Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Do Categorical Emotions Have Intensity? Modeling Emotional Intensity Transitions by Task, Peer, and Environment Causes

Ryan LI[a], Ashwin T S[a], Divya MEREDDY[a], & Gautam BISWAS[a]
[a]*Computer Science, Vanderbilt University, USA*
*ashwindixit9@gmail.com

Abstract: Emotions shape students' learning experiences. While research links positive emotions to better outcomes and negative emotions to poorer ones, few studies examine how emotions change in intensity and cause during collaborative learning. This paper introduces and operationalizes a tripartite cause framework—Task, Peer, and Environment—and connects graded learning-centered emotions to their triggers. Using EcoJourneys, a narrative-centered, game-based learning environment, we annotated the emotional states of 21 elementary school students, capturing shifts among five core emotions (Engagement, Boredom, Confusion, Frustration, Delight) across four intensity levels. We investigated four research questions: which causes trigger transitions; how causes affect state duration; whether cause effects vary by intensity; and how theory-grounded causes (CVT, SDT, EST) shape intensity transitions. Our findings show that most transitions increase Engagement intensity. Peer causes frequently support recovery from moderate Confusion to higher Engagement, Task causes often drive within-engagement transitions, and environmental influences show more sustained effects. This work offers a theory-grounded account of emotional intensity dynamics in collaborative learning and informs the design of emotion-aware learning technologies.

Keywords: Learning-Centered Emotions, Collaborative Learning, Emotional Intensity, Emotional Transitions, Emotional Analysis, EcoJourneys, External Causes

1. Introduction

Emotions are integral to understanding students' learning, particularly in Self-Regulated Learning (SRL) (Zimmerman, 1986) and Socially Shared Regulation of Learning (SSRL) (Järvelä *et al.*, 2015). They complement cognitive states and provide insight into engagement, group dynamics, and collaborative learning regulation. In classrooms, learning-centered emotions—Engagement, Boredom, Confusion, Frustration, and Delight—are prevalent compared to universal emotions, forming the basis for affective research in education (D'Mello & Graesser, 2012; TS and Biswas, 2024).

Previous studies treated emotions as discrete states (TS and Biswas, 2024; Ashwin *et al.*, 2024). However, observations indicate emotions like Confusion and Engagement exhibit varying intensities (D'Mello & Graesser, 2014). Mild Confusion may signal curiosity, while intense Confusion implies a higher cognitive load that can hinder learning (D'Mello & Graesser, 2014). These differences necessitate modeling emotions by type and intensity, prompting exploration of emotional dynamics through graded intensities and transitions (Reisenzein & Junge, 2024; Williams *et al.*, 2023). However, modeling intensity alone does not make interventions actionable. Prior approaches emphasize internal appraisals of emotion (e.g., Control–Value Theory, CVT; Pekrun, 2006) and, in empirical practice, often do not link intensity changes to external drivers such as task design, peer interaction, or environmental conditions. To enable targeted, designable interventions, we examine external causes of intensity transitions—Task, Peer, and Environment.

In this paper, we introduce and operationalize a tripartite cause framework—Task, Peer, and Environment—for analyzing emotional intensity transitions in collaborative learning. We ground these causes in established theory: Control–Value Theory for task-related influences (Pekrun, 2006), Self-Determination Theory (SDT) for peer-related influences (Deci & Ryan, 2012), and Ecological Systems Theory (EST) for environmental influences (Bronfenbrenner, 1979). We select CVT, SDT, and EST as a minimal, complementary set: CVT captures task appraisals (control, value) underlying shifts among engagement, confusion, and frustration; SDT explains how autonomy, competence, and relatedness shape peer-driven affect; and EST models contextual conditions (e.g., noise, interruptions, fatigue) that modulate attention and regulation. Other theories (e.g., ACT, CLT, SCT, EVT) emphasize different phenomena or omit one or more of the task/peer/environment pathways, making them less suitable as a compact explanatory set for our focus on actionable external causes.

