Affect Dynamics and Behavioral Patterns in Intelligent Learning Environments

Patricia A. Jaques[1], **Felipe de Morais**[2], **Tiago R. Kautzmann**[3]

[1]Universidade Federal do Paraná - PPGInf/UFPR
Universidade Federal de Pelotas - PGCC/UFPEL
patricia@inf.ufpr.br

[2]Universidade do Vale do Rio dos Sinos - PPGCA/UNISINOS
felipmorais@unisinos.br

[3]Instituto Federal do Paraná (IFPR)
tiago.kautzmann@ifpr.edu.br

Abstract. *This research examines the interplay between emotions and learning behaviors in Intelligent Learning Environments (ILEs) using data from Brazilian students. We analyze associations between emotions—including confusion, frustration, boredom, and engagement—and learning behaviors, focusing on the influence of gender and the duration of these emotions and behaviors. Data from 30 students in a math step-based tutoring system were annotated by human coders, with transition probabilities assessed using the L metric. Key findings indicate that engagement and confusion correlate with active task involvement, while boredom and frustration are linked to task withdrawal. Additionally, gender-specific dynamics between emotions and behaviors were observed. These results highlight the potential to use the observed relationships between emotions and behaviors to enhance personalized learning experiences and inform the design of adaptive learning environments.*

Resumo. *Esta pesquisa examina a interação entre emoções e comportamentos de aprendizagem em Ambientes de Aprendizagem Inteligentes (ILEs) utilizando dados de estudantes brasileiros. Analisamos as associações entre emoções—incluindo confusão, frustração, tédio e engajamento—e comportamentos de aprendizagem, focando na influência do gênero e na duração dessas emoções e comportamentos. Dados de 30 estudantes em um sistema tutor baseado em etapas de matemática foram anotados por codificadores humanos, com probabilidades de transição avaliadas utilizando a métrica L. Os principais achados indicam que o engajamento e a confusão correlacionam-se com o envolvimento ativo nas tarefas, enquanto o tédio e a frustração estão ligados à retirada das tarefas. Além disso, foram observadas dinâmicas específicas de gênero entre emoções e comportamentos. Esses resultados destacam o potencial de usar as relações observadas entre emoções e comportamentos para aprimorar experiências de aprendizagem personalizadas e informar o design de ambientes de aprendizagem adaptativos.*

1. Introduction

Understanding affective states is crucial in educational contexts, as they significantly influence cognition and learning outcomes. Emotions such as confusion, frustration, bore-

DOI: 10.5753/sbie.2024.241613

dom, and curiosity, play a pivotal role in shaping attention, motivation, cognitive strategies, and self-regulation in learners [Pekrun 2014]. This is particularly pertinent in Intelligent Learning Environments (ILEs), where the dynamics of emotional experiences can profoundly impact pedagogical effectiveness [Graesser et al. 2014].

The model proposed by D'Mello and Graesser [Graesser and D'Mello 2011] highlights that unmanaged confusion can lead to frustration and boredom, negatively impacting learning. Subsequent studies support that confusion often transitions to negative emotions [Fredrickson 1998]. Karumbaiah et al.'s [Karumbaiah et al. 2021] reevaluation of affect dynamics, employing an adjusted L metric, has illuminated the influence of cultural context on emotional transitions. Moreover, gender and the duration of emotions have been identified as significant factors in learning, with evidence pointing to gender-specific variations in emotional experiences and their impacts on education [Morais and Jaques 2023]. Notably, research indicates that female students may experience higher levels of anxiety and negative emotions compared to their male counterparts, who exhibit lower levels of anxiety and experience emotions such as pride, enjoyment, and boredom differently [Hembree 1988, Zeidner 1998, Frenzel et al. 2007].

The relationship between emotions and learning is further nuanced by factors such as behavioral patterns in students. Learning-oriented behaviors, which encompass students' actions and attitudes towards learning tasks, have been shown to be significantly associated with learning outcomes [Perkins 1965]. Research has linked specific emotions with learning behaviors and outcomes. For instance, boredom has been associated with poorer learning and problematic behaviors like gaming the system, while off-task behaviors can sometimes reengage students and alleviate negative emotions [Baker 2011, Baker et al. 2011, Sabourin et al. 2011, Baker et al. 2007].

This study explores the complex interplay between emotions and behaviors in education, focusing on the duration of emotions and gender influence. While previous research examined transitions between emotional and behavioral states, this study innovatively emphasizes how the sustained presence of an emotion can influence behaviors and how prolonged emotional experiences affect the likelihood of specific behavioral outcomes. By focusing on Brazilian students, this study addresses a gap in exploring diverse cultural backgrounds in educational research, particularly in Latin America. It offers insights into adapting universal findings to different cultural settings, highlighting the need to customize educational strategies for each student population.

Our methodology utilizes data collection from Brazilian students through PAT2Math, a step-based Math Intelligent Tutoring System (ITS) [Jaques et al. 2013]. We focus on annotating students' learning emotions (engagement, confusion, frustration, and boredom) and behaviors (on-task, off-task, participating in task-related conversations, using the system but not on a specific task, and accessing external educational resources) using the EmAP-ML protocol [Morais et al. 2019]. The analysis includes evaluating the transition probabilities between emotions and learning behaviors using the adjusted L statistic [Karumbaiah et al. 2021]. The analysis also considers patterns related to gender and the duration of emotions. Statistical tests are employed to determine the significance of these transitions. This study is guided by four primary research questions:

- RQ1: How are students' behaviors and emotional states in an ILE associated with each other in the context of Brazilian learners?

