Affective Sequences and Student Actions within Reasoning Mind

Jaclyn Ocumpaugh[1] Ryan S. Baker,[1] Shamya Karumbaiah,[1] Scott A. Crossley[2] and Matthew Labrum[3]

[1] University of Pennsylvania
[2] Georgia State University
[2] Imagine Learning

ojaclyn@upenn.edu

Abstract. Now that the modeling of affective states is beginning to mature, understanding affect dynamics has become an increasingly realistic endeavor. However, the results from empirical studies have not always matched those of theoretical models, which raises questions as to why. In this study, we explore the relationship between affective sequences that have been previously explored in the literature and the activities students may engage in when interacting with Reasoning Mind, a blended learning system for elementary mathematics. The strongest correlations are found for students who shift from engaged concentration to frustration, making fewer actions in the system. While confusion is generally associated with positive patterns, and frustration and boredom have unexpectedly similar implications for student activity.

1   Introduction

Efforts to understand the relationship between affective states and learning have been underway in earnest since researchers were first able to model affective constructs through the use of sensors (e.g., Grafsgaard et al., 2014; Bosch et al., 2016) and interaction-based modeling (e.g., Baker & Ocumpaugh, 2014). Theoretical models of how students' experiences of affect change over time, such as those suggested by D'Mello & Graesser (2012), have guided much of the discussion, but empirical findings have also had considerable impact on the literature, with researchers suggesting, for example, that brief instances of confusion and frustration may have different effects than extended confusion and frustration (Liu et al., 2013), that different affective states tend to persist for different amounts of time (D'Mello & Graesser, 2011; Botelho et al., 2018), and that differences in affective sequences can have substantial impacts on learning (Andres et al., 2019). Work has suggested that affect can influence how students choose to interact with a learning system. For instance, researchers have found that negative affective states such as boredom tend to precede disengaged behaviors such as gaming the system (Baker et al., 2010; Sabourin et al., 2011).

The field has studied how affect manifests within AIED systems, and there have been several attempts to influence affect through the design of AIED systems (Arroyo et al., 2011; Grawemeyer et al., 2017, Karumbaiah et al., 2017). However, there has been relatively limited work to determine how the existing design of AIED systems interacts affect – i.e. how features not specifically intended to be affect-responsive nonetheless connect with the affect students experience. In one example, Slater and colleagues (2016) investigated how the textual features of math problems within the ASSISTments platform. They found relatively minor effects, perhaps due to the relatively minor differences in content they studied. However, for systems that alternate between very different pedagogies (e.g., shifting between games and workbook-style content), it is reasonable to expect that affective experiences may be influenced more substantially by the kinds of tasks students are being asked to complete and that affect may also drive students' choices.

This paper looks at the affective sequences of students using one such system, Imagine Learning's Reasoning Mind, which provides a blended learning curriculum in mathematics that engages elementary-aged students in tasks that range from basic instruction to challenge problems to speed games to non-academic activities. Specifically, this study looks at how the prevelance of affective patterns that have been studied in previous research—including D'Mello & Graesser (2012) and Andres et al. (2019)—correlates with different activities students may engage in when using Reasoning Mind.

2   Previous Research

Theorists working on the role of academic emotions have long suggested the need to understand both their antecedents and their consequences (see discussion in Pekrun, 2006; Pekrun and Linnenbrink-Garcia 2012). Therefore, it is becoming increasingly important to understand how student affect relates to their engagement within different kinds of learning environments.

Empirical investigations of academic emotions have produced a number of interesting results, including findings that it is better to be frustrated than bored (Baker et al., 2010). Researchers have also shown that both confusion and frustration appear to have Goldilocks effects on learning, where either too little or too much can be detrimental (Liu et al, 2013). These lead to important questions about when a system should intervene to resolve confusion and frustration and when a student should be allowed or even guided to shift into these affective states (Lehman & Graesser, 2015).

As AIED environments have developed as research tools that can provide fine-grained temporal data on the shifts in student cognition and emotion, there has emerged a growing interest in *affect dynamics*— the study of how affect shifts and develops over time. One of the most prominent models of affective dynamics comes to us from D'Mello and Graesser (2012). In this model, two sequences are hypothesized to be related to learning. The first, which is thought to encourage learning, involves a student cycling between engaged concentration and confusion (and back again). The second, which is thought to inhibit learning, involves a student cycling from engaged concentration to confusion to frustration, and finally to boredom.