Using multimodal data from 21 students in a narrative-centered, collaborative learning environment (EcoJourneys), we analyze how these causes shape transition types, durations, and intensity patterns. Recognizing emotion-intensity transitions is not only theoretically significant but also critical for designing responsive educational technologies. Adaptive learning systems must respond differently to mild disengagement versus deep frustration, and detecting the intensity of an emotion can inform tailored interventions. Modeling transitions across intensity levels enables real-time emotional awareness that supports personalized scaffolding, making these systems more responsive to the evolving affective states of learners. Accordingly, we test how these causes shape transition types, durations, and intensity patterns in collaborative learning.

Specifically, this study addresses the following research questions:

RQ1: What causes among Task, Peer, and Environment trigger emotional transitions?

RQ2: How do these causes affect the duration of emotional states?

RQ3: Do different causes trigger transitions at different intensities of the same emotion?

RQ4: How do Task, Peer, and Environment causes, as conceptualized through CVT, SDT, and EST, shape emotional intensity transitions in collaborative learning?

By addressing these questions, we aim to enhance understanding of collaborative learning dynamics and inform the design of affect-sensitive technologies.

2. Related Work

Emotion Intensity in Learning Theories: Theories have been proposed that prompt the investigation of different intensities that can exist within the same emotion. Control-Value Theory (CVT) explains how internal appraisals influence emotional experiences during learning (Pekrun, 2006). Extensions of CVT include physiological signals to estimate emotion intensity (Barradas et al., 2025), acknowledging graded emotions. However, CVT focuses on internal appraisals and provides limited guidance on external influences, such as task structure or peer interaction. The Circumplex Model (Russell, 1980; Akpanoko et al. 2024; Akpanoko and Biswas, 2024) supports a continuous representation of emotions in valence-arousal space. Yet, empirical applications often fail to distinguish functional variations within emotions or connect shifts in emotion to external causes, such as the impact of emotions on the overall learning process. This leaves gaps in understanding how intensity evolves in collaborative settings. CVT and the Circumplex model drive the investigation of within-emotion intensity and its development. However, empirical research still underexplores the links between changes in intensity and external factors such as tasks, peers, or the environment.

Modeling Emotional Transitions: Markov models are instrumental in analyzing emotional transitions, predicting the likelihood of one emotion following another. Cipresso *et al.* (2023) used a Markov chain model to compute transitional probabilities between emotions like Stressed, Engaged, Bored, and Relaxed, demonstrating the utility of probabilistic models in learning contexts. Prasetio *et al.* (2020) further integrated Time-Delay Neural Networks with Markov chains to predict stress and emotion transitions, highlighting temporal aspects in affect modeling. In the context of learning, emotional state modeling has employed Hidden Markov Models using multimodal data, such as emotion detection from facial expressions (Schmidt *et al.*, 2010; Bosch *et al.*,2016), facial thermal imaging (Liu & Wang, 2011), and sentiment analysis (Ho & Cao, 2012). These studies often use discrete categorical frameworks and overlook transitions within the same emotion at varying intensities, such as escalating from mild to intense Confusion.

Causal Attribution of Emotional Transitions. Prior work in learning sciences and educational psychology points to three broad classes of external influences on students' affect during collaborative learning: task-related, peer-related, and environment-related factors. Control-Value Theory explains how task appraisals (control and value) shape engagement, confusion, and frustration (Pekrun, 2006). Self-Determination Theory highlights how peer interactions and the fulfillment of competence and relatedness needs regulate motivation and affect (Deci & Ryan, 2012; Järvelä *et al.*, 2015). Ecological perspectives emphasize how environmental conditions (e.g., noise, visual clutter, interruptions) modulate attention and emotional regulation (Bronfenbrenner, 1979; Godwin & Fisher, 2011; Fisher *et al.*, 2014; Gheller *et al.*, 2023). Empirical studies show that well-calibrated tasks sustain engagement, whereas poorly calibrated tasks elicit confusion or frustration (Özhan & Kocadere, 2020; Gijlers & de Jong, 2013). Peer scaffolding can facilitate recovery from negative states (Fu *et al.*, 2009; Järvelä *et al.*, 2015), and environmental distractions are linked to disengagement and boredom (Shernoff, 2013; Godwin & Fisher, 2011; Fisher *et al.*, 2014; Gheller *et al.*, 2023). While CVT and SDT support task- and peer-related influences, and ecological perspectives highlight environmental factors, the literature typically examines these influences in isolation and models emotions as discrete categories. This paper introduces and operationalizes a unified tripartite cause framework—Task, Peer, and Environment—and links these causes to fine-grained emotional transitions across intensity levels.