- RQ2: What are the gender-related patterns in the association between emotional states and learning behaviors in ILEs?
- RQ3: How does the duration of emotions relate to learning behaviors and vice versa, particularly in the context of Brazilian students?
- RQ4: How do the patterns of association between emotional states and learning behaviors in Brazilian students compare with international findings? Do these patterns replicate, diverge, or reveal unique insights when contrasted with established international patterns?

Addressing these questions will enhance our understanding of emotional dynamics in education and inform culturally sensitive teaching strategies. By including data from Latin American students, the study explores unique demographic aspects and broadens the generalization of research findings, which have been mostly based on North American and Filipino contexts. This approach ensures that insights reflect a wider range of cultural and contextual nuances in learning.

2. Related Work

The model of affect dynamics proposed by D'Mello and Graesser [D'Mello and Graesser 2012] stands as a significant contribution to the field. This model illustrates the progression of student emotions during learning, highlighting the critical role of cognitive disequilibrium and confusion. The model suggests that confusion, if appropriately regulated, can lead to re-engagement and learning. However, unaddressed confusion may devolve into frustration or boredom, stymieing the learning process. This model underlines two essential aspects: the necessity for ILEs to challenge students, fostering critical thinking, and the importance of regulating negative emotions like frustration and boredom through pedagogical interventions.

A recent analysis by [Karumbaiah et al. 2021] reviewed studies applying the L metric in affect dynamics. Their re-analysis, considering an adjusted version of this metric, revealed notable differences in affect transitions across different cultural contexts, namely the USA and the Philippines. This finding is significant as it suggests that cultural and contextual factors play a crucial role in affect dynamics, a hypothesis that our study seeks to explore further in the Brazilian context.

The relationship between affect and behaviors, while extensively studied, often lacks a focus on affect dynamics. Notable works in this area include studies by [Rodrigo et al. 2009, Pardos et al. 2014, Fancsali 2014, Kostyuk et al. 2018]. However, these studies do not integrate the affect dynamics perspective. In contrast, Baker et al.'s research in different ILEs across two countries delves into the interplay between students' affect and behaviors, highlighting the persistence of boredom and its impact on student engagement [Baker et al. 2010]. Further expanding on this, Baker and colleagues [Baker et al. 2011, Baker et al. 2007] investigated the co-occurrence and sequences of affect and behaviors in various educational settings. Their findings emphasize the intricate relationship between specific emotions, like boredom and confusion, and student behaviors such as being off-task or gaming the system.

The study by Morais et al. [Morais and Jaques 2023] offers a comprehensive analysis of the dynamics of emotions in Brazilian students using a Math ITS. The research highlights how affective transitions vary based on gender and the duration of emotions, uncovering distinct emotional response patterns between male and female students. Additionally, it presents the importance of considering the duration of emotions, illustrating how short and long-term emotions differently impact learning processes. These findings emphasize the necessity for educational technology designs that are culturally and emotionally sensitive to the context of Brazilian students.

Upon synthesizing these studies, several key conclusions emerge: while the affect dynamics model by D'Mello and Graesser offers substantial contributions, its applicability across different contexts is not universal, highlighting the significant influence of students' cultural backgrounds on affect dynamics. Our research pushes the boundaries of understanding the intricate relationship between emotions and behaviors within educational environments by emphasizing the previously unexplored roles of emotion duration and gender. Our focus on the duration of emotions aims to unravel how the prolonged presence of a specific emotional state might trigger certain behaviors and, conversely, how the extended experience of an emotion might change its likelihood of leading to specific behavioral outcomes.

Focusing on Brazilian students, this study fills a gap in educational research on Latin American contexts. It provides insights into how universal findings should be adjusted for different cultures, emphasizing tailored educational strategies. No prior research has developed a statistically significant affect dynamics model for Brazilian students or comprehensively considered gender, learning behaviors, and emotion duration together.

3. Research Method

This section outlines the methodological approach adopted in our study to investigate the interplay between students' emotional states and their learning behaviors in the context of an ILE. The methodological design encompasses a data collection, emotion and behavior annotation, and statistical analysis, further elaborated in the following subsections.

3.1. Dataset of Emotions and Behaviors

Our study utilized the dataset from [Morais and Jaques 2023], which includes observations from 30 students—16 females and 14 males, aged 12 to 13, enrolled in two seventh-grade classes at a private elementary school. Participants engaged in ten sessions with PAT2Math, a web-based Math Intelligent Tutoring System (ITS), each session lasting approximately 40 minutes. The version of PAT2Math used includes gamification elements such as rankings and scores. Webcam recordings captured facial expressions, vocal nuances, and computer screen interactions in detail, providing a rich view of student emotions and interactions with the ITS. The annotation of emotions and behaviors was performed by the study's researchers, who manually recorded these observations in five-second intervals, known as clips. The annotations complied with the EmAP-ML protocol, employing the tool provided by the protocol [Morais et al. 2019], allowing for the documentation of one to two emotions and one to two behaviors in each clip.

Building on this foundation, the annotation process categorized emotions into four primary learning-centered types: **engagement**, **confusion**, **frustration**, and **boredom**, with all other emotions being classified as **other**. Additionally, it identified five distinct student behaviors: **On-task**, indicating direct focus on the current task; **On-task-conversation**, highlighting task-related discussions with teachers or peers; **On-task-resource**, signifying the utilization of external resources such as notes or notebooks; **On-system**, denoting interactions with the ITS not directly related to a specific task; and **Off-task**, referring to instances where students were not actively participating in the designated system task.

