How Anxiety Affects Affect: A Quantitative Ethnographic Investigation using Affect Detectors and Data-Targeted Interviews

J. M. Alexandra L. Andres[1], Stephen Hutt[1], Jaclyn Ocumpaugh[1], Ryan S. Baker[1], Nidhi Nasiar[1] and Chelsea Porter[1]

[1]University of Pennsylvania, Philadelphia PA 19104, USA
aandres@upenn.edu

Abstract. This study integrates the analysis of student interaction data with classroom interviews in order to better understand how students' trait-level anxiety relates to in-the-moment measures of students' affective experiences (i.e., boredom, confusion, delight, engaged concentration, and frustration). We first use quantitative data to drive the data collection, with the interviews being triggered by previously-developed models of student emotions, allowing us to focus interviews during times where specific emotional experiences of interest have just occurred. We then analyze the log data, finding differences in how students with high and low science anxiety manage their emotions (and subsequent behaviors). Finally, we connect these behaviors back to the strategies students articulate in the interviews, finding that many students with higher anxiety scores have an external locus of control, are less likely to seek evidence that they are wrong, and are more likely to make major changes to their solutions in response to their frustration.

1 Introduction

As technology advances, it is becoming possible to scale up the collection of rich, ethnographic data that was typically only available in small-scale studies. In education, qualitative methods such as student interviews can provide critical information about how students learn [1] and in turn, improve learning theory. However, interview methods are resource-intensive and suffer from the "one-shot" problem [2]. Out of context, students may not be able to provide accurate insight on their cognition in a specific key situation hours earlier [3]. Even if interviews are conducted during class, in real-time, in a class of 25 students working quietly, it can be challenging to identify relatively rare events of interest, especially if such events are short-lived. By taking a quantitative approach to directing student interviews we can collect qualitative data from fleeting, yet educationally relevant, experiences.

One such fleeting experience known to impact student education is in-the-moment affect. Educationally relevant affective states (e.g., boredom, confusion, delight, engaged concentration, and frustration [4]) have been shown to have both short and long-term impacts on student outcomes. Short-term impacts include effects on learning gains [5, 6], analytical reasoning [7], motivation [8], and self-efficacy [9] with long-term impacts including college enrollment [10] and career trajectories [11–13]. However, less work has explored whether the learning outcomes predicted by these momentary affective experiences might also be modulated by more enduring affective experiences such as trait-level anxiety.

On their own, both state (in-the-moment) and trait-level anxiety have been shown to impact learning [14] as well as other types of affective experiences and emotion regulation [15]. There has been considerable attention to test anxiety in educational research [16], but work has also considered anxiety surrounding specific subjects or fields [17, 18], including science anxiety – the focus of this paper. Science anxiety is distinct from general anxiety [18]. It can be caused by an array of sources, including lack of role models, gender and racial stereotyping, and the stereotyping of scientists in the popular media. Students suffering from science anxiety are often calm and productive in their nonscience courses, including their mathematics courses, but experience anxiety symptoms in a science setting [19].

Given the context-specific and highly internal nature of anxiety and student regulation (both emotion and learning), these constructs lend themselves to study via interviews. Indeed, students (especially younger students) may find it challenging to articulate metacognitive strategies using traditional self-report instruments such as surveys [20]. This study employs a quantitative ethnography methodology that uses real-time affect detection [4,5] to direct *in situ*, open-ended interviews where students were encouraged to talk about their learning experiences, strategies, and interests. These interviews are then analyzed to examine the interplay between student affect and anxiety. By leveraging real-time affect detection, we target interviews at key moments in the learning process, such as theoretically-aligned affect patterns (e.g., the inhibitory or facilitative cycles identified in [4]). Through this approach, we scale up qualitative data collection to whole classrooms at a time, generating a rich dataset of qualitative (interview) and quantitative (interaction and affect) data. We leverage this combination to explore how anxiety alters student motivation, self-efficacy, and interaction.

1.1 Quantitative Ethnography Research on Emotion

Quantitative Ethnography (QE) is a methodology that blends quantitative and qualitative methods to enable scalable qualitative analysis, using an integrated representation across qualitative and quantitative methods [22].Within education research, the combination of ethnographic and statistical tools supports a deeper examination of what learners do and why [22].

