL 32816 USA
[2]LEAD Graduate School and Research Network, Geschwister-Scholl-Platz, 72074 Tübingen, Germany
[3]Leibniz-Institut fur Wissensmedien (IWM), Schleichstraße 6, 72076 Tübingen, Germany

Abstract

Emotional experiences have a significant impact on learning about complex topics. Yet, challenges exist because emotions are typically operationalized as end products, excluding if, how, and when emotions change during learning and their relation to metacognition and performance with advanced learning technologies such as intelligent tutoring systems (ITSs). In this paper, we addressed these challenges by capturing and analyzing 117 college students' concurrent and self-reported emotions at 3 time points during learning with MetaTutor, an ITS. Analyses revealed negative relationships between increases in boredom, metacognitive monitoring accuracy, and performance. We also found that if confusion persisted over time during learning, it was detrimental to performance. These findings provide implications for designing affect-sensitive ITSs which foster emotion-regulation and metacognitive monitoring based on changes in emotions during learning to optimize performance.

Keywords: Emotions; Self-regulated learning; Intelligent Tutoring Systems; Metacognition

Introduction

Studies suggest emotions impact learners' ability to understand complex topics with advanced learning technologies (Lajoie, Pekrun, Azevedo, & Leighton, in press), where positive and negative emotions may influence learners differently (Wortha, Azevedo, Taub, & Narciss, 2019). Positive emotions, such as excitement, were found to be positively associated with learning processes such as metacognitive monitoring and performance during learning compared to negative emotions (Ahmed, Van der Werf, Kuyper, & Minnaert, 2013). Yet, other studies revealed positive relationships between negative emotions, learning processes, and performance. For instance, D'Mello, Lehman, Pekrun, and Graesser (2014) found that when learners were confused, it resulted in better performance on post- and transfer tests compared to learners who were not confused. Similarly, a study found that negative emotions such as frustration were not detrimental to learning, but rather the persistence of a negative emotion (e.g., boredom) throughout the learning session was more indicative of poor performance (Baker, D'Mello, Rodrigo, & Graesser, 2010). However, another study found evidence suggesting that boredom was negatively related to SRL processes and performance (Obergriesser & Stoeger, in press).

Mixed findings on whether negative emotions are beneficial or detrimental to learning processes and performance begs a methodological question: how are studies capturing and defining emotions? Traditionally, in educational, cognitive, and psychological studies, emotions are captured before and after a learning session using self-report questionnaires, quantifying emotions as static states that do not change across tasks and learning (Price, Mudrick, Taub, & Azevedo, 2018). Other studies use observational methods to capture students' emotions (Baker, Ocumpaugh, & Andres, 2019). Recently, studies have begun using facial-recognition software such as Taub and colleagues (2019) to define emotions. They examined whether positive and negative emotions were related to self-regulated learning (SRL) processes such as metacognitive monitoring accuracy, cognition strategy use, and performance. Their results showed that frustration was positively related to learning processes and performance (Taub et al., in press). However, gaps remain in this methodological approach as emotions were not captured as dynamic, fluctuating states that change over time during learning to examine its impact on learning processes such as metacognitive monitoring accuracy and performance.

Theoretical Frameworks

We grounded our work in (1) the information-processing theory (IPT) of SRL because it describes learning as a cyclic, non-linear process that occurs over time (Winne, 2018), and (2) the model of affective dynamics (D'Mello & Graesser, 2012) because it explains how and why emotions arise during learning with technologies such as an ITS, while defining emotions as dynamic and fluctuating of which depends on the learner and the information they are learning about.