3. Methodology

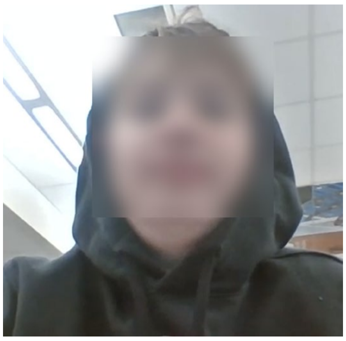
Learning Environment and Data: EcoJourneys is a game-based collaborative learning environment where students investigate the illness in tilapia on an island in the Philippines (Acosta *et al.*, 2024). It includes a tutorial and three inquiry quests, with activities like Deduce and TIDE (Talk, Investigate, Deduce, Explain) that promote evidence-based reasoning. In this IRB-approved study, 21 elementary students participated over five days, engaging in about an hour of gameplay daily. All students and guardians provided informed consent before data collection, which included multimodal recordings of video, speech, and logs to capture emotional and behavioral indicators during the collaborative process. To enable analysis aligned with our research questions, we annotate both emotional categories and their intensity levels to model graded transitions rather than only discrete states. This supports examining transition types, durations, and intensity patterns in subsequent analyses (RQ1–RQ3), and provides the basis for theory-grounded interpretation (RQ4).

3.1 *Emotion Annotation Scheme and Intensity Definition*

We annotated students' emotional states using synchronized video, audio, and screen recordings on the *EcoJourneys* platform (Figure 1). Annotations focused on individual task engagement and group interactions as students worked in teams of 3–4 on inquiry-driven activities. We highlighted five learning-centered affective states—Engagement, Confusion, Frustration, Boredom, and Delight—defined by D'Mello & Graesser (2012).

Emotions in learning environments vary in intensity, with observable behaviors and verbal cues indicating different affective levels (Gupta *et al.*, 2016; Xu *et al.*, 2023). Facial Action Coding System (FACS; Ekman & Friesen, 1978), along with other modalities such as hand gestures (Vicario & Newman, 2013; Ashwin and Guddeti, 2020a), body postures (Dael *et al.*, 2012; Ashwin and Guddeti, 2020b), are examples of visual cues that we can directly observe from the learners. Transcribed text, pitch, and tone are examples of verbal cues that we can hear (Fonteles *et al.*, 2024). These multimodal signals help distinguish emotional intensity functionally and perceptually, in which similar observations have been noted in earlier studies on affective expression in learning (D'Mello & Graesser, 2014; Bosch *et al.*, 2016).

Guided by these insights, we observed and classified emotional expressions into four intensity levels: *low, moderate low, moderate high, and high*, based on facial, body, and verbal indicators. These categories emerged through iterative observation and were later formalized. This clustering revealed stable patterns among learners, supported by theory-driven distinctions in intensity thresholds, especially for Confusion (D'Mello & Graesser, 2014). Only three intensity levels for Boredom were noted in our dataset. Two trained annotators independently coded emotion categories, intensity levels, and causes, achieving high inter-rater reliability (Cohen's κ = 0.91), confirming consistency in our intensity judgments.



(a) (b)

Figure 1(a) (on the left) A scene where the avatar is controlled by the students. Figure *1(b)* (on the right) The student is using the EcoJourneys software to complete the task. Note that the student's face is not shown for IRB compliance.

Table 1. *Emotion Intensity Scale (Used Across All Emotions)*

Level	Description of Each Emotion Intensity
1	Low – Minimal expression, short duration, subtle or isolated cue
2	Moderate low – Clearly expressed but not dominant, 1–2 cues observed
3	Moderate high – Noticeable intensity, 2+ cues, co-occurring behaviors
4	High – Dominant emotion with rich multimodal expression and sustained cues