Subsequent research has sometimes found evidence for the two cycles proposed by D'Mello and Graesser (2012), and in fact, efforts to promote confusion actually lead to positive learning outcomes (Lehman & Graesser, 2015), but the cycles themselves are less common than had originally been thought. A recent synthesis of published research on affective dynamics (Karumbaiah et al., 2018) shows that the relative frequency of the transitions captured by these sequences is often below chance. This raises important questions for those hoping to build interventions triggered by affective sequences. If these sequences are tied to learning outcomes, but they are unlikely to occur, we need to know what behaviors within a learning environment might mediate their appearance.

3   Reasoning Mind

3.1   Reasoning Mind

Imagine Learning's Reasoning Mind is an intelligent tutoring system for mathematics that was used by over 100,000 pre-K to 8th grade students, primarily in the Southern United States. Research has shown that Reasoning Mind is associated with higher state standardized test scores (Waxman & Houston, 2012) and engagement measures (Ocumpaugh et al., 2013).

Reasoning Mind activities are organized within the context of a virtual environment known as *RM City*, where students can navigate from building to building to participate in multiple modes, including: *City Landscape* (navigation page), *Guided* Study (theory and tests on math concepts), *Office* (teacher-assigned topics), *My Place* (students use points to purchase decorations for their virtual room) and *Game Room* (students participate in speed games, that require them to race against a speed meter, or solve math puzzles, like those found in the *Riddle Machine*). Content is further classified according to function and difficulty. *Theory* problems guide students to learn math concepts through animations and exercises. *Notes Test* check comprehension at the end of segments of theory material, requiring a review crucial concepts while reinforcing good note-taking practices. *A-level* problems reflect a fundamental understanding of basic material, while *B-level* problems may require multiple skills to complete multiple steps. *C-level* problems are conceptually advanced, requiring higher order thinking skills.

Fig. 1. Reasoning Mind's pedagogical agent, the Genie (left) and *RM City* (where *City Landscape Actions* happen, right)

4 Methods

4.1 Students

This study examines data from 796 Texas students who used Reasoning Mind as part of their regular 2nd to 6th-grade mathematics instruction during the 2017-18 school year.

4.2 Activities (Type of System Usage) Considered

Reasoning Mind students are offered a wide range of activities within the system. These include activities related to the primary modes of instruction, from the most basic problems (*A-level* Actions, *A-level* Accuracy, and *Guided Study* Actions) that all students must complete to more challenging problems (*B/C-level* Actions) which are often optional. They also include measures related to behaviors that vary in terms of their instructional content. For example, the speed-drills (*Game Room Actions*) review learning modules but do not provide instruction on new content. Meanwhile, the number of actions spent in the *RM City, (*or *City Landscape Actions*) tell us how often a student is switching between tasks, which may indicate either completion or dissatisfaction with the learning environment. Finally, we also consider the how students are spending the points they earn in Reasoning Mind's virtual store (*Items Purchased*).

Four activities chosen for this analysis represent a range in the type of usage that students using Reasoning Mind encounter: *Guided Study Actions, B/C-Level Actions, City Landscape Actions,* and *Items Purchased.* The first two represent actions that involve learning, while the latter may be less indicative of learning (although students are not able to purchase items unless they have earned points through positive learning behaviors). These activities also represent a range in the amount of choice a student has in whether they participate in that activity. Finally, they were carefully selected in order to exclude any actions that might have contributed to the BROMP-based interaction detectors developed by Kostyuk (2018). (A-level Actions, for example, are a part of several of Kostyuk's affect detectors, and so they were excluded in order to avoid circularity problems in the analysis.)