IPT of SRL describes learners as taking an active role in their learning consisting of goal setting and information-processing events that occur during four phases of learning (Winne, 2018). Each phase describes how learners navigate and complete tasks, represented in a feedback loop that is interrelated and dependent on the learners' SRL process use. The phases consist of (1) task definition, where learners must define what is required to complete the task, (2) goals and planning, where learners must set goals and initiate plans to complete the task based on its requirements, (3) enact strategies, and (4) metacognitively monitor and adapt strategies initiated based on the learners' feedback loop, reflecting their motivation, beliefs, and continuous evaluation of progress toward achieving their set goals in phase 2. Previous research using IPT of SRL has found evidence supporting that the more often SRL is used, the better learners perform (Azevedo, Taub, & Mudrick, 2018). As such, we analyzed variables related to phase 4 in this study (i.e., metacognitive monitoring accuracy), because phase 4 depends on information gathered in phases 1-3 and plays a critical role in one's ability to self-regulate their learning. However, IPT of SRL does not account for how emotions experienced during learning might change, so we used the model of affective dynamics to explain how emotions occur during learning with an ITS and develop our research questions/hypotheses.

D'Mello and Graesser (2012) describe that the model of affective dynamics is based on cognitive disequilibrium, which is critical for deep, conceptual understanding of information. Cognitive disequilibrium is defined as uncertainty stemming from impasses (e.g., being confused about a biology concept) in completing goals (e.g., understanding the path of blood flow) during learning. The model assumes that learners are in a constant state of (1) engagement or (2) disengagement (e.g., boredom) while learning, distinguished based on whether learners are working toward their goals (i.e., engagement) or not (i.e., disengagement). When learners reach cognitive disequilibrium upon encountering an impasse, they experience confusion and must engage in activities (e.g., metacognitive monitoring) to resolve the impasse and alleviate their confusion. Once the impasse is resolved, cognitive equilibrium is restored. Although, if the impasse is left unresolved after the learner attempted to work through their confusion, they get stuck, leading to frustration and eventually boredom and disengagement. For example, if learners were confused about the path of blood flow during learning with MetaTutor, they must metacognitively monitor their confusion (e.g., enact a judgment of learning to assess how well their perceived understanding of blood flow is aligned with their actual understanding) and then regulate by adapting their strategies (e.g., reread information on path of blood flow). Since empirical evidence has found mixed results regarding the relations between negative emotions, SRL processes, and performance, to what degree do changes in emotions (e.g., increases in confusion at 3 times points) impact learners' metacognitive monitoring accuracy and performance?

Fusing these models together helps explain what might give rise to an impasse during learning, leading to cognitive disequilibrium and a subsequent emotional experience. IPT of SRL describes the phases of learning such as metacognitively monitoring their understanding of information. Learners adapt their strategies based on information acquired from their judgment of learning based on impasses presented during learning (e.g., realizing they do not understand as much about the content as they had predicted, etc.). This impasse gives rise to cognitive disequilibrium and an emotional experience, occurring at any point during a learning session. Based on methodological gaps related to capturing emotions as states that do not change and mixed findings in literature about the role of negative emotions during learning, it is critical to examine how emotions might change during learning to gain insight into their relation to metacognitive monitoring accuracy and performance after learning with an intelligent tutoring system. In our study, those gaps were addressed.

Current Study

Confusion, boredom, and frustration were captured and analyzed at 3 time points during learning based on assumptions explained in the model of affective dynamics (see Coding and Scoring for justification on using 3 time points), and IPT of SRL was used to examine and interpret relations between emotional changes, judgments of learning accuracy (i.e., metacognitive monitoring), and performance after learning with MetaTutor.

Our research questions and hypotheses include (1) **To what extent are there relationships between metacognitive monitoring accuracy and post-test scores while controlling for pre-test scores and condition after learning with MetaTutor?** We hypothesize there will be significant, positive relationships between metacognitive monitoring accuracy and post-test scores while controlling for pre-test scores and condition after learning with MetaTutor. (2) **Are there differences in how changes in confusion, boredom, and frustration are distributed at 3 time points during learning with MetaTutor?** We hypothesize there will differences in how changes in confusion, boredom, and frustration are distributed at 3 time points during learning with MetaTutor. (3) **To what extent are there relationships between if, when, and how confusion, boredom, and frustration change at 3 time points and post-test scores while controlling for pre-test scores and condition after learning with MetaTutor?** We hypothesize there will be relationships between if, when, and how confusion, boredom, and frustration change at 3 time points and post-test scores while controlling for pre-test scores and condition after learning with MetaTutor, where changes in confusion will be positively related to post-test scores and no changes in confusion will be negatively related to post-test scores. Additionally, we hypothesize that changes in boredom and frustration will be negatively related to post-test scores, whereas no changes in boredom and frustration will be positively related to post-test scores. (4) **To what extent are there relationships between changes in confusion, boredom, and frustration at 3 time points and metacognitive monitoring accuracy while controlling for condition during learning with MetaTutor?** We hypothesize there will be relationships between changes in confusion, boredom, and frustration change at 3 time points and metacognitive monitoring accuracy while controlling for condition after learning with MetaTutor. We do not provide a direction because the theoretical frameworks used do not indicate how changes in negative emotions impact metacognitive monitoring accuracy in conjunction with the mixed findings in literature.