Although Epistemic Network Analysis is the most common method used for QE [23], the key quality of QE research is a "focus on validity and the linkages and consistency between quantitative models and qualitative analysis, both of which are sound on their own, and both of which are attending to the same mechanisms at work in the same set of data." [23, p.2]. Much of the existing work in QE has taken complex data, encoded it (by a human or automated system [24], and then analyzed it using a combination of qualitative and quantitative methods attending to the same themes (e. g. [25]). In this paper, we instead use quantitative analysis to drive the collection of complex, rich data. We then use further quantitative analysis both to understand the phenomena of interest and drive the selection of cases to study qualitatively. By doing so, we analyze our themes around anxiety and – as will be discussed below – frustration using a blend of quantitative and qualitative methods.

Our work studying emotion using QE joins an increasing scientific discourse demonstrating that QE can contribute to theoretical understanding of emotion in context. Martin et al. [26] present one of the first analyses of emotion within a QE framework. Their research used qualitative information (interviews and conversations) in tandem with codified changes in facial expression and body motion to identify moments in which students demonstrated changes in understanding. Their findings demonstrated how instances of joy coincided with periods of increased interaction with the learning platform and knowledge elaboration among participants. The framework adopted within the study displayed potential for the interlacing of human coders and computational techniques towards the development of whole-body analysis that would be capable of identifying proxies of learning in informal learning environments.

Espino et al. [27] also used QE to explore the range of emotions, their valence, and their antecedents during a transition to online learning. Student emotions were analyzed in relation to specific areas of change affected by a switch to online learning (e.g., workload, schedule, teacher interaction). Their findings revealed that in response to factors of focus, instructional format, and workload, anxiety was the most commonly experienced negative emotion. The quantitative results of their study are supplemented by excerpts of student responses that reflect upon their experiences with a new instructional format. This work indicates the key importance of anxiety within online learning and the potential of QE methods for research in this area.

1.2 Other Relevant Past Research on Anxiety

Characterized as a loosely coupled ensemble of cognitive, affective, somatic arousal, and behavioral components, anxiety is evoked in response to mental representations of future threats or danger [28, 29]. Anxiety remains an insufficiently studied emotion in education, primarily studied in terms of test anxiety [30] or anxiety about subject domains [31–33], rather than in terms of its immediate manifestations or impacts.

Recent research has shown that anxiety inhibits working memory [34] and can manifest as avoidance behaviors [35] and increased apprehension around learning [36]. These impacts can, in turn, impede optimum performance [37] and potentially limit future college and career choices [38]. In academic settings, the effects of subject anxiety are perhaps most studied in mathematics (see review in [39]) but have also been examined in other areas, including learning English as a foreign language [40] and science [9, 32, 41].

Prior research has demonstrated the value of quantitative data for understanding anxiety in educational settings. Studies have shown changes in learner experiences in response to anxiety in terms of physiological [42], motivational [41], self-regulated learning [43], and performance [29] components. However, quantitative analysis alone may not be enough to study a state as internal as student anxiety. In addition, quantitative analysis by itself cannot fully capture the experiences and personal justifications that influence student actions, calling for a more blended approach.

2 Methods

2.1 Betty's Brain & Affect Detection

Betty's Brain is an open-ended, computer-based learning environment that uses a learning-by-teaching paradigm to teach complex scientific processes [44]. Betty's Brain, shown in Fig. 1, asks students to teach a simulated virtual agent (Betty) about scientific phenomena (e.g., climate change, ecosystems, thermoregulation) by constructing concept maps that demonstrate the causal relationships involved. As students construct their concept map, they must navigate through multiple hypermedia information sources where they can read about a variety of concepts and their relations. They choose how often to test Betty's knowledge, and they may elect to interact with a virtual mentor agent (a simulated experienced teacher named Mr. Davis) if they are having trouble teaching Betty.

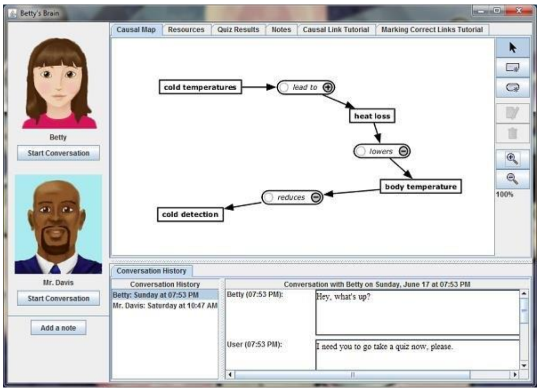

Fig 1. Screenshot showing a partial concept map (right) being used to teach the student agent (Betty, top left), with help from the mentor-agent (Mr. Davis, bottom left).