4.3 Affective Models and Sequences Considered

Models of Affective States. Affective states studied in this paper are modeled using detectors built by Kostyuk et al., (2018). These cross-validated, interaction-based detectors (e.g., Baker & Ocumpaugh 2014) were developed using the BROMP protocol for classroom observation (Ocumpaugh et al., 2015). Table 1 shows detectors for four academic emotions (boredom, confusion, engaged concentration, and frustration) and for off-task behavior. Although detector performance was relatively weak, the scale of data was sufficient to derive theoretically expected predictions for learning outcomes (Kostyuk et al., 2018). The distribution of affect predictions were re-scaled to bring the low incidence affective states back to the original distributions: Bor (13.7%), Eng (78.8%), Con (31.1%), and Fru (1.1%).

Table 1. Affective Models (from Kostyuk et al., 2018).

	Algorithm	AUC
Boredom (Bor)	Random Forest	0.60
Engaged Concentration (Eng)	Gradient Boosting Machine (GBM)	0.61
Confusion (Con)	Stepwise forward selection linear regression	0.53
Frustration (Fru)	Random Forest	0.65
Off-task (Off)	Stepwise forward selection linear regression	0.64

Affective Sequences. A considerable body of research has emerged using D'Mello's L, a likelihood metric for studying individual transitions (D'Mello & Graesser, 2012). However, this metric does not handle multi-state sequences, and recent research suggests that L requires corrections in order to be valid (Karumbaiah et al., 2019). Therefore, we take a different approach.

Instead, this study investigates affective sequences that were selected based on two previous publications. Specifically, we include the two cycles from D'Mello & Graesser's (2012) theoretical model, as well as include 16 sequences found to be important in Andres et al.'s (2019) exploration of affective dynamics in Betty's Brain, which (like this study) also made use of BROMP-based detectors.

Specifically, Andres et al. (2019) examined 12 "three-step" transitions where the first step was repeated (e.g., Eng-Eng-Bor, or Fru-Fru-Con) as well as four homogenous "four-step" transitions, which repeated the same affective state across the entire sequence (e.g., Bor-Bor-Bor-Bor). We also investigate two four-step sequences that involve off-task behavior (also modeled using Kostyuk's (2018) BROMP-based detectors), based on evidence that off-task behavior is more strongly negatively correlated with learning outcomes in Reasoning Mind than other interactive learning environments (Kostyuk et al., 2018). In total, this study investigates 20 affect sequences.

For each affect sequence, prevalence is computed using the method in Andres et al. (2019). Prevalence is the total number of times a pattern occurred within a given student's data divided by the total number of times it could have occurred in that data. The sequences involving only engaged concentration or confusion show the highest prevalence with Eng-Eng-Eng-Eng at 63.4% and Con-Con-Con-Con at 13.1%. This is followed by the sequences that have Bor with Eng-Eng-Bor at 6.2% and Bor-Bor-Con at 1.3%. Lastly, the sequences with frustration show the lowest prevalence with Eng-Eng-Fru at 0.29% and Eng-Con-Fru-Bor at 0.02%.

4.4 Analysis

Spearman's Rho (ρ) was used to correlate the prevalence of 20 affect sequences studied to the 4 types of student activities (types of usage) within Reasoning mind. Spearman's Rho is a non-parametric correlation coefficient that is often used when assumptions of normality cannot be applied across an entire data set. Because this analysis resulted in 80 separate statistical tests (20 affective states x 4 activity types within the system), Benjamini and Hochberg's (B&H; 1995) post-hoc FDR correction was applied. P-values in the results section are only marked as significant if they remained significant after the B&H procedure was applied.

5   Results

Results for the relationship between the prevalence of affective sequences and the different types of activities within Reasoning Mind are given in Table 2, where they are organized by the type of affective sequence being considered. These include (1) the D'Mello and Graesser cycles (both the facilitative and the inhibitory), (2) the sequences using the BROMP-based off-task detector, and then (3) the sequences studied by Andres et al., (2019). The latter is organized by the dominant affect in each sequence (i.e., the one that appears most frequently), with the homogenous four-step sequences (i.e., Eng-Eng-Eng-Eng) given in the order of the D'Mello and Graesser's inhibitory cycle (i.e., engaged, followed by confusion, followed by frustration, followed by boredom). However, readers will see that the results do not fully fit this model's predictions.

Table 2. Correlations between prevalence of affective sequences and types of student actions. Items that are non-significant after the B&H correction was applied are given in gray-scale.