Methods

Participants and Materials

A total of 199 college students were recruited from three, large North American universities and completed a 2-day, quasi-experimental study with MetaTutor. In this paper, we analyzed 117 college students (58% female; 72% Caucasian) because they met our inclusion criteria (see Coding and Scoring). Ages ranged from 18 to 35 years (M=20.19, SD=2.21) where students were compensated $10/hour in a study lasting up to 2.5 hours.

To assess knowledge of biology, a 30-item, pre- (M=0.59, SD=0.15) and equivalent post-test (M=0.70, SD=0.14) was administered before and after learning with MetaTutor. The Emotions Value (EV) questionnaire was administered every 14 minutes during learning to capture concurrent emotional states using a 5-point, Likert scale measuring 19 emotions defined as discrete, basic, and learner-centered and 2 items measuring task value and perceived ability to perform (Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015). These items were grounded in theories and empirical research on emotions (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). In this paper, we focus on self-reported confusion, boredom, and frustration to align with the model of affective dynamics.

MetaTutor

MetaTutor is a multi-agent ITS designed to teach students about the circulatory system using 47 pages of text and diagrams with four embedded pedagogical agents (PAs; Figure 1). Each PA was designed to mirror tutoring behaviors such as providing feedback on performance (e.g., quiz scores) and prompting SRL processes and strategies during learning (Azevedo et al., 2018). Specifically, MetaTutor was designed with two conditions where students in the control condition did not receive feedback on performance or prompts from PAs to use SRL processes compared to students in the experimental (i.e., prompt and feedback) condition.

Each agent specialized in a component of SRL–e.g., Gavin the Guide introduced students to MetaTutor by highlighting tools, whereas Pam the Planner emphasized planning activities and creating sub-goals. Mary the Monitor prompted students to monitor progress toward achieving goals and Sam the Strategizer asked students to use strategies such as summarizing content. All students were informed the overall objective of learning with MetaTutor was to gain as much knowledge about the circulatory system during the learning session. Before learning, all students were required to set at least two sub-goals (e.g., components of the circulatory system) related to the objective, aligning with IPT of SRL (Winne, 2018). MetaTutor was also designed with interface elements that fostered SRL including (1) a table of contents, (2) a timer reflecting the amount of time left, and (3) progress bars representing the amount of progress made toward each sub-goal as well as the overall objective.

Tools designed to foster metacognition. Mary the Monitor was designed to foster metacognition during learning for students in the experimental condition (Azevedo, Witherspoon, Chauncey, Burkett, Fike, 2009). Mary prompted judgments of learning (JOLs; see top-right in Figure 1), content evaluations, and feeling of knowing judgments to trigger and foster metacognitive processes during and after students viewed content related to the objective and created sub-goals. When students used metacognitive processes such as a JOL, they did so by selecting an SRL process or strategy on the SRL palette (right-hand side of Figure 1). Students were not prompted to use metacognitive processes in the control condition, but in both conditions participants were free to initiate these processes using the SRL palette.