3.2. L Metric and Statistical Procedure

Our study employed the adjusted L metric, as refined by Matayoshi and Karumbaiah [Matayoshi et al. 2020] from D'Mello and Graesser's original concept [D'Mello et al. 2007], to assess transitions between affective states (confusion, frustration, boredom, engagement, and others) and behaviors (On-task, On-task-conversation, On-task-resource, On-system, and Off-task). This metric, which ranges from $-\infty$ to 1, accounts for the removal of self-transitions, adjusting the chance-level transition base rate from zero to 0.0625.

For statistical analysis, we applied two primary methods. The *t-test* evaluated the significance of emotion-behavior transitions, while the Wilcoxon rank-sum test, chosen based on Shapiro-Wilk normality test results, compared behavior duration times across emotions and genders, suitable for our non-normally distributed data. To address potential false positives from multiple comparisons with the *t-test*, we applied the Benjamini-Hochberg (BH) post hoc correction with an alpha level of .05, ensuring robustness in our findings related to affect dynamics analysis [Karumbaiah et al. 2021, Morais and Jaques 2023].

4. Results

To answer RQ1, we assessed whether student behaviors predict emotions or vice versa, and calculated co-occurrence probabilities between each pair of emotions and behaviors. Table 1 shows the probability that an emotion predicts a behavior's onset or continuation, while Table 2 presents the probability that a behavior predicts the following emotion. Table 3 exhibits the co-occurrence probabilities of each emotion-behavior pair. To answer RQ2, we analyzed all transitions considering students' gender. Tables 1, 2, and 3 present results based on gender: male (M), female (F), and all (A) students. Throughout the tables in this section, we use the following acronyms to represent the emotional states: ENG for engagement, BOR for boredom, CON for conversation, OTH for other emotions, and FRU for frustration. The first value in each table cell indicates the L metric result, the second value (in parentheses) is the adjusted *p-value* per the BH post hoc method, and the third value (in square brackets) shows the number of transitions or co-occurrences. To ensure reliability, the L metric analysis was limited to relationships with more than 10 occurrences, avoiding the disproportionate influence of rare events. Thus, L statistics and significance are not shown for occurrences below 11. Bold values indicate statistically significant transitions per the BH post hoc method. As described in Section 3.2, an L value below 0.0625 is considered low probability, while a value above this threshold indicates a higher transition probability than chance.

4.1. Dynamics between Emotions and Behaviors

According to Table 1, engagement significantly predicts an increase in on-task behavior (L = .26), on-system behavior for males (L = .17), and off-task behavior (L = .12), while

Table 1. Statistical results of the dynamics between emotions and behaviors.

	G	On-task	On-system	On-task Resource	Off-task	On-task Conver.
ENG	A	**.26 (.00) [193]**	**.15 (.00) [95]**	.04 (.11) [31]	**.12 (.03) [73]**	**.03 (.04) [22]**
	F	**.26 (.00) [106]**	.13 (.05) [49]	.06 (.68) [21]	.12 (.07) [36]	.04 (.41) [14]
	M	**.27 (.00) [87]**	**.17 (.01) [46]**	[10]	.12 (.22) [37]	[8]
CON	A	**.40 (.00) [110]**	**-.02 (.00) [13]**	.04 (.33) [17]	.10 (.39) [32]	.14 (.12) [28]
	F	**.43 (.00) [65]**	.01 (.13) [11]	[6]	.09 (.49) [22]	.11 (.44) [13]
	M	**.37 (.01) [45]**	[2]	.08 (.74) [11]	[10]	.16 (.16) [15]
FRU	A	**.47 (.01) [20]**	[3]	[1]	[5]	[7]
	F	[8]	[1]	[1]	[3]	[5]
	M	**.62 (.02) [12]**	[2]	[0]	[2]	[2]
BOR	A	**.61 (.00) [38]**	[7]	[2]	[8]	[0]
	F	**.60 (.00) [25]**	[5]	[2]	[4]	[0]
	M	**.61 (.01) [13]**	[2]	[0]	[4]	[0]
OTH	A	**.58 (.00) [125]**	.05 (.74) [33]	[10]	.03 (.26) [24]	**.01 (.00) [11]**
	F	**.56 (.00) [55]**	.04 (.47) [16]	[4]	.07 (.95) [16]	[9]
	M	**.60 (.00) [70]**	.08 (.64) [17]	[6]	[8]	[2]

it reduces the likelihood of engaging in on-task-conversation. Confusion leads to an increase in on-task behavior (L = .40) and is negatively associated with on-system behavior (L = -.02). Frustration is a strong predictor of on-task behavior in males (L = .62). Boredom notably increases the likelihood of on-task behavior (L = .61). Emotions classified as ‘other’ are also found to precede on-task behavior (L = .58), indicating a general trend towards task engagement irrespective of the specific emotional state.

4.2. Dynamics between Behaviors and Emotions

Table 2 shows the probability of a behavior predicting the onset of an emotion. Again, the results were computed according to the student’s gender. According to Table 2, engaging in on-task behavior significantly increases the likelihood of experiencing engagement (L = .16), confusion (L = .17), and other emotional states (L = .20), showing a clear link between this behavior and a range of emotional responses. Notably, being on-task decreases the likelihood of experiencing frustration for male students (L = .03) and boredom for both genders (L = .02). Interacting with the ITS without direct focus on a specific task (on-system behavior) strongly predicts engagement (L = .65), while also decreasing the likelihood of confusion (L = -.05) and boredom (L=.02). Utilizing external resources (on-task-resource behavior) strongly predicts engagement (L = .60). Conversely, being off-task is associated with an increased probability of engagement (L = .45). Lastly, engaging in on-task conversation is linked to increased engagement for female students (L = .43).