Real-time affect detection was implemented using models that had been previously integrated into Betty's Brain [21]. Detectors had been trained on ground truth incidences of boredom (BOR), confusion (CON), engaged concentration (ENG), delight (DEL), and frustration (FRU) collected using the BROMP method for conducting quantitative field observations [45]. Each detector generated an affect prediction (a probability between 0 and 1) every 20 seconds based upon previous student interactions.

2.2 Dataset

The dataset was collected at an urban Tennessee middle school, from 99 sixth-graders who used Betty's Brain in their regular science class. This school's population is 60% White, 25% Black, 9% Asian, and 5% Hispanic, with 8% enrolled in the free/reduced lunch program. Individual demographics were not collected. The same students used Betty's Brain to complete two science inquiry scenarios conducted in December 2018 (study 1) and February 2019 (study 2). In study 1, students spent four daily sessions (approx. 50 min/day) using Betty's Brain to complete a causal map about climate change. In study 2, students spent three daily sessions modeling thermoregulation. During both studies, qualitative interviews were used to explore how students were working with the system during specific affective experiences. At the end of study 2, science anxiety surveys were administered alongside questions about students' perceptions of difficulty and familiarity with the topic and questions about Mr. Davis. Between the two studies, minor changes were made to Betty's Brain, including small changes to make Mr. Davis seem more polite. All other procedures were identical.

2.3 Student Interviews

A total of 594 interviews (358 from study 1 and 236 from study 2) were recorded during classroom observations. These interviews were triggered by an app called the Quick Red Fox (QRF; [43]), which used the automated detectors of student affect (see 2.1) to prompt individual interviews. Interviewers were signaled via QRF, which automatically recorded metadata (i.e., timestamps and user IDs).

A prioritization algorithm was used to select which student should be interviewed in cases when multiple students displayed interesting patterns at roughly the same time. We prioritized certain triggers over others in cases where multiple patterns were present in the detector outputs (summarized in Table 1), prioritizing the two affect cycles outlined by D'Mello & Graesser [4] and other theoretically interesting transitions.

Table 1. Top 5 affect patterns in order of prioritization

Rank	Pattern
1	Engaged Concentration → Confusion → Delight → Engaged Concentration (Facilitative Cycle [4])
2	Engaged Concentration → Confusion → Frustration → Boredom (Inhibitory Cycle [4])
3	Confusion → Delight
4	Confusion → Frustration
5	Frustration → Engaged Concentration

This prioritization algorithm also helped to ensure that students were not repeatedly chosen for interviews in a short period of time. If interviewers were not comfortable interrupting a student for an interview, they could skip the prompt, and the system would suggest another student.

Interviewers sought a helpful but non-authoritative approach when speaking with students. They often asked students what their strategies were (if any) for getting through the system. As new patterns and information emerged, questions designed to elicit information about intrinsic interest (e.g., "What kinds of books do you like to read and why?" or "What do you want to be when you grow up?") were sometimes added to see whether or not students would mention science interest outside of their current environment. Overall, however, students were encouraged to speak about what they wanted and to provide feedback about their experience with the software.

Interviews were transcribed, preserving critical metadata such as timestamps and filenames (that contained additional date and time information) as well as information on the interviewer and student being interviewed. These transcriptions were then synchronized with other data streams (e.g., affect predictions and interactions logs) to provide additional context for analysis.

2.4 Anxiety Measure

A revised version of the Math Anxiety Scale (MAS; [46]) was used to collect students' science anxiety scores at the end of study 2. MAS was chosen as it is suitable for younger learners [47] and has been shown to have good test-retest reliability [48]. Recent research with this scale has shown good internal consistency, with Cronbach's α 0.87, calculated from the responses of 250 secondary school students [49].

Questions were adapted to ask students to reflect upon science topics instead of mathematics. The scale collects responses involving learners' general disposition about science beyond their interaction with the current learning environment. This scale contains ten items (five of which are reverse coded) that inquire about students' experiences and opinions on studying science. Items are rated on a 5-point Likert scale where higher scores indicate higher anxiety [46, 47]. Student anxiety scores are calculated by adding the values from each of these ten items. In this study, 95 students completed this survey, leading to the exclusion of four students. Anxiety scores ranged from 10 to 44 with a mean score of 22.35 (SD = 8.24), where 46 students had above-average anxiety (48.42%) and 49 below-average anxiety (51.58%).