Two trained annotators independently coded the emotional states and causes using a standardized rubric. Initial training involved calibration rounds with practice videos, where discrepancies were discussed and resolved until inter-rater reliability exceeded 0.9. During annotation, each coder worked independently and flagged ambiguous cases for joint review with a third expert. Regular consistency checks and review meetings maintained label accuracy throughout the dataset. This structured workflow ensured reliable, theory-aligned labeling across students and sessions. Intensity levels were determined through manual, multimodal observation of student behaviors, based on the idea that higher-intensity emotions are expressed through multiple cues, such as facial expressions, body posture, and vocal signals. Lower-intensity states generally have a single, brief cue. Observed patterns were grouped into four distinct levels for each emotion. Each entry included a timestamp with start time, end time, emotion, intensity level, and cause (e.g., Confusion 2 → Engagement 3, Cause: Peer). This approach allowed systematic analysis of emotion changes in relation to their causal context and intensity fluctuations. Table 1 summarizes the intensity scale used across all emotions.

Cause Annotation Scheme: To capture the context of each emotional transition, annotators assigned a cause label: Task, Peer, Environment, or a combined label (Peer/Task, Task/Environment, Peer/Environment) when influences co-occurred. Task was coded when difficulties in progressing (e.g., ambiguity, navigation/usability issues) precipitated the shift; Peer when interactions with group members altered affect (e.g., help that resolved confusion, social comparison, miscommunication); and Environment when external distractions (e.g., classroom noise, instructor prompts, self-initiated off-task behavior) were primary drivers. We permitted overlapping cause annotations to reflect co-occurring influences in collaborative settings. For interpretation in RQ4, we align Task with CVT, Peer with SDT, and Environment with EST; the cause framework itself is introduced in this paper. Operational cues for each emotion at each intensity level are listed in Table 2.

Table 2. *Emotion-Specific Intensity Descriptors*

Emotion	Level 1	Level 2	Level 3	Level 4
Engagement	Gaze unfocused, glancing around	Semi-focused, occasional distractions	Focused, limited peer interaction while working	Fully focused, immersed, no external attention shifts
Confusion	Brief hesitation, single cue	Repeated questions, procedural uncertainty	Multiple failed attempts, verbal + nonverbal cues	Facial strain, physical gestures, vocal escalation
Frustration	Mild complaint or sigh	Fast typing, annoyed tone	Physical agitation, task-related frustration	Raised voice, visible anger, peer-directed frustration
Boredom	Stretching, looking away briefly	Longer disengagement, chatting	Extended inattention, no task input	Complete withdrawal, slouched posture, no interaction with task or group *(not observed)*
Delight	Subtle smile or posture shift	Light laughter or excitement	Verbal joy, leaning forward	Expressive body language, vocal exclamations

Table 3. *Cause Category Definitions*

Cause	Definition of the Causes
Task	Triggered by interactions with the learning software or task content
Peer	Triggered by verbal/nonverbal interactions with group members
Environment	Triggered by factors outside the task or group, e.g., classroom noise, instructor intervention
Peer/Task	Interaction with peers about task progress or content
Task/Environment	Task engagement is influenced by the physical or instructional environment
Peer/Environment	Peer interaction involving environmental topics or students outside the group

Annotation Structure and Transition Format: In this experiment, we document types of transitions, including emotion categories, their intensity levels, causes, frequencies, and durations. An annotation example is (0:30, 1:20, Engagement, 4, Peers), indicating the student exhibited a high level of Engagement from 30 seconds to 1 minute and 20 seconds, influenced by peers. To annotate a transition, we consider the next instance. For example, (0:30, 1:20, Engagement, 4, Peers) and (1:20, 1:35, Confusion, 3, Task) can be noted as Engagement 4

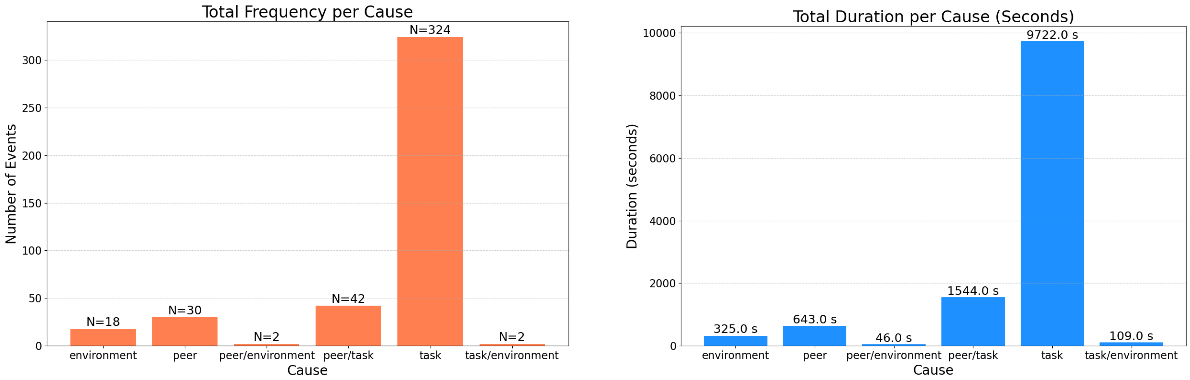
→ Confusion 3 due to Task (as the transition to Confusion 3 stems from the task itself). Definitions and examples of the cause categories appear in Table 3.