4.3. Co-occurrence between Emotions and Behaviors

Table 3 shows the probability of co-occurrence of one emotion and one behavior, i.e., the likelihood of the pair emotion and behavior occurring simultaneously. According to Table 3, the behavior of being on-task is significantly associated with experiencing engagement (L = .64) and confusion (L = .24). Notably, being on-task shows a reduced likelihood of coinciding with boredom (L = .04) and other emotions (L = .03). On-system behavior, or interacting with the ITS without a direct focus on tasks, is predominantly associated with

Table 2. Statistical results of the dynamics between behaviors and emotions.

	G	ENG	CON	FRU	BOR	OTH
On-task	A	**.16 (.00) [173]**	**.17 (.00) [131]**	**.03 (.01) [29]**	**.02 (.00) [28]**	**.20 (.00) [133]**
	F	**.16 (.01) [96]**	**.17 (.00) [71]**	.04 (.09) [15]	**.02 (.00) [17]**	**.20 (.00) [65]**
	M	**.17 (.05) [77]**	**.17 (.01) [60]**	**.03 (.03) [14]**	**.02 (.03) [11]**	**.20 (.00) [68]**
On-system	A	**.65 (.00) [96]**	**-.05 (.00) [11]**	[3]	**.02 (.03) [12]**	.04 (.55) [27]
	F	**.59 (.00) [47]**	[8]	[0]	[10]	.06 (.97) [17]
	M	**.71 (.00) [49]**	[3]	[3]	[2]	[10]
On-task res.	A	**.60 (.00) [40]**	.13 (.42) [14]	[0]	[1]	[7]
	F	**.68 (.00) [22]**	[7]	[0]	[1]	[4]
	M	**.52 (.03) [18]**	[7]	[0]	[0]	[3]
Off-task	A	**.45 (.00) [67]**	.06 (.95) [29]	[0]	.08 (.65) [15]	.07 (.91) [27]
	F	**.53 (.00) [39]**	.12 (.37) [22]	[0]	[8]	[8]
	M	**.35 (.04) [28]**	[7]	[0]	[7]	.20 (.19) [19]
On-task conv.	A	**.41 (.00) [32]**	.09 (.63) [14]	[4]	[1]	.12 (.47) [14]
	F	**.43 (.01) [19]**	[7]	[3]	[1]	[10]
	M	.38 (.13) [13]	[7]	[1]	[0]	[4]

engagement for male students (L = .28) and with the emotion classified as other (L = .59) across genders. For female students, on-system behavior also significantly coincides with boredom (L = .22). The use of external resources (on-task resource) strongly correlates with high student engagement (L = .82), indicating that engaged students actively seek additional support, like using notebooks for problem-solving or reviewing notes. Conversely, off-task behavior is associated with other emotions (L = .76). Lastly, engaging in on-task conversations is uniquely associated with experiencing confusion (L = .66), highlighting the role of social interaction in the emergence of this emotional state.

Table 3. Statistical results of the co-occurrence between emotions and behaviors.

	G	ENG	CON	FRU	BOR	OTH
On-task	A	**.64 (.00) [308]**	**.24 (.00) [127]**	.05 (.24) [25]	**.04 (.01) [19]**	**.03 (.00) [15]**
	F	**.66 (.00) [170]**	**.24 (.00) [68]**	.04 (.21) [12]	[10]	[4]
	M	**.61 (.00) [138]**	**.24 (.00) [59]**	.05 (.68) [13]	[9]	.05 (.32) [11]
On-system	A	**.18 (.02) [30]**	.05 (.76) [12]	[3]	**.15 (.01) [24]**	**.59 (.00) [80]**
	F	.10 (.42) [13]	[10]	[2]	**.22 (.00) [19]**	**.56 (.00) [38]**
	M	**.28 (.03) [17]**	[2]	[1]	[5]	**.62 (.00) [42]**
On-task res.	A	**.82 (.00) [47]**	.14 (.12) [12]	[0]	[1]	[2]
	F	**.85 (.00) [27]**	[4]	[0]	[1]	[2]
	M	**.78 (.00) [20]**	[8]	[0]	[0]	[0]
Off-task	A	.08 (.51) [15]	[9]	[4]	.07 (.86) [11]	**.76 (.00) [99]**
	F	[9]	[9]	[2]	[6]	**.69 (.00) [51]**
	M	[6]	[0]	[2]	[5]	**.84 (.00) [48]**
On-task conv.	A	.15 (.11) [14]	**.66 (.00) [40]**	[4]	[0]	[7]
	F	[7]	**.72 (.00) [26]**	[2]	[0]	[5]
	M	[7]	**.59 (.01) [14]**	[2]	[0]	[2]

4.4. Duration of Behaviors and Emotions

To answer our RQ3, we computed the results of the duration time of the students in the learning behaviors and in the emotions when considering the behaviors. First, Figure 1 shows the duration time distribution in behaviors according to the student's gender. For the analysis of the duration of emotions and behavior distributions, we conducted the Shapiro-Wilk normality test to assess normal distribution. Given that the data did not exhibit a normal distribution, we utilized the Wilcoxon rank-sum test for comparison. Therefore, all *p-value* (p) are based on the Wilcoxon rank-sum test with a $.05$ significance level. Also, for each distribution, we described the sample mean ($\bar{X}$), the sample median (M), and the sample standard deviation (SD).