3 Results

This study integrates quantitative and qualitative methods to understand how anxiety manifests when students interact with Betty's Brain. Specifically, we quantitatively examined how affect differs over time for anxious students, and how anxious students interacted with the system differently than other students. Then, a qualitative analysis of students' metacognitive awareness (as seen in student interviews) provided a richer, thicker description of the factors leading to quantitative differences between students with different anxiety levels.

3.1 Changes in Affect Over Time

We first consider the impact anxiety might have on affect over time, looking at how far into a learning session (class period) a student was. For each affect inference, we calculated how long it occurred after the student's first interaction that day. We then examined the affective changes and patterns that emerge in the students over the course of each session in relation to their anxiety scores.

Linear mixed-effect models (LMEM) were used to assess how student affect changes over time. These models consider fixed effects, random effects, and noise as linearly contributing to the dependent variable (e.g., affective states; [50]). In these models, anxiety scores and the amount of time since the start of each session were both treated as fixed effects and student IDs as random effects.

Table 2 summarizes the LMEM results across the two studies. In study 1, we observed a significant interaction between anxiety and time into the session when predicting frustration. Student anxiety had no effect on frustration levels early in the session, but later in the session, those levels diverged, and students with above-average levels of anxiety are significantly more likely to get frustrated than their lower-anxiety peers. This result was replicated in study 2, where students with higher anxiety were also significantly less likely to experience boredom, engaged concentration, and delight later in the session.

Table 2. Beta Coefficients for LMEM Predicting Affect (N = 95, $p < .05$) in Study 1 (S1) and Study 2 (S2) (Boredom, Confusion, Frustration, Engagement, Delight). Significant ($p< 0.05$) coefficients shown in bold

	Dependent Variables (S1)					Dependent Variables (S2)				
Predictors	BOR	CON	FRU	ENG	DEL	BOR	CON	FRU	ENG	DEL
(Intercept)	**-0.05**	**-0.05**	**0.01**	**-0.02**	**0.1**	**0.03**	**0**	**0.01**	**0.06**	**0.08**
Anxiety Score	-0.01	-0.03	0.1	0.01	-0.04	0.01	0.04	0.1	0	0
Time Into Session	**-0.28**	**0.35**	**0.43**	**0.02**	0.04	**-0.25**	**0.38**	**0.35**	**0.05**	**0.13**
Interaction (Anxiety* Time Into Session)	0	0	**0.03**	0	0.01	**-0.04**	0	**0.01**	**-0.05**	**-0.01**

MANOVA was conducted to examine the effects of anxiety on frustration across both studies. The results from this analysis revealed that anxiety was associated with differences in the range ($F(1, 93) = 4.88$, $p < 0.01$) and SD ($F(1, 93) = 3.16$, $p = 0.01$) of frustration predictions but not the means ($F(1, 93) = 1.57$, $p = 0.21$), suggesting that students with higher anxiety experience a greater variance of frustration feelings.

We also examined whether anxiety was associated with different patterns of affect across days (i.e., were high anxiety students more likely to feel frustrated on Day 3 of the study than Day 1?). Within this analysis, we averaged each student's affect predictions within each day, and used mixed-effect models to predict the proportion of an affective state by the interaction of anxiety (high/low) and day of the study, with student ID as a random effect, creating one model per affective state per study. For engaged concentration, anxiety showed a significant interaction with day of study ($p = .04$). A simple slopes analysis indicated that higher anxiety students were less likely to experience engaged concentration in later stages of the study. However, no effect of day of study was observed for frustration ($p = .51$) or any of the other momentary affective states investigated in this study.

3.2 Behavioral Differences Based on Affect

In order to further explore the relationship between frustration and anxiety, a second thread of analysis was conducted examining student actions following instances of frustration. We focused on what the students did immediately following an instance of frustration, and then further explored the correctness of those actions (where that analysis was appropriate).