4. Results

We analyzed 418 emotional transitions, totaling 12,389 seconds of activity from 21 students. Off-task periods (1,975 seconds) and undetectable facial recordings (391 seconds) were excluded. Below are findings regarding our four research questions.

RQ1: *What causes among Task, Peer, and Environment trigger emotional transitions*?

We first quantify which causes are associated with observed emotional transitions (Figure 2). Transitions were primarily triggered by Task-related causes (324 transitions, 77.51%), followed by Peer (30), Environment (18), Peer/Task (42), and minimal contributions from Peer/Environment (2) and Task/Environment (2). Task causes accounted for the most unique transitions (81 types), compared to Peer (13) and Peer/Task (23). Most transitions led to increased Engagement. For instance, under Task, the top transitions were Engagement 2 → Engagement 3 (10%), Engagement 3 → Engagement 2 (9%), and Engagement 3 → Engagement 4 (7%). Among Peer-triggered transitions, the most common was Confusion 2 → Engagement 3 (17%), indicating peer collaboration is vital for emotional recovery. A classroom study reported a similar cause distribution (Task ≫ Peer > Environment), reinforcing that the prevalence of Task-triggered transitions is a robust pattern in collaborative learning settings.



(a) *(b)*

Figure 2. Distribution of (a) total frequency and (b) total duration counts of causes across all transition types.

RQ2: *How do these causes affect the duration of emotional states*?

We next examine how cause categories relate to the duration of emotional states. In terms of total duration, Task-based transitions dominated (9,722 seconds), followed by Peer/Task (1,544s), Peer (643s), Environment (325s), Task/Environment (109s), and Peer/Environment (46s). Notably, Environment-related causes had a long average duration per instance (mean = 59.87s), suggesting that environmental distractions may have more sustained effects on emotional states.

Transitions from Engagement 2 → Engagement 3 accounted for the most extended cumulative duration (2,798 seconds), indicating that students spent significant time increasing attention during collaborative gameplay. Task-triggered transitions involving non-Engagement states into Engagement also had high durations, with examples such as Boredom 1 → Engagement 4 (131s, single instance) and Boredom 1 → Engagement 3 (91s, single instance). A one-way ANOVA showed marginally significant differences in transition duration across cause types: $F(5, 412) = 2.42$, $p = 0.065$, suggesting that the distribution of durations varies depending on the causal trigger.

RQ3: *Do different causes trigger transitions at different intensities of the same emotion*?

We analyzed transitions across intensity levels using a Markov Chain framework. Task-related causes primarily led to moderate-to-high levels of Engagement, including Engagement 2 → Engagement 3, Engagement 3 → Engagement 4, and Engagement 3 → Engagement 2. Peer causes often helped students recover from Confusion into higher Engagement, most notably Confusion 2 → Engagement 3. High-intensity Frustration 4 and Confusion 4 were only observed under Task, Peer, and Peer/Task conditions, indicating that negative high-intensity states emerge more frequently from direct interaction with content or social dynamics. In contrast, Delight 4 only appeared in transitions related to Task and Environment.

A Chi-square test examining the association between cause and emotion category revealed that Engagement transitions were most diversely caused: $\chi^2(10, N = 418) = 11.65$, $p = 0.0399$, *Cramér's V* = 0.1851, indicating a medium effect size. This suggests that Engagement intensity transitions are particularly sensitive to multiple contextual factors.