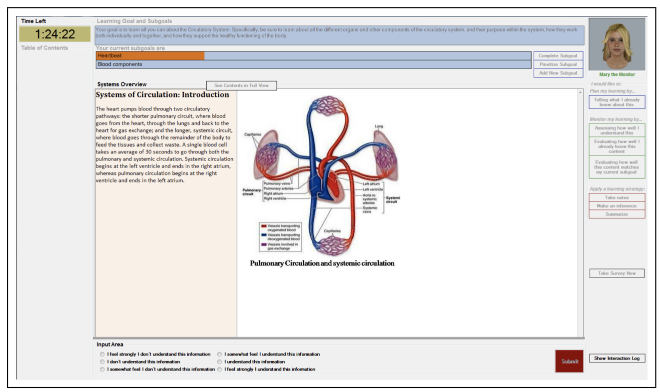

Figure 1: MetaTutor interface

Procedure

The study lasted 2 days. On day 1, students were randomly assigned to 1 of 2 conditions: (1) prompt and feedback (n=69) and (2) control (n=48). More students were in the experimental condition because they more often than not met our inclusion criterion for analyses (see Participants and Materials). We suspect the nature of the experimental condition required students to stay in the environment longer than the control (e.g., agent prompts, etc.), resulting in completing more EVs every 14 minutes. Following consent, students completed demographic questions, questionnaires that gauged emotions and motivation, as well as a 30-item, multiple choice pretest. On the second day, students returned to the lab space and were calibrated to an eye tracker, electrodermal activity bracelet, and facial recognition software. Next, they were required to set two sub-goals before learning with MetaTutor and then started the learning session. Both conditions required students to complete the EV approximately every 14 minutes over the learning session, which we used to capture if, how, and when emotions changed during learning. Once the learning session was ended, students completed a similar, counterbalanced 30-item, multiple-choice post-test and self-report questionnaires that gauged motivation and emotion regulation. Next, students were paid, debriefed, and thanked for their participation.

Coding and Scoring

Performance variables were operationalized using pre- and post-test scores, such that pre-test scores defined students' level of prior knowledge which plays a crucial role in metacognitive monitoring accuracy and knowledge acquisition (i.e., post-test scores) according to IPT of SRL and previous empirical studies. JOLs were used to define metacognitive monitoring accuracy, while changes in EV scores at 3 time points captured differences in self-reported confusion, boredom, and frustration during learning with MetaTutor. These variables were used to assess relations between emotional changes, self-regulation–i.e., metacognitive monitoring accuracy and performance with MetaTutor.

Emotional changes over 3 time points. MetaTutor prompted all students to complete the EV approximately every 14 minutes during learning. We analyzed EV scores captured at (1) 0 minutes (i.e., before learning), (2) 14 minutes (i.e., at the start of learning), (3) 28 minutes (roughly half-way through learning), and (4) 42 minutes into learning (i.e., the late stages of learning). Students were included in our analyses if they completed the EV at these instances. To capture changes in emotions, we calculated whether confusion, boredom, and frustration (1) increased (e.g., confusion at 0 minutes was less than confusion at 14 minutes), (2) decreased (e.g., confusion at 0 minutes was greater than confusion at 14 minutes), or (3) did not change (e.g., confusion at 0 minutes was equal to confusion at 14 minutes) from time points (1) 0 to 14 minutes, (2) 14 to 28, and (3) 28 to 42 minutes. All participants were assigned to a group that either (1) increased, (2) decreased, or (3) did not change for each emotion based on the initial self-reported emotion at 3 time points.

Performance. Pre- and post-test scores were calculated separately by summing all of the correct items and dividing by 30 to generate pre- and post-test ratio scores.

Metacognitive monitoring accuracy. Metacognitive accuracy was calculated using judgments of learning (JOLs) initiated using the SRL palette on the MetaTutor interface. Once a JOL was initiated, students reported, *How well do you feel you understand the content on this page?* on a 6-point scale (1=*I feel I strongly do not understand*; 6=*I feel strongly I understand*) based on previously-viewed content. After reporting, students completed a 3-item quiz based on the page they had made the JOL to assess how aligned their rating was with actual performance–i.e., quiz score. We used a coding scheme that calculated the accuracy of JOLs by assessing 3 components, where (1) 50% of points were allocated based on how aligned the JOL was with the quiz (e.g., student reported little understanding and performed poorly on the quiz, their metacognitive monitoring was more accurate than if they reported little understanding but performed well on the quiz; (2) 25% of the points were allocated based on quiz score; and (3) 25% of the points were allocated based on JOL rating (Taub et al., in press). Students in the experimental condition generated an average of 27% (*SD*=0.31) accuracy, while students in the control generated an average of 59% (*SD*=0.26) accuracy.