We examined actions in the 20 seconds following an instance of frustration (defined as when frustration had the highest probability among the five affective states). This time interval was chosen to align with both the initial BROMP field observations [45] and the detectors, which generate new affect predictions every 20 seconds [21]. Some students never had a frustration prediction; thus, the filtered data set includes 69 students for study 1 (33 above-average anxiety and 36 below-average anxiety students) and 41 students for the second study (21 above-average anxiety and 20 below-average anxiety students).

We calculated the frequency of 38 types of student interactions (examples in Table 3) within Betty's Brain in each 20-second window, averaged this across students, and then mapped data to each student's pre-test score and anxiety score, including their label as an above-average or below-average anxiety learner. T-tests were performed in order to determine if there were differences in actions between the anxiety groups.

Table 3. Examples of Types of Student Actions (S1 = Study 1; S2 = Study 2)

Action	Definition
Causal Map Elements Moved	the student moved (rearranged, but did not delete) a set of map elements (i.e., concepts and/or links) in the causal map; N = 2049 (S1), 1080 (S2)
Delete Causal Link	the student deleted a causal link between two labels on the causal map; N = 49 (S1), 73 (S2)
Delete Entity	the student deleted a concept (and any links connected to it) from the causal map; N = 252 (S1), 76 (S2)
Quiz Taken	the student gave the teachable agent a quiz, which tests the accuracy of the causal map; N = 94 (S1), 84 (S2)

Only study 2 showed significant differences between above-average and below-average anxiety learners. Students with higher anxiety had a higher tendency to move causal map elements ($t(20) = -2.18$, $p = .04$) and delete entities ($t(20) = -2.42$, $p = .02$) following predictions of frustration while students with below-average anxiety were more likely to delete causal links ($t(19) = 3.37$, $p < .001$) and take quizzes ($t(19) = 2.17$, $p = .04$) after frustration predictions. However, given the substantial number of potential behaviors examined here, after a Benjamini & Hochberg correction [51], only the finding around deletion of causal links remained statistically significant – the other findings must be treated as suggestive but inconclusive.

Further analysis was conducted to determine whether students who were deleting parts of their causal map were doing so in an effective or ineffective manner.

Specifically, Mann-Whitney tests were used to determine whether there were meaningful differences in the proportion of effective and ineffective deletion actions between high versus low anxiety students. Deleting correct answers was treated as ineffective, and deleting incorrect answers was treated as effective.

These results reveal a marginally significant difference for the deletion of concept map entities (Delete Entity actions). High-anxiety students made fewer effective deletions of entities (the labels on the map) than low anxiety students ($U = 12$, $z = -1.78$, $p = 0.07$). However, no difference between groups was found for deletion behavior involving causal links ($U = 17$, $z = -0.47$, $p = 0.63$). The presence of a marginal difference between some deletion actions between groups suggests that less-anxious students may be making more systematic changes to their concept map than more-anxious students.

3.3 Interview Findings

We next turned to interview transcripts to consider any potential differences between high and low anxiety students. We analyzed all the interviews that were conducted within 80 seconds of a frustration prediction (a sub-sample of 28 interviews out of the original 594 conducted). The analysis focused on exploring the relationship between frustration and different anxiety levels from the experiences of students. The difference in behaviors between frustrated/more anxious students and frustrated/less anxious students also corresponded to differences in the interview data. This data for this analysis was drawn from 93 of the students in the original sample, as six students above did not have any interviews. Below, we discuss in detail several representative responses, drawn from 7 of these students. These responses provided particularly illustrative evidence around anxiety and frustration.

Within the analyses below, pseudonyms were assigned using http://random-name-generator.info/, which generates names based on the frequencies within all U.S. census data, ignoring local community or subgroup variation, and ignoring the actual gender or age of the student. Three students with higher anxiety are referred to pseudonymously as Rebecca, Elmer, and Mandy while four students with lower anxiety are given the pseudonyms of Stephanie, Dwight, Marty, and Nancy.

One core difference in behavior noted above was the difference in what more anxious and less anxious students delete when they become frustrated. Less anxious students deleted individual causal links when they were frustrated – focusing on the specific errors in their understanding. By contrast, more anxious students deleted concepts, clearing away significant parts of their map, a less productive behavior. Rebecca articulates this strategy in her interview, stating that

> *I'm just getting really frustrated, I had to delete most of my web because I was getting really confused at it… When I did the quiz, I got 3 of them wrong and 3 of them that he didn't –they didn't know… And so, when I did that I was like, which one did I get wrong, even though I pressed on it, I tried changing the answer but I still got it wrong, so I'm just like, I need to start back over, because then that'll be the easiest for me.*

Rebecca's comments indicate that she felt she could not iterate on her map when some incorrect answers were present and therefore thought it would be easiest to start from scratch. It is possible that her anxiety caused her to feel that she could not productively iterate her solution.