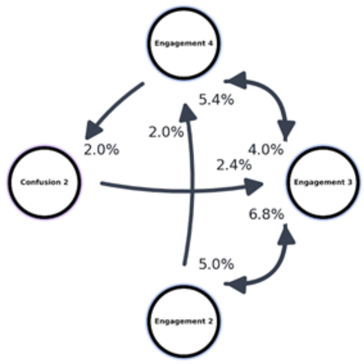


Figure 3. Frequent intensity transitions (≥2% of all transitions).

In Figure 3, nodes denote emotion + intensity level; directed edges denote observed transitions. Labels show % of all transitions (N=418) and edge thickness is proportional to frequency. Transitions are aggregated across causes, complementing Figure 2 (cause distribution) and Figure 3 (durations). Most frequent transitions are within the Engagement spectrum (e.g., Engagement 2→3, 3→4) and recovery from Confusion 2→Engagement 3, complementing cause and duration summaries in Figures 2–3.

RQ4: *How do Task, Peer, and Environment causes (CVT, SDT, EST) shape emotional intensity transitions*?

Observed patterns align with the theoretical lenses used for interpretation. Task-related causes dominated transitions (324/418; 77.51%) and most often produced moderate-to-high Engagement intensity, including Engagement 2→3 (10%), Engagement 3→2 (9%), and Engagement 3→4 (7%). Peer-related causes frequently supported recovery from negative states, notably Confusion 2→Engagement 3 (17% of peer-triggered transitions). Environment-related transitions were fewer (18 total) yet showed a longer average duration per instance (≈59.87s), indicating more sustained contextual effects. Interpreted through CVT/SDT/EST, these results provide a concise, theory-grounded account of how external causes shape emotional intensity transitions in collaborative learning.

Patterns observed here are consistent with prior evidence from a narrative-centered classroom study of middle-school learners. That study reported that Task-related causes were the dominant triggers of emotional change (71.3%), with Peer at 23.2% and Environment at 7.1%; within categories, Task most often co-occurred with Engagement (60%), Peer with Engagement/Delight (50%/30%), and Environment with Boredom (60%). These distributions and within-category associations mirror our findings that Task chiefly drives within-Engagement intensity increases, Peer frequently supports recovery (e.g., Confusion 2 → Engagement 3), and Environment relates to disengagement-prone states. Together, the alignment across studies strengthens the interpretation that cause type systematically shapes intensity transitions in collaborative learning.

5. Discussion

This study examined how emotional transitions caused by Task, Peer, and Environment influence the type, intensity, and duration of a student's emotional state. By annotating multimodal data with emotion and cause, we provide a more nuanced understanding of affective dynamics in collaborative learning.

Our findings show that while Task-related causes are the most common, Peer-related causes effectively promote transitions from negative to positive states, especially from moderate Confusion to high Engagement. This supports claims from SRL and SSRL frameworks emphasizing that peer scaffolding and social regulation can stabilize or elevate emotional states (Järvelä *et al.*, 2015). In contrast, Environment-related causes, though less frequent, result in longer transition durations, suggesting that disruptive stimuli like noise or interruptions have lasting effects on emotional regulation. These findings align with ecological systems theory (Bronfenbrenner, 1979), which posits that environmental conditions significantly impact self-regulation from many perspectives. They highlight the need for thoughtfully designed learning environments. Transitions toward higher engagement were common across cause categories, yet only Task-related triggers were linked to high-intensity negative emotions like Frustration 4 and Confusion 4. This indicates that poorly aligned tasks, including those with characteristics such as redundancy or conflicting features, can disengage learners and induce frustration. Therefore, task calibration must balance challenge and support, especially in adaptive systems.

From a modeling perspective, intensity levels underscore the difference between mild and strong emotional expressions. For example, Confusion 1 may signify curiosity, while Confusion 4 can reflect overload. Recognizing these differences allows for targeted interventions—e.g., possibly addressing Task-induced Confusion at Level 2 with hints, while Level 4 may need peer or instructor support. Finally, Engagement emerged as the most diverse in causes, reinforcing that it is not a simple "on-task" signal but a dynamic state shaped by social, task-based, and environmental factors. This insight suggests that affect-aware learning systems should not only detect discrete emotional states but also understand transitions and context-specific causes.