Results

Data were extracted from logfiles using 'Numpy' (Oliphant, 2006) and 'Pandas'(McKinney, 2011) packages in Python. The dataset was processed in R Studio (R version 3.5.1) using 'dplyr' (Wickham, Francois, Henry, & Müller, 2015). We conducted statistical analyses using SAS software (Version 9.4 for Windows). Homogeneous variance and normality assumptions were met prior to conducting the analyses.

RQ1: To what extent are there relationships between metacognitive monitoring accuracy and post-test scores while controlling for pre-test scores and condition during learning with MetaTutor?

A partial Pearson correlation was calculated to assess whether there were positive relationships between metacognitive monitoring accuracy and post-test scores while controlling for pre-test scores and condition. We found a significant, positive partial correlation between metacognitive monitoring accuracy and post-test scores, r=0.26, p=0.006, indicating a positive relationship between metacognitive monitoring accuracy and post-test scores (Figure 2). These findings are consistent with our hypothesis.

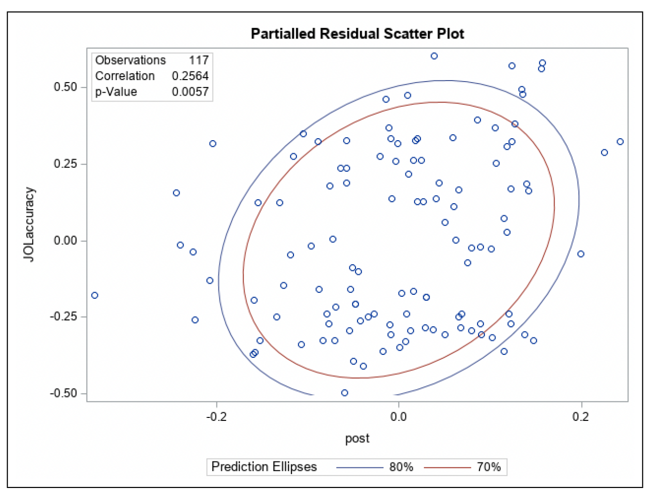

Figure 2: Partial correlation between metacognitive accuracy and post-test scores.

RQ2: Are there differences in how changes in confusion, boredom, and frustration are distributed over 3 time points during learning with MetaTutor?

Three Chi-squares were calculated to examine whether there were differences in the distribution of changes in confusion, boredom, and frustration at 3 time points during learning with MetaTutor. Using a Bonferroni correction ($p<0.05/3=0.017$), differences were found in the number of (1) increases, (2) decreases, and (3) no changes in confusion, boredom, and frustration at time points 1, 2, and 3. The results suggested differences in how emotional changes were distributed at time points 1-3 (i.e., 0 to 42 minutes) during learning with MetaTutor (see Table 1 and Figure 3). These findings support our hypothesis.

Table 1: Differences in emotional changes at time points 1-3

Time point	χ^2	p
1	24.44	<0.001
2	16.21	0.002
3	50.89	<0.001

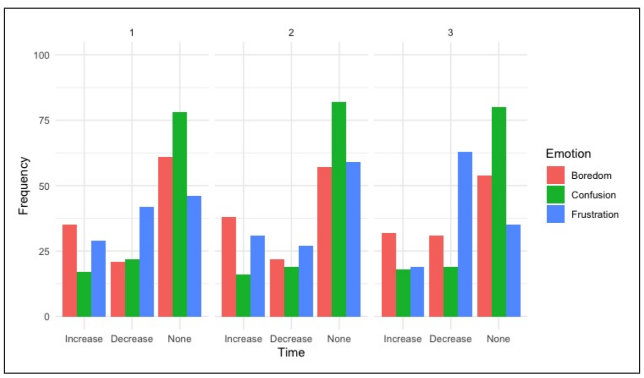

Figure 3: Frequency of emotions between 3 time points

RQ3: To what extent are there relationships between changes or no changes in confusion, boredom, and frustration over 3 time points and post-test scores while controlling for pre-test scores and condition after learning with MetaTutor?