	Learning Actions				Non-Learning Actions			
	Guided Study Problems		B/C-Level Problems		City Landscape		N of Items Purchased	
	Rho	p	Rho	p	Rho	p	Rho	p
Eng-Con-Con-Eng	-0.02	0.4	-0.02	0.54	0.03	0.32	-0.06	0.06
Eng-Con-Fru-Bor	-0.21	0	-0.1	0.01	-0.17	0	0.04	0.24
Off-Off-Off-Off	-0.23	0	-0.09	0.01	-0.19	0	0.03	0.45
Off-Off-Off-Eng	-0.13	0	-0.03	0.33	-0.18	0	-0.02	0.52
Eng-Eng-Eng-Eng	0.16	0	0.09	0	0.05	0.14	-0.05	0.18
Eng-Eng-Bor	-0.16	0	-0.08	0.01	-0.11	0	0.09	0.01
Eng-Eng-Con	0	0.83	0	0.97	0.06	0.1	-0.07	0.03
Eng-Eng-Fru	-0.29	0	-0.12	0	-0.31	0	0.11	0
Con-Con-Con-Con	-0.13	0	-0.07	0.03	-0.04	0.26	-0.02	0.66
Con-Con-Bor	-0.22	0	-0.13	0	-0.08	0.01	0.07	0.03
Con-Con-Eng	-0.05	0.13	-0.03	0.31	0.03	0.48	-0.04	0.21
Con-Con-Fru	-0.28	0	-0.13	0	-0.2	0	0.11	0
Fru-Fru-Fru-Fru	-0.24	0	-0.09	0.01	-0.18	0	0.07	0.05
Fru-Fru-Bor	-0.23	0	-0.1	0	-0.17	0	0.05	0.18
Fru-Fru-Con	-0.28	0	-0.12	0	-0.21	0	0.08	0.02
Fru-Fru-Eng	-0.27	0	-0.1	0	-0.23	0	0.08	0.02
Bor-Bor-Bor-Bor	-0.26	0	-0.13	0	-0.13	0	0.07	0.03
Bor-Bor-Con	-0.25	0	-0.13	0	-0.09	0	0.1	0
Bor-Bor-Eng	-0.21	0	-0.12	0	-0.11	0	0.06	0.08
Bor-Bor-Fru	-0.28	0	-0.14	0	-0.17	0	0.07	0.02

5.1   Off-Task Sequences

Two sequences that were constructed using the BROMP-based off-task detectors were included in these analyses in order to explore findings suggesting that off-task behavior is correlated more negatively in learning in Reasoning Mind than in other interactive learning environments (Kostyuk et al., 2018). The first, Off-Off-Off-Off, was negatively correlated with three of the activity types: *City Landscape Actions* (ρ = -.19),

Guided Study Actions (ρ = -.23), and *B/C-Level Actions* (ρ = -.09). Interestingly, this effect was nearly twice as strong for *Guided Study Actions* as it was for *B- and C-Level Actions*, which may be because students spend less time in that mode overall. Off-Off-Off-Off was not, however, significantly correlated with the *Number of Items Purchased*, perhaps because students can only purchase items if they spend enough time on task to earn the points to do so.

When we changed the fourth step from off-task to engaged, only two of the correlations remained significant. While the *City Landscape Actions* correlation only changed slightly (ρ = -.18), the *Guided Study Actions* correlation was half as strong for Off-Off-Off-Eng (ρ = -.13) as it was for the homogenous four-step sequence. This suggests that even a slight reduction in the duration of off-task behavior improves the outcomes for Reasoning Mind students, in line with Pardos et al.'s (2014) findings.

5.2   D'Mello & Graesser's (2012) Sequences

As discussed above, D'Mello and Graesser (2012) they theorized a number of different transitions between affective states that were thought to be relevant to learning. In this section, we explore results related to their facilitative and inhibitory sequences.

D'Mello and Graesser's facilitative sequence, in which a student cycles between engaged concentration and confusion, is operationalized here as Eng-Con-Con-Eng. As the results in Table 2 show, this sequence has no statistically-significant relationship to any of the activities within Reasoning Mind. Likewise, the three-step patterns related to this sequence, Eng-Eng-Con and Con-Con-Eng, show similar results. The former is only weakly significantly related to the *Number of Items Purchased* (ρ = -0.07), where it shows the only negative correlation with that activity. The latter, like the main D'Mello and Graesser facilitative sequence, has no significant relationships with any of the activity types.