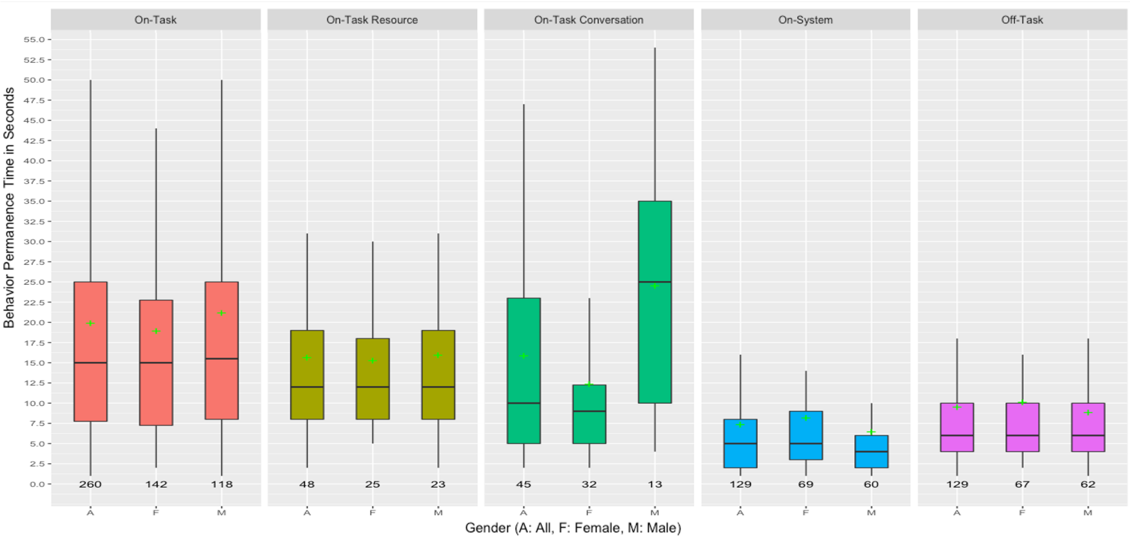

Figure 1. Distribution of the duration time in behaviors.

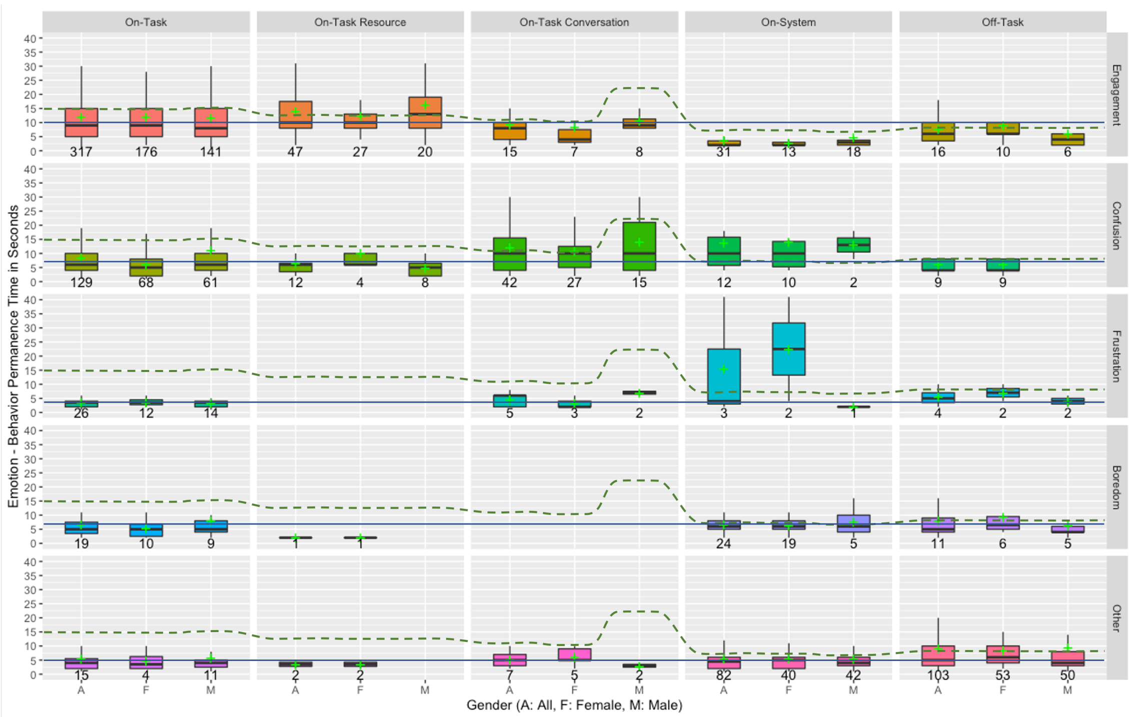

Figure 2. Distribution of the duration time in emotions and behaviors.

According to the box-plots in Figure 1, the student's gender did not significantly affect behavior duration, except for on-task conversation, where male students ($\bar{X} = 24.5s, M = 25s, SD = 16.9s$) spent more time than female students ($\bar{X} = 12.4s, M = 9s, SD = 12.2s$) ($p = .0154$). Students spent more time on-task ($\bar{X} = 19.9s, M =$

$15s, SD = 22.5s$) compared to other behaviors, except on-task-resource ($p < .05$). On-task-resource ($\bar{X} = 15.6s, M = 12s, SD = 13.3s$) had a similar duration to on-task and on-task conversation but lasted longer than on-system and off-task ($p < .0001$). On-task conversation ($\bar{X} = 15.9s, M = 10s, SD = 14.6s$) was shorter than on-task ($p = .0478$) but longer than on-system and off-task ($p < .001$). On-system duration ($\bar{X} = 7.4s, M = 5s, SD = 9.3s$) was shorter than all other behaviors ($p < .01$). Finally, off-task ($\bar{X} = 9.5s, M = 6s, SD = 11.9s$) was shorter than all behaviors except on-system ($p < .001$).