To examine whether there were differences in post-test scores between changes in emotions during learning at 3 time points, three separate Analyses of Covariance (ANCOVA) were calculated for each 3 time point (i.e., (1) 0 to 14 minutes, (2) 14 to 28 minutes, and (3) 28 to 42 minutes). Using a Bonferroni correction (0.05/3=0.017), where pre-test scores and condition were included as covariates, the first model (i.e., time point 1) revealed significant differences in post-test scores between increases, decreases, and no changes in boredom that occurred from 0 to 14 minutes during learning, $F(5,111)=19.55$, $p<0.001$, where boredom, pre-test scores, and condition explained 47% of the variance in post-test scores. The fitted model estimated that average post-test scores decreased by 0.23 points when increases in boredom occurred between 0 to 14 minutes (relative to decreases or no changes) during learning with MetaTutor.

Significant differences were also found in post-test scores between increases, decreases, and no changes in confusion from 14 to 28 minutes during learning, $F(5,111)=20.54$, $p<0.001$, where confusion, pre-test scores, and condition explained 48% of the variance in post-test scores. The fitted model estimated that average post-test scores decreased by 0.23 points when no changes in confusion occurred between 14 to 28 minutes (relative to increases or decreases) during learning with MetaTutor. These findings partially support our hypotheses where we expected changes in emotions at 3 time points to be related to post-test scores while controlling for pre-test scores and condition.

RQ4: To what extent are there relationships between if, when, and how confusion, boredom, and frustration change over 3 time points and metacognitive monitoring accuracy while controlling for condition during learning with MetaTutor?

Three separate ANCOVAs were calculated for each time point (i.e., (1) 0 to 14 minutes, (2) 14 to 28 minutes, and (3) 28 to 42 minutes) between emotional change groups (i.e., (1) increases, (2) decreases, and (3) no changes) using a Bonferroni correction (0.05/3=0.017). Condition was included as a covariate in the model since a *t*-Test indicated significant differences in metacognitive monitoring accuracy between experimental ($M=0.28$, $SD=0.31$) and control conditions ($M=0.59$, $SD=0.26$). Pre-test was not included as a covariate in this research question because there were no differences in pre-test scores and metacognitive monitoring accuracy ($p<0.05$).

Analyses revealed significant differences in metacognitive monitoring accuracy between changes in boredom that occurred from 14 to 28 minutes, $F(5,111)=8.09$, $p<0.001$, where boredom and condition explained 27% of the variance in metacognitive monitoring accuracy. Specifically, we found a significant interaction when increases in boredom that occurred from 14 to 28 minutes during learning was negatively related to metacognitive monitoring accuracy for learners in the control condition. The fitted model estimated that average metacognitive monitoring accuracy decreased by 0.36 points when increases in boredom occurred between 14 to 28 minutes during learning with MetaTutor for learners in the control condition.

Discussion

To address gaps in research on emotions during learning, we examined if, how, and when confusion, boredom, and frustration changed at 3 time points during learning and their relation to metacognitive monitoring accuracy and performance after learning with MetaTutor. We found a positive relationship between metacognitive monitoring accuracy and post-test scores, while controlling for pre-test scores and condition. This finding was consistent with our hypothesis, IPT of SRL (Winne, 2018), and previous research (Taub et al., in press). We also found significant differences between changes in confusion, boredom, and frustration occurring between (1) 0 to 14 minutes, (2) 14 to 28 minutes, and (3) 28 to 42 minutes during learning with MetaTutor. These findings were consistent with our hypotheses and previous research (Wortha et al., 2019).