D'Mello and Graesser's inhibitory sequence is operationalized here as Eng-Con-Fru-Bor. Its results are similar to Off-Off-Off-Eng, as it shows non-significant relationships with two of the action types (*B/C-Level Actions* and *Number of Items purchased*) and negative relationships for the other two (ρ = -.17 for *City Landscape Actions* and ρ = -.21 for *Guided Study Actions*).

5.3   Engaged Concentration Sequences

Four sequences in this study are composed primarily of engaged concentration, and these sequences demonstrate some of the most divergent results. Much of this divergence is driven by results from Eng-Eng-Fru, which, when compared to all other affective sequences in this study, shows the strongest (negative) correlations with *City Landscape Actions* (ρ = -.31) and with *Guided Study Actions* (ρ = -.29). Eng-Eng-Fru also shows one of the strongest correlations with *B/C-Level Actions* (ρ = -.12 compared to max ρ = -.14). These results are stronger than those of Eng-Eng-Bor: *City Landscape Actions* (ρ = -.11), *Guided Study Actions* (ρ = -.16), and *B/C-Level Actions* (ρ = -.08). Compared to Eng-Eng-Con, which does not have significant relationships with these actions, the results for Eng-Eng-Fru and Eng-Eng-Bor suggest that students who skip

confusion when transitioning from engaged concentration have lower levels of positive behaviors.

Skipping confusion (i.e., not going through the Eng-Eng-Con transition) and going to boredom (i.e., Eng-Eng-Bor) or frustration (i.e., Eng-Eng-Fru) also shows differences for the *Number of Items Purchased.* While the sequence with confusion shows a negative relationship with this action type (ρ = -.07), the sequences with boredom and frustration are positive (ρ = .09, .11, respectively).

Finally, Eng-Eng-Eng-Eng is significant for only two activity types. Notably, in contrast to the results for Eng-Eng-Fru and Eng-Eng-Bor, Eng-Eng-Eng-Eng is positively correlated with *Guided Study Actions* (ρ = .16), and *B/C-Level Actions* (ρ = .09). In fact, these are the only positive correlations in the whole study that are not related to the *Number of Items Purchased*.

5.4 Confusion Sequences

Sequences involving confusion also show some divergence in their relationships with activity types, though not as extreme as those for engaged concentration. Most of the significant relationships between sequences composed primarily of confusion and activity types are negative. As with the results for the engaged concentration sequences, the exceptions to this pattern are for the *Number of Items Purchased*, which may sometimes be driven by a desire to go off-task, but also require a student to have successfully completed a significant amount of work.

In general, these results show that Con-Con-Con-Con is weakly negatively correlated to learning activities (ρ = -.16 for *Guided Study Actions* and ρ = -.07 for *B/C-Level Actions*). (The relationship between Con-Con-Con-Con and *City Landscape Actions* is not significant.) The relationships for Con-Con-Bor and Con-Con-Fru are slightly stronger: *City Landscape Actions* (ρ = -.08, -.20, respectively), *Guided Study Actions* (ρ = -.22, -.28, respectively), and *B/C-Level Actions* (ρ = -.13, -.13, respectively). These results are not inconsistent with findings that confusion is beneficial to learning (i.e. Lehman & Graesser 2015; Liu et al., 2013), but contrast with findings that suggest that it is better to be frustrated than bored (i.e. Baker et al., 2010). Interestingly, Con-Con-Eng is not significantly related to these learning activities. While this result is surprising, it is consistent with the results for the facilitative D'Mello and Graesser sequence.

5.5 Frustration and Boredom Sequences

Nearly all of the relationships between sequences composed primarily of frustration and activity types are significantly significant, and the same is true for those sequences composed primarily of boredom. As with the results for confusion sequences, these show negative relationships with *City Landscape Actions*, *Guided Study Actions*, and *B/C-Level Actions* and positive relationships with the *Number of Items Purchased*.