Taken together, these findings show that Task-related transitions are consistent with Control–Value Theory (CVT), shaping movement toward moderate-to-high Engagement as well as occasional high-intensity Confusion/Frustration; Peer-related influences align with Self-Determination Theory (SDT), supporting recovery transitions such as Confusion 2 → Engagement 3; and Environment-related influences reflect Ecological Systems Theory (EST), exhibiting longer average durations indicative of sustained contextual effects. This mapping provides a concise explanation of how causes shape emotional intensity transitions in collaborative learning (answers RQ4).

5.1 Limitations

This study proposed a framework linking emotional intensity levels with external causes in collaborative learning, but it has limitations. The dataset was small (21 students) and manually coded, leading to limited generalizability despite high inter-rater agreement ($\kappa = 0.91$). Some transitions, like Boredom Level 4, were likely absent due to the game structure. Cause attribution relied on visible cues from video and audio, making these interpretive despite prior validation. Future studies should include system logs or gaze data for more explicit cause identification. All findings are correlational, with patterns emerging, such as peer interactions aiding recovery and tasks triggering diverse transitions, but causal mechanisms need validation through real-time or experimental studies. Despite these constraints, this study offers three key insights. First, categorical learning of emotions shows meaningful intensity levels. Second, transitions are influenced by identifiable causes: Task, Peer, and Environment. Third, modeling transitions reveals dynamics often missed in discrete emotion labeling. This work contributes to developing emotion-aware learning technologies that adapt to emotion types and their contextual evolution, supporting more collaborative and emotionally intelligent learning environments.

6. Conclusion

This study introduced a framework for modeling emotional transitions in collaborative learning by examining five categorical learning-centered emotions and their perceived intensity influenced by external causes—Task, Peer, and Environment. Through detailed annotations of 418 emotional transitions, we found that intensity transitions are patterned and context-sensitive. Task-related causes were most frequent and diverse, resulting in high Engagement and intense negative states. Conversely, Peer-related causes helped recovery from negative emotions, particularly moderate Confusion. Environmental causes, though less frequent, were linked to lasting emotional states.

Our findings underscore the need to integrate emotion intensity and causal context in affect-aware learning technologies. Instead of focusing only on emotional categories, intelligent systems should recognize nuanced emotional trajectories and triggers. Understanding how emotions evolve allows educators and designers to better support emotional regulation and engagement in collaborative learning. Future research should involve larger datasets, explore group-level emotion dynamics, and integrate predictive, real-time adaptive mechanisms addressing emotional intensity and causality.

Acknowledgements

This research was supported by the National Science Foundation AI Institute Grant No. DRL-2112635. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Acosta, H., Lee, S., Mott, B., Bae, H., Glazewski, K., Hmelo-Silver, C., & Lester, J. (2024). Multimodal Learning Analytics for Predicting Student Collaboration Satisfaction in Collaborative Game-Based Learning.

Akpanoko, C. E., Cordell, G., & Biswas, G. (2024). Investigating the relations between students' affective states and the coherence in their activities in open-ended learning environments. In Proceedings of the 17th International Conference on Educational Data Mining (pp. 511-517).

Akpanoko, C. E., & Biswas, G. (2024). The interplay of affective states and cognitive processes in an open-ended learning environment: A case study. In Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024, pp. 873-880. International Society of the Learning Sciences.

Ashwin, T. S., & Guddeti, R. M. R. (2020). Affective database for e-learning and classroom environments using Indian students' faces, hand gestures and body postures. Future Generation Computer Systems, 108, 334-348.

Ashwin, T. S., & Guddeti, R. M. R. (2020). Impact of inquiry interventions on students in e-learning and classroom environments using affective computing framework. User Modeling and User-Adapted Interaction, 30(5), 759-801.

Ashwin, T. S., SNYDER, C., AKPANOKO, C. E., Srigowri, M. P., & BISWAS, G. (2024, November). Combining Multimodal Analyses of Students' Emotional and Cognitive States to Understand Their Learning Behaviors. In International Conference on Computers in Education.

Barradas, I., Tschiesner, R., & Peer, A. (2025). Dynamic emotion intensity estimation from physiological signals facilitating interpretation via appraisal theory. *PloS one*, *20*(1), e0315929.