Next, we found negative relationships between post-test scores, changes in boredom, and no changes in confusion during learning with MetaTutor while controlling for pre-test scores and condition, where on average, post-test scores decreased by 0.23 points when increases in boredom occurred between 0 to 14 minutes during learning. These findings were consistent with our hypothesis where we expected increases in boredom to negatively relate to post-test scores while controlling for pre-test scores and condition, in addition to the model of affective dynamics (D'Mello & Graesser, 2012) and previous literature (Obergriesser & Stoeger, in press). For instance, students may have reached an impasse (e.g., not understanding a concept in biology) during 0 to 14 minutes of learning. If students were unable to resolve the impasse, they may have gotten stuck (i.e., frustrated), entering cognitive disequilibrium which theoretically leads to boredom and disengagement, potentially explaining lower post-test scores after learning. We also found that, on average, post-test scores decreased by 0.23 points when no changes in confusion occurred between 14 to 28 minutes during learning. This finding was consistent with our hypothesis, the model of affective dynamics (D'Mello & Graesser, 2012), and previous literature (Ahmed et al., 2013). For instance, if students reported no changes in confusion, they would not have reached a state of cognitive disequilibrium where they would need to assess and monitor the impasse to understand what they are confused about, potentially propelling a deeper, conceptual understanding of biology. However, we did not find any relationships between changes in frustration and post-test scores. This finding was inconsistent with our hypothesis, the model of affective dynamics (D'Mello & Graesser, 2012), and previous literature (Taub et al., in press). A possible explanation could be that administering the EV every 14 minutes may have missed the window in which confusion transitioned to frustration, or before frustration transitioned to boredom.

Last, we found a negative relationship between metacognitive monitoring accuracy and changes in boredom, where on average, metacognitive monitoring accuracy decreased by 0.36 points when increases in boredom occurred between 14 to 28 minutes particularly for the control condition. Not only was this result consistent with our hypothesis, but also the model of affective dynamics (D'Mello & Graesser, 2012), IPT of SRL (Winne, 2018), and previous studies (Baker et al., 2010). This finding was particularly interesting because we found a negative relationship between post-test scores and increases in boredom from 0 to 14 minutes. Yet, when boredom continued to increase from 14 to 28 minutes–i.e., suggesting that boredom persisted from 0 to 28 minutes during learning, it was related to lower metacognitive monitoring accuracy, consistent with Obergriesser et al. (in press). However, this result begs a question about the role of the pedagogical agent, and its ability to help down-regulate emotions that may be detrimental to learning. We would also like to highlight our results did not reveal a relationship between changes in frustration and metacognitive monitoring. This finding was inconsistent with previous research and the model of affective dynamics (D'Mello & Graesser, 2012), suggesting that frustration may not play a role in SRL or performance during learning with MetaTutor.

Limitations

Capturing changes in emotions at 14 minute intervals may not have been an accurate threshold to measure emotional changes (e.g., emotions may have changed before 14 minutes). Also, the trichotomy used does not consider the degree of change or emotional baseline. Additionally, we did not assess dynamics between emotional changes–e.g., examining if emotions such as confusion that persist for too long led to a transition into frustration or boredom.

Future Directions and Implications

Findings from this study suggest future research should examine if and how changes in emotions at certain stages during learning (e.g., beginning or middle of learning) might influence changes in emotions and other SRL processes–e.g., cognition and motivation (Cloude, Taub, & Azevedo, 2018; Cloude, Taub, Lester, & Azevedo, 2019), at later stages during learning with ITSs. For instance, do changes in emotions influence interest, and how is cognitive load impacted by emotions that persist for too long or too little? Sequential pattern mining could be a novel technique for examining if, how, and when emotions may change during learning and their relation to later emotions, SRL, and performance. Understanding the temporal dynamics of emotions and their causal relations to other learning-related processes (e.g., metacognitive monitoring accuracy) could provide insight into designing affect-sensitive ITSs that intervene when emotional changes reach a temporal threshold–e.g., confusion persisting from 0 to 28 minutes, which may be detrimental to SRL and performance.