For *City Landscape Actions* and *Guided Study Actions,* the relationships with frustration sequences tend to be stronger than those with boredom sequences, which also contradicts the idea that frustration is better for learning than boredom. However, this difference is small, and for *B/C-Level Actions*, that relationship is reversed. That is, frustration and boredom both appear to be negatively associated with learning-related

activities (and positively associated with non-learning activities), but overall there is little separation between them.

6   Conclusion

In this paper, we investigate how affective patterns connect to student activity choices within Reasoning Mind. We find that the strongest patterns involve students who shift from engaged concentration to frustration. These students interact less with the environment than other students, although they do spend more of their points purchasing virtual decorations for their *My Place* room. We also find that confusion is generally associated with positive behavioral patterns. Somewhat surprisingly, frustration and boredom generally correlate to the same usage patterns. Also, inhibitive sequences emerging from D'Mello and Graesser's (2012) theoretical model are relatively weakly associated with activities within the system, while the facilitative sequence is not significantly associated with any of the activities considered in this study.

The findings here, in concert with Karumbaiah et al.'s (2018) research synthesis of affect dynamics research, which found that few patterns were more likely than chance across studies, potentially raise concerns about the generalizability of findings from previous research. However, they also point to the need for a more comprehensive understanding of the relationship between affective dynamics, behavioral patterns, and learning outcomes, as these findings suggest that these relationships may not be as straight-forward as we once thought.

Overall, the findings here suggest that there are relationships between student affect and the activities they engage in within a learning system. It is not entirely clear what the direction of the effects is from our current evidence – are students with specific affective patterns choosing different activities? Or are the activities driving the affective patterns? A more in-depth temporal analysis may be able to shed more light on this issue, but these issues are complex; affect may develop, and shape interaction choices but also shape the future affect itself (i.e. D'Mello & Graesser, 2012; Botelho et al., 2018). What our findings indicate is that usage choices and affect are connected in many ways.

Overall, these findings point to the need for a more comprehensive understanding of the relationship between affective dynamics, behavioral patterns, and learning outcomes, as the findings here suggest that the relationship may not be as straight-forward as might have been thought. Fully understanding these interconnections – and the role that the design of AIED systems plays – is an important area for future research, and an important step towards AIED systems that are fully sensitive to the shifts in students' affect and how these shifts in turn impact behavior.

References

1. Andres, J.M.A.L., Ocumpaugh, J., Baker, R., Slater, S., Paquette, S., Jiang, Y., Karumbaiah, S., Bosch, N., Munshi, A., Moore, A., Biswas, G. (2019) Affect Sequences and Learning in Betty's Brain. *Proceedings of the 9th International Learning Analytics and Knowledge Conference*, 383-390.

2. Arroyo, I., Woolf, B. P., Cooper, D. G., Burleson, W., & Muldner, K. (2011, July). The impact of animated pedagogical agents on girls' and boys' emotions, attitudes, behaviors and learning. In *2011 IEEE 11th International Conference on Advanced Learning Technologies* (pp. 506-510). IEEE.

3. Baker, R.S.J.d., 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 (4), 223-241.

4. Baker, R.S.J.d., 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 (4), 223-241.

5. Baker, R.S.J.d., Ocumpaugh, J. (2014) Interaction-Based Affect Detection in Educational Software. In R.A. Calvo, S.K. D'Mello, J. Gratch, A. Kappas (Eds.), *The Oxford Handbook of Affective Computing*. Oxford, UK: Oxford University Press.

6. Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), pp. 289-300.

7. Bosch, N., D'Mello, S.K., Ocumpaugh, J., Baker, R.S., Shute, V. (2016) Using video to automatically detect learner affect in computer-enabled classrooms. *ACM Transactions on Interactive Intelligent Systems*, 6 (2).

8. Botelho, A.F., Baker, R., Ocumpaugh, J., Heffernan, N. (2018) Studying Affect Dynamics and Chronometry Using Sensor-Free Detectors. *Proceedings of the 11th International Conference on Educational Data Mining*, 157-166.

9. Crossley, S., Bradfield, F., Ocumpaugh, J., Dascalu, M., Labrum, M., & Baker, R. S. (2018). Modeling math identity and math success through sentiment analysis and linguistic features. In International Conference on Educational Data Mining, pages 11–20. EDM Society Press.