We computed the duration of emotions and behaviors for students, considering their gender, as shown in Figure 2. A straight line indicates the median emotion duration for all students, while a dotted line represents the median behavior duration. The grid in Figure 2 displays the duration distribution for combinations of student emotions and behaviors. Rows show behaviors for each emotion, and columns show emotions for each behavior. The vertical axis indicates duration in seconds, and the horizontal axis shows the time distribution by gender. The number below each lane indicates the sample count. Unlike the analysis of dynamics between behaviors and emotions, this graph includes occurrences less than 11 because it only shows the duration of behaviors and emotions, without statistical analysis.

Based on Figure 2, emotion duration varied by behavior. Students generally stayed engaged for a median of 10 seconds ($\bar{X} = 16.3s, SD = 17.2s$). Engagement duration was shorter on-task ($\bar{X} = 11.2s, M = 9s, SD = 11.2s$) compared to on-task-resource ($\bar{X} = 14.1s, M = 10s, SD = 12.8s$) ($p = .0289$), but longer than on-system ($\bar{X} = 3.9s, M = 2s, SD = 4.6s$) ($p < .0001$) and off-task ($\bar{X} = 7.9s, M = 6s, SD = 6.2s$) ($p = .0597$). During engagement, on-system behavior duration was shorter than on-task-resource ($p < .0001$), on-task-conversation ($p < .001$), and off-task ($p = .0013$). On-task-resource lasted longer than on-task-conversation ($p = .0328$) and off-task ($p = .0098$). No significant gender differences were found. Engagement without behaviors lasted longer than engagement combined with on-task ($p = .0073$), on-task-conversation ($p = .0638$), on-system ($p < .0001$), and off-task ($p = .0136$).

In general, students remained confused for a median of 7 seconds ($\bar{X} = 10.9s, SD = 13.2s$), with longer confusion during on-task-conversation ($\bar{X} = 12.1s, M = 10s, SD = 10.4s$) and on-system ($\bar{X} = 13.9s, M = 10s, SD = 11.9s$) compared to other behaviors ($p < .02$). Confusion during on-task ($\bar{X} = 8.5s, M = 6s, SD = 12.6s$), on-task-resource ($\bar{X} = 6.7s, M = 6s, SD = 5.4s$), and off-task ($\bar{X} = 6.2s, M = 4s, SD = 4.9s$) behaviors showed similar durations. Duration was similar between on-task-conversation and on-system while confused. Female students were less confused than males during on-task ($p = .0229$). Confusion without behaviors lasted longer than when combined with on-task ($p = .0094$), on-task-resource ($p = .0977$), and off-task ($p = .0821$), was shorter than on-system ($p = .0714$), with no difference with on-task-conversation.

Frustration had a short duration with a 4-second median ($\bar{X} = 5.3s, SD = 7.2s$), similar when measured during on-task behavior ($\bar{X} = 3.5s, M = 3.5s, SD = 1.4s$). Other behaviors showed varying durations due to limited samples of frustration combined with on-system, on-task-resource, on-task-conversation, and off-task. No samples were found for on-task-resources. Frustration without behaviors lasted longer than during on-task-resource ($p = .0094$). Other behaviors had similar durations. No significant gender differences were found in the emotion-behavior analysis.

Boredom had a short duration with a 6-second median ($\bar{X} = 7.9s, SD = 5.5s$). Bored students spent similar time on-task ($\bar{X} = 6.6s, M = 5s, SD = 6.3s$), on-system ($\bar{X} = 6.7s, M = 6s, SD = 3.6s$), and off-task ($\bar{X} = 8s, M = 5s, SD = 6.6s$). Boredom without behaviors lasted longer than during on-task ($p = .0807$, marginally significant). We lacked enough samples for on-task-resources and on-task-conversations. No significant gender differences were found in the emotion-behavior analysis.

Students remained in the other state for a 5-second median ($\bar{X} = 8.5s, SD = 12.4s$). In combination with behaviors, the duration in the other state was similar for on-task ($\bar{X} = 5.5s, M = 4s, SD = 6.2s$), on-system ($\bar{X} = 5.6s, M = 4.5s, SD = 5.1s$), and on-task-conversation ($\bar{X} = 5.3s, M = 5s, SD = 3.1s$). Students in the other state spent less time on-task than off-task ($\bar{X} = 9s, M = 5s, SD = 14s, p = .0504$) and less time on-system than off-task ($p = .0278$). There were insufficient samples for on-task-resource behavior. The other state without behaviors lasted longer than during on-task ($p = .0387$) and shorter than on-system ($p = .0136$). No significant gender differences were found in the emotion-behavior analysis.

Figure 2 shows that students were engaged and working on their tasks, either on-task or on-task-resource. When confused, they sought help from classmates or teachers or did something else in the system. Students in on-task-resource did not experience frustration, which was most common during on-task activities. Boredom was not experienced during on-task-resource and on-task-conversation but did occur during on-system, on-task, and off-task behaviors.

5. Discussion

This section answers the research questions based on the results reported in Section 4, exploring the dynamics between emotional states and learning behaviors in an ILE for Brazilian students.