Another core difference in the quantitative results above was the tendency of less anxious students to take quizzes after becoming frustrated. For example, Stephanie, a less-anxious student, reports, “Like, I’m tryin’ to – just tryna figure out what mistakes I could improve. So I can get a better grade on the other test,” showing the interviewer her quiz result. Similarly, Dwight, another less-anxious student, notes:

> *Um, so, I’ve given him a lot of tests, and the ones that he gets wrong, I’m just trying to go through and like…see where I can add something, or like, if it’s wrong… And then that helps me, and I just did that and he got a higher grade, so it’s been working.*

Like the behavioral findings, these interviews reflect a greater willingness among lower-anxiety students to collect evidence on where and how they are wrong.

By contrast, students with higher science anxiety are apprehensive about quizzes. Not only do they not take as many quizzes (as documented in the behavioral findings); even when they do, they do not use them very productively. For example, Elmer, a more anxious student, focuses on taking quizzes where he will get a better score:

> *Um, so, I realized that I hadn’t taken an everything quiz, and I was like, well, of course I haven’t, because I’m not getting anything good on the other quizzes, and then I took one, and I got one right, and the other ones are just getting everything wrong… So I’ve just been taking everything quizzes [about the whole concept map] instead of the separate quizzes [about individual parts of the map] because I’ve been getting better scores.*

Jon, another more anxious student, reports a similar strategy: “Um, what I’m doing is just quizzing my guy. And…I’m quizzing him on the…bodily responses to cold. Because I think I’ll do better than that one.” Instead of using the quizzes to target improvements to their causal map, these students appear to be using the quizzes in areas they are confident in to calm the distress they are feeling about not having all of the answers.

More generally, the more anxious students seemed to interpret struggles with the material as communicating something more general about their competence. For example, during various interviews over the course of both studies, Mandy made general statements about science (e.g., “It’s hard,” or “science is hard…”) or the specific topic she was studying in Betty’s Brain (e.g., “I don’t really get thermoregulation…”).

Another difference that emerged within the interview data was differences in responses to the pedagogical agent, Mr. Davis. For example, when the interviewer asked Rebecca (a more anxious student) if Mr. Davis gave her anything useful, she responded,

> *He said to just remember that we’re goin – that when we’re…doing this, we’re learning and are going to get more smart…It made me feel even more frustrated. Because, like, I want, like, really useful information, not really motivation because I don’t need the motivation, I get frustrated because I don’t have the information.*

Upon further prompting, Rebecca noted that Mr. Davis had given her information but that she had not found it useful:

> *He told me to go to the [virtual textbook] if I’m having any trouble… I went there but I don’t really understand it… it says, “plants increase oxygen.” It doesn’t really tell how it increases the oxygen. It doesn’t really tell how you use it or anything. It’s not really being specific.*

Elmer, another more anxious student, commented that "He's had one thing that he said that was useful, but everything else has just been like, 'Hm, let's take a look at your progress.'" In other words, high-anxiety students seemed to be unable to process the help they were being given.

By contrast, several less anxious students seemed better able to work with the advice being provided by Mr. Davis. For example, Marty reported, "Well, I know, like, what to do if I don't understand, like, how to do something. Like I can ask Mr. Davis if I don't understand how." Similarly, Nancy told an interviewer,

> *Um, he's been helpful, he's, um. What I really like is that he's not just recommending pages in the teacher's book, he's telling us where our stuff is wrong, more exact… What I really like is that he's not just recommending pages in the teacher's book, he's telling us where our stuff is wrong, more exact.*

Dwight, another less anxious student, also stated "he told me where I should look for something and then that helped. It was pretty helpful."

In general, the findings in these interviews demonstrate a pattern where students with higher anxiety appear to be less able to regulate their learning process, perhaps because they see it as something they are unable to fully control. This may cause them to either overlook or deliberately skip activities that might be essential for their learning, possibly because they do not recognize the opportunities provided to them.