10. DeFalco, J.A., Rowe, J.P., Paquette, L., Georgoulas-Sherry, V., Brawner, K., Mott, B.W., Baker, R.S., Lester, J.C. (2018) Detecting and Addressing Frustration in a Serious Game for Military Training. *International Journal of Artificial Intelligence and Education, 28* (2), 152-193.

11. D'Mello, S., & Graesser, A. (2011). The half-life of cognitive-affective states during complex learning. *Cognition & Emotion*, *25*(7), 1299-1308.

12. D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2), 145-157

13. Grafsgaard, J., Wiggins, J., Boyer, K. E., Wiebe, E., & Lester, J. (2014, July). Predicting learning and affect from multimodal data streams in task-oriented tutorial dialogue. In *Educational Data Mining 2014*.

14. Grawemeyer, B., Mavrikis, M., Holmes, W., Gutiérrez-Santos, S., Wiedmann, M., & Rummel, N. (2017). Affective learning: improving engagement and enhancing learning with affect-aware feedback. *User Modeling and User-Adapted Interaction*, *27*(1), 119-158.

15. Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American journal of Physics, 66(1), 64-74.
16. Karumbaiah, S., Lizarralde, R., Allessio, D., Woolf, B., Arroyo, I., & Wixon, N. (2017). Addressing Student Behavior and Affect with Empathy and Growth Mindset. *International Educational Data Mining Society*.
17. Karumbaiah, S., Andres, J. M. A. L., Botelho, A. F., Baker, R. S., & Ocumpaugh, J. (2018). The implications of a subtle difference in the calculation of affect dynamics. In *26th International Conference for Computers in Education*.
18. Karumbaiah, S., Baker, R.S., Ocumpaugh, J. (2019) The Case of Self-Transitions in Affective Dynamics. *Proceedings of the 20th International Conference on Artificial Intelligence in Education*, 172-181.
19. Kostyuk, V., Almeda, M.V., Baker, R.S. (2018) Correlating Affect and Behavior in Reasoning Mind with State Test Achievement. In *Proceedings of the International Conference on Learning Analytics and Knowledge*, 26-30.
20. Lehman, B., & Graesser, A. (2015, June). To resolve or not to resolve? That is the big question about confusion. In *International Conference on Artificial Intelligence in Education* (pp. 216-225). Springer, Cham.
21. Liu, Z., Pataranutaporn, V., Ocumpaugh, J., Baker, R.S.J.d. (2013) Sequences of Frustration and Confusion, and Learning. *Proceedings of the 6th International Conference on Educational Data Mining*, 114-120.
22. Mulqueeny, K., Kostyuk, V., Baker, R.S., Ocumpaugh, J. (2015) Incorporating Effective e-Learning Principles to Improve Student Engagement in Middle-School Mathematics. *International Journal of STEM Education, 2* (15).
23. Ocumpaugh, J., Baker, R.S., Rodrigo, M.M.T. (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.
24. Ocumpaugh, J., Baker, R. S., Gaudino, S., Labrum, M. J., and Dezendorf, T. (2013). Field observations of engagement in reasoning mind. In International Conference on Artificial Intelligence in Education, pages 624–627. Springer.
25. Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., & Gowda, S. M. (2014). Affective States and State Tests: Investigating How Affect and Engagement during the School Year Predict End-of-Year Learning Outcomes. *Journal of Learning Analytics*, *1*(1), 107-128.
26. Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. Educational Psychology Review, 18(4), 315-341.
27. Sabourin, J., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011, June). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In *International Conference on Artificial Intelligence in Education* (pp. 534-536). Springer, Berlin, Heidelberg.
28. Slater, S., Ocumpaugh, J., Baker, R., Scupelli, P., Inventado, P.S., Heffernan, N. (2016) Semantic Features of Math Problems: Relationships to Student Learning and Engagement. *Proceedings of the 9th International Conference on Educational Data Mining.*, 223-230.
29. Waxman, H. C. and Houston, W. R. (2012). Evaluation of the 2010-2011 Reasoning Mind program in Beaumont ISD. Unpublished manuscript.