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 (TiiS)*, *6*(2), 1-26.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Harvard University Press.

Cipresso, P., Borghesi, F., & Chirico, A. (2023). Affects affect affects: A Markov chain. *Frontiers in Psychology*, *14*, 1162655.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, 12(5), 1085.

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. *Handbook of theories of social psychology, 1*(20), 416-436.
D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145-157.
D'Mello, S. K., & Graesser, A. C. (2014). Confusion. In *International handbook of emotions in education* (pp. 289-310). *Routledge*.
Ekman, P., & Friesen, W. V. (1978). Facial Action Coding System. *Environmental Psychology & Nonverbal Behavior*.
Fisher, A. V., Godwin, K. E., & Seltman, H. (2014). Visual environment, attention allocation, and learning in young children: When too much of a good thing may be bad. *Psychological science, 25*(7), 1362-1370.
Fonteles, J., Davalos, E., Ashwin, T. S., Zhang, Y., Zhou, M., Ayalon, E., ... & Biswas, G. (2024, July). A first step in using machine learning methods to enhance interaction analysis for embodied learning environments. In International Conference on Artificial Intelligence in Education (pp. 3-16). Cham: Springer Nature Switzerland.
Fu, F. L., Wu, Y. L., & Ho, H. C. (2009). An investigation of cooperative pedagogic design for knowledge creation in web-based learning. *Computers & Education*, *53*(3), 550-562.
Gheller, F., Spicciarelli, G., Scimemi, P., & Arfé, B. (2023). The effects of noise on children's cognitive performance: A systematic review. *Environment and Behavior*, *55*(8-10), 698-734.
Gijlers, H., & de Jong, T. (2013). Using concept maps to facilitate collaborative simulation-based inquiry learning. *Journal of the learning sciences*, *22*(3), 340-374.
Godwin, K., & Fisher, A. (2011). Allocation of attention in classroom environments: Consequences for learning. In *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 33, No. 33).
Gupta, A., Jaiswal, R., Adhikari, S., & Balasubramanian, V. (2016). DAISEE: dataset for affective states in e-learning environments. *arXiv preprint arXiv:1609.01885*, 1-22.
Ho, D. T., & Cao, T. H. (2012, September). A high-order hidden Markov model for emotion detection from textual data. In *Pacific rim knowledge acquisition workshop* (pp. 94-105). Berlin, Heidelberg: Springer Berlin Heidelberg.
Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., ... & Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: Designing for CSCL regulation tools. *Educational Technology Research and Development*, *63*, 125-142.
Liu, Z., & Wang, S. (2011, October). Emotion recognition using hidden Markov models from facial temperature sequence. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 240-247). Berlin, Heidelberg: Springer Berlin Heidelberg.
Özhan, Ş. Ç., & Kocadere, S. A. (2020). The effects of flow, emotional engagement, and motivation on success in a gamified online learning environment. *Journal of Educational Computing Research*, *57*(8), 2006-2031.
Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review,* 18, 315-341.
Prasetio, B. H., Tamura, H., & Tanno, K. (2020). Deep time-delay Markov network for prediction and modeling the stress and emotions state transition. *Scientific Reports*, *10*(1), 18071.
Reisenzein, R., & Junge, M. (2024). Measuring the intensity of emotions. *Frontiers in Psychology*, *15*, 1437843.
Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.
Shernoff, D. J. (2013). Optimal learning environments to promote student engagement.
TS, A., & Biswas, G. (2024, November). Relating students cognitive processes and learner-centered emotions: An advanced deep learning approach. In Proceedings of the 26th International Conference on Multimodal Interaction (pp. 575-584).
Vicario, C. M., & Newman, A. (2013). Emotions affect the recognition of hand gestures. *Frontiers in Human Neuroscience*, 7, 906.
Williams, T. F., Vehabovic, N., & Simms, L. J. (2023). Developing and validating a facial emotion recognition task with graded intensity. Assessment, 30(3), 761-781.
Xu, Y., Li, Y., Chen, Y., Bao, H., & Zheng, Y. (2023). Spontaneous visual database for detecting learning-centered emotions during online learning. *Image and Vision Computing*, 136, 104739.
Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses?. *Contemporary educational psychology*, 11(4), 307-313.