RQ1: Association Between Behaviors and Emotional States. Engagement, a key emotional state, predominantly leads to on-task behavior, showing deep immersion in activities. Engagement also influences on-system, on-task resource, and off-task behaviors, but negatively relates to on-task conversation, suggesting engaged students may not seek help. Conversely, on-task behavior frequently results in engagement, creating a positive feedback loop. On-system and on-task resource behaviors also lead to engagement, highlighting the system's effectiveness. Confusion drives task focus, likely to resolve uncertainties, and negatively relates to on-system behaviors. Frustration leads to on-task behavior, showing a determined effort to overcome difficulties. Boredom increases on-task behavior, likely due to the gamified elements of the ILE, providing distraction and enjoyment. 'Other' emotions also precede on-task behavior, indicating a general trend toward task engagement. Co-occurrence analysis shows that on-task behavior often accompanies engagement and confusion, highlighting the dual role of task involvement in maintaining focus and presenting challenges. On-system behavior is primarily associated with engagement, suggesting these activities sustain interest and prevent confusion. Off-task behavior aligns with various emotions, indicating diverse states during disengagement. On-task conversation is uniquely associated with confusion, underscoring the role of social interaction in seeking help and resolving uncertainties.

RQ2: Gender Differences in the Interaction Between Emotional States and Learning Behaviors. Gender differences were also found in the interaction between emotional states and learning behaviors. Male students exhibit a strong relationship between on-system behavior and engagement, indicating that exploring gamified elements like rankings and scores is highly motivating. This preference for competitive aspects enhances their engagement and likely sustains their interest. In contrast, female students display a unique relationship between off-task behavior and engagement, suggesting they may use off-task behaviors as a strategy to manage frustration and re-engage with learning tasks. For both genders, on-task behavior leads to engagement, but it is negatively associated with frustration primarily in males, suggesting they are more likely to become frustrated when their on-task efforts are unsuccessful. On-system behavior is negatively associated with confusion and boredom for both genders, but males tend to stay more engaged while exploring the system, possibly because they find the gamified features more compelling. Co-occurrence patterns reveal that on-system behavior often co-occurs with engagement for males, highlighting the role of gamified elements in sustaining their interest. For females, on-system behavior frequently aligns with boredom, suggesting a different interaction with the system's features. Additionally, off-task behavior is associated with various emotional states, indicating that females might experience a broader range of emotions when disengaged from tasks. Finally, on-task conversation is uniquely associated with confusion, emphasizing the importance of social interaction in clarifying doubts and resolving uncertainties, a dynamic observed across both genders.

RQ3: Duration of Emotions and Learning Behaviors. The duration of emotions and learning behaviors reveals key patterns. Engagement, the most sustained emotion, lasts longest during on-task and on-task-resource behaviors, indicating deep emotional involvement when students are focused on tasks or using external resources. Conversely, engagement is shortest during on-system behaviors, showing less emotional involvement in non-task-related activities. Confusion lasts longer during on-task conversations and on-system behaviors as students seek help or explore the system to resolve doubts. Female students spend less time confused during on-task activities compared to males. Frustration is generally brief, especially during on-task behavior, suggesting that focused work helps alleviate frustration quickly. Boredom is also short-lived but occurs during on-task, on-system, and off-task behaviors, indicating these activities may not be sufficiently engaging. Notably, boredom is less common during on-task-resource and on-task-conversation behaviors, highlighting these activities as more engaging for students.

RQ4: Comparison with International Findings. Comparing our findings with international studies reveals both similarities and divergences. The affect dynamics model by [D'Mello and Graesser 2012] emphasizes cognitive disequilibrium and confusion in learning, suggesting that properly managed confusion can lead to re-engagement. Our study supports this as confusion often leads to on-task behavior, but negatively correlates with on-system behavior, indicating task-focused efforts to resolve uncertainties. The re-analysis of [Karumbaiah et al. 2021] using the L metric shows cultural context affects affect dynamics, aligning with our unique patterns in Brazilian students. We further highlight the role of sustained engagement during on-task and on-task-resource behaviors, fostering deeper emotional involvement. Previous studies on affect and behaviors by [Rodrigo et al. 2009, Pardos et al. 2014, Fancsali 2014, Kostyuk et al. 2018] often overlook affect dynamics. Consistent with [Baker et al. 2010, Baker et al. 2011, Baker et al. 2007], we emphasize boredom's impact on engagement, but also note that boredom can lead to on-task behavior due to ILE gamification elements providing distraction and enjoyment. The study of [Morais and Jaques 2023] on Brazilian students highlights gender and emotion duration's importance. Our research confirms and expands these findings, showing engagement and confusion as the most sustained emotions, deeply affecting learning behaviors. Female students spend less time confused during on-task activities, indicating gender-specific emotional responses.

6. Conclusion

This study highlighted the nuanced relationships between students' emotions, behaviors, and gender in ILE learning. Our statistical analysis emphasizes the significant impact of gender and learning behaviors on emotions. Unlike previous research, our findings suggest that ILE characteristics, such as gamification, play a crucial role in managing emotions like boredom, underscoring the importance of ILE design in shaping learning experiences. Our results highlight the potential for ILEs to leverage affect dynamics and learning behaviors to create more adaptive and engaging environments [D'Mello 2020]. Identifying that specific student actions can predict emotional shifts underscores the importance of integrating these insights into ILE design. For instance, the observation that students benefit from brief breaks after periods of sustained engagement suggests a strategy where ILEs could proactively provide cognitive breaks or engaging activities to preemptively address or mitigate negative emotional states. Similarly, addressing boredom during on-system activities, particularly among female students, suggests targeted interventions to re-engage students before disengagement intensifies.

These insights not only contribute to the development of more nuanced affect detectors but also offer a foundation for creating ILEs that are responsive to the diverse and dynamic needs of learners. By integrating behavioral predictions with emotional state detection, we can enhance the accuracy and relevance of affective interventions, promoting an engaging and effective learning experience.