Our next steps are to analyze the temporal dynamics between self-reported measures, physiology using sensors, and facial expressions of emotions and their relation to SRL processes and performance over the sequence of learning activities using latent growth modeling. Future studies should also analyze other data channels to capture changes in emotions during learning that go beyond self-reports. Capturing eye-gaze and logfiles might reveal what a student is doing in relation to changes in their emotions during learning, potentially highlighting if, what, when, and how the student is learning before, during, and after a particular emotion. For example, imagine a student is fixating on irrelevant content during learning with MetaTutor and experiences confusion. Implications of our findings could provide suggestions for building a system that could pinpoint what might have led a learner to experience a particular emotion. The system could use the emotional trigger to guide intelligent, adaptive scaffolding and feedback, such as redirecting the student's attention to relevant material and prompting their use of an emotional-regulation strategy to reduce confusion if it is detrimental to SRL, learning, and performance.

Acknowledgments

This research was funded by the National Science Foundation (DRL1661202) and the Social Sciences and Humanities Research Council of Canada (SSHRC 895-2011-1006). The authors would like to thank members of the SMART Lab at UCF for their assistance.

References

Ahmed, W., Van der Werf, G., Kuyper, H., & Minnaert, A. (2013). Emotions, self-regulated learning, and achievement in mathematics: A growth curve analysis. *Journal of Educational Psychology*, *105*, 150–161.

Azevedo, R., Taub, M., & Mudrick, N. V. (2018). Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies. In J. A. Greene & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 254–270). New York, NY: Routledge.

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*, 223–241.

Baker, R. S., Ocumpaugh, J. L., & Andres, J. (2019). Bromp quantitative field observations: A review. In R. S. Feldman (Ed.), *Learning science: Theory, research, and practice.* New York, NY: McGraw-Hill.

Cloude, E. B., Taub, M., & Azevedo, R. (2018). Investigating the role of goal orientation: Metacognitive and cognitive strategy use and learning with intelligent tutoring systems. In R. Nkambou, R. Azevedo, & J. Vassileva (Eds.), *Proceedings of the 14th international conference on intelligent tutoring systems* (pp. 44–53). Netherlands, Amsterdam: Springer.

Cloude, E. B., Taub, M., Lester, J., & Azevedo, R. (2019). The role of achievement goal orientation on metacognitive process use in game-based learning. In S. Sotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), *Proceedings of the 20th international conference on artificial intelligence in education* (pp. 36–40). Cham, Switzerland: Springer.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*, 145–157.

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior*, *48*, 615–625.

Lajoie, S. P., Pekrun, R., Azevedo, R., & Leighton, J. P. (in press). Understanding and measuring emotions in technology-rich learning environments. *Learning and Instruction*.

McKinney, W. (2011). pandas: A foundational python library for data analysis and statistics. *Python for High Performance and Scientific Computing*, *14*.

Obergriesser, S., & Stoeger, H. (in press). Students' emotions of enjoyment and boredom and their use of cognitive learning strategies–how do they affect one another? *Learning and Instruction*.

Oliphant, T. E. (2006). *A guide to numpy* (Vol. 1). Scotts Valley, CA: Trelgol Publishing.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (aeq). *Contemporary Educational Psychology*, *36*, 36–48.

Price, M. J., Mudrick, N. V., Taub, M., & Azevedo, R. (2018). The role of negative emotions and emotion regulation on self-regulated learning with metatutor. In R. Nkambou, R. Azevedo, & J. Vassileva (Eds.), *Proceedings of the international conference on intelligent tutoring systems* (pp. 170–179). Netherlands, Amsterdam: Springer.

Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., & Price, M. J. (in press). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learning and Instruction*.

Wickham, H., Francois, R., Henry, L., & Müller, K. (2015). *dplyr: A grammar of data manipulation. r package version 0.4. 3.* Retrieved from `http://CRAN.R-project.org/package=dplyr`.

Winne, P. H. (2018). Cognition and metacognition within self-regulated learning. In J. A. Greene & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 52–64). NY, New York: Routledge.

Wortha, F., Azevedo, R., Taub, M., & Narciss, S. (2019). Multiple negative emotions during learning with digital learning environments–evidence on their detrimental effect on learning from two methodological approaches. *Frontiers in Psychology*, *10*, 2678.