Acknowledging the study's limitations is crucial. Its focus on a specific student demographic and a single ILE highlights the need for broader research. The EmAP-ML protocol [Morais et al. 2019] used for collecting student emotions and behaviors showed good inter-rater agreement (Cohen's Kappa $> .6$) but later relied on single-coder annotations, which may introduce bias. Although the protocol's short annotation clips (5 seconds) capture detailed emotions and behaviors, this method is time-consuming and challenges large data set collection.Our use of the L statistic revealed significant directional relationships between states, indicating transitions that merit further study. However, this evidence does not imply causality. Future research should conduct experiments to rigorously investigate the causal nature of these transitions.

As future work, there is an evident opportunity to apply these insights in the development of adaptive learning technologies. Incorporating automated affect detection, accounting for learner diversity, and undertaking longitudinal and comparative studies are pivotal steps toward crafting educational experiences that are not only more personalized and engaging but also inclusive and effective for every student.

Acknowledgment

This study was financed in part by the following research funding agencies: CAPES (Finance Code 001), FAPERGS (granting $17/2551 - 0001203 - 8$) and CNPq (granting $306005/2020 - 4$) from Brazil. We would like to acknowledge that this article was proofread using ChatGPT 4.0 with prompt engineering designed to enhance clarity and flow of the authors' initial manuscript. Following this proofreading, the article underwent a thorough review by the authors, who take full responsibility for its content.

References

Baker, R. (2011). Gaming the System: A Retrospective Look. *Philippine Computing Journal*, 6(2):9–13.

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., and Graesser, A. C. (2010). Better to be frustrated than bored: ... *Int. J. of HCS*, 68(4):223–241.

Baker, R. S., Moore, G. R., Wagner, A. Z., Kalka, J., Salvi, A., Karabinos, M., Ashe, C. A., and Yaron, D. (2011). The dynamics between student affect and behavior occurring outside of educational software. In *ACII*, pages 14–24. Springer.

Baker, R. S., Rodrigo, M. M. T., and Xolocotzin, U. E. (2007). The dynamics of affective transitions in simulation problem-solving environments. In *ACII*, pages 666–677.

D'Mello, S., Graesser, A., and Taylor, R. S. (2007). Monitoring affective trajectories during complex learning. In *Annu. Meet. Cogn. Sci. Soc.*, volume 29.

D'Mello, S. K. (2020). Big data in the science of learning. In *Big data in psychological research*, pages 203–225. APA.

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

Fancsali, S. (2014). Causal discovery with models: behavior, affect, and learning in cognitive tutor algebra. In *Educational Data Mining 2014*. EDMS.

Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology*, 2(3):300–319.

Frenzel, A. C., Pekrun, R., and Goetz, T. (2007). Girls and mathematics—a "hopeless" issue? a control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, 22(4):497–514.

Graesser, A. and D'Mello, S. K. (2011). Theoretical perspectives on affect and deep learning. In *New Perspectives on Affect and Learning Tec.*, pages 11–21. Springer, NY.

Graesser, A. C., D'Mello, S. K., and Strain, A. C. (2014). Emotions in advanced learning technologies. In *International handbook of emotions in education*, pages 483–503. Routledge, New York.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58(1):47–77.

Jaques, P. A., Seffrin, H., Rubi, G., Morais, F., Ghilardi, C., Bittencourt, I. I., and Isotani, S. (2013). Rule-based expert systems to support step-by-step guidance in algebraic problem solving: The case of the tutor PAT2math. *Expert Systems with Applications*, 40(14):5456–5465.

Karumbaiah, S., Baker, R. B., Ocumpaugh, J., and Andres, A. (2021). A re-analysis and synthesis of data on affect dynamics in learning. *IEEE TAC*.

Kostyuk, V., Almeda, M. V., and Baker, R. S. (2018). Correlating affect and behavior in reasoning mind with state test achievement. In *LAK*, pages 26–30.

Matayoshi, J., Karumbaiah, S., et al. (2020). Adjusting the l statistic when self-transitions are excluded in affect dynamics. *J. of EDM*, 12(4):1–23.

Morais, F. and Jaques, P. A. (2023). The dynamics of brazilian students' emotions in digital learning systems. *IJAIED*.

Morais, F., Kautzmann, T. R., Bittencourt, I. I., and Jaques, P. A. (2019). EmAP-ML: A protocol of Emotions and behaviors Annotation for Machine Learning labels. In *EC-TEL*, pages 495–509, Netherlands. Springer.

Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., and Gowda, S. M. (2014). Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes. *JLA*, 1(1):107–128.

Pekrun, R. (2014). Emotions and learning. Technical Report Educational Practices Series-24, UNESCO International Bureau of Education.

Perkins, H. V. (1965). Classroom behavior and underachievement. *American Educational Research Journal*, 2(1):1–12.

Rodrigo, M. M. T., Baker, R. S., Jadud, M. C., Amarra, A. C. M., Dy, T., Espejo-Lahoz, M. B. V., Lim, S. A. L., Pascua, S. A., Sugay, J. O., and Tabanao, E. S. (2009). Affective and behavioral predictors of novice programmer achievement. In *ACM SIGCSE Conf. on Innovation and technology in CSE*, pages 156–160.

Sabourin, J., Rowe, J. P., Mott, B. W., and Lester, J. C. (2011). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In *AIED*, pages 534–536. Springer.

Zeidner, M. (1998). *Test Anxiety: The State of the Art*. Springer, New York.