4 Discussion

This study employs a quantitative ethnographic approach, including the analysis of students' metacognitive strategies (articulated in ethnographic interviews) to better understand how student anxiety influences behaviors in computer-based learning. We do this by examining multiple aspects of learner experiences involving both state-level and trait-level anxiety. By using real-time affect detection to identify critical points in the learning process, we guided qualitative data collection to explore how anxiety alters student motivation, self-efficacy, and interaction. In this work, we analyze two studies where students used Betty's Brain to construct causal maps of complex scientific phenomena in middle school science classes, comparing students who score above and below the cohort's average on a science anxiety scale.

Across both studies, the findings show that trait-level anxiety measures help quantitatively differentiate state-level affective experiences, which influences both students' behaviors and their metacognitive descriptions of their strategies. Specifically, students with higher trait-level science anxiety are more likely to experience frustration, but only in the second half of a daily one-hour session. At the beginning of the day, both more- and less-anxious students are equally likely to be frustrated; as the day goes on, frustration is more likely for high anxiety students than low-anxiety students.

Interestingly, while only frustration showed temporal divergence in the first study, two other states showed divergence in the second study. Specifically, higher-anxiety students were less likely to experience engaged concentration and delight in the second half of each learning session. This complements the finding that they also experienced higher levels of frustration during this time period. These changes are potentially driven by anxious fixation on challenges where students' engagement is impeded by increased focus on the source of frustration [52]; anxiety may also impede experiences of delight when students are unable to achieve their own standards of success [53].

Analysis of student interviews that followed instances of frustration provided additional insights into students' experiences. Excerpts from student interviews ground the quantitative data by providing in-depth information pertaining to participants' experiences and viewpoints on their use of Betty's Brain. From this we are better able to contextualize codified affect and behavior in elaborated narratives [54].

Previous research has demonstrated that anxiety may manifest in a variety of ways during learning, such as speeding through problems [35] and disengagement from activities [55]. We found that higher anxiety learners were more likely to make inefficient changes as a result of their deletions and edits, indicative of uncertainty that can lead to concern with unimportant details and inability to understand knowledge content [56]. We posit that the behavioral differences may only emerge within the second study due to the effect of the novelty of the interactive platform in study 1, which wore off in study 2, resulting in more pronounced engagement effects [57].

4.1 Limitations and Future Work

One potential limitation of this work is the different timescales of crucial measurements. While we measured trait-level anxiety via a survey, we considered state-level affect with predictions being generated every 20 seconds. Future work should examine anxiety at a similar time scale to affect, though this would require a state-level anxiety measure. Anxiety is generally not included in classroom observation affect protocols [45], as it is harder to measure in this fashion, so additional effort will be needed to collect the training data needed to develop an automated detector of anxiety.

If state-level anxiety could be automatically measured, this would also open up possibilities for future qualitative work. Using the same targeted interview approach used in this paper, interviewers could be directed to students at moments of peak anxiety. Such an approach could yield a rich qualitative dataset that could be used to better understand the interaction between anxiety, affect, and student learning, at a finer grain size, contributing both to affect and learning theory [58].

4.2 Conclusion

This study used interaction data, affect measurements, and qualitative data from targeted student interviews, to study how trait-level anxiety relates to affect and behavior during a complex learning activity. Our results provide insight into the relationship between anxiety and emotion, most notably frustration, during learning experiences. More-anxious students were more likely to be frustrated and less likely to be engaged and delighted later in the sessions. Students with higher levels of anxiety had a higher tendency to delete entire sections of their work when frustrated, while frustrated students with lower anxiety were more likely to delete individual elements and take quizzes to better understand the quality of their solutions. This work takes a step towards better understanding anxiety in an ecologically valid setting, and brings us closer to better supporting high anxiety students using learning technology in the classroom.

References

1. Schofield, J.W.: Computers and classroom culture. Cambridge University Press (1995)
2. Wessel, D.: The Potential of Computer-Assisted Direct Observation Apps. Int. J. Interact. Mob. Technol. 9, (2015)
3. Stetler, C.B., Caramanica, L.: Evaluation of an evidence-based practice initia-tive: Outcomes, strengths and limitations of a retrospec-tive, conceptually-based approach. Worldviews Evidence-Based Nurs. 4, 187–199 (2007)
4. D'Mello, S.K., Graesser, A.: Dynamics of affective states during complex learning. Learn. Instr. 22, 145–157 (2012). https://doi.org/10.1016/j.learninstruc.2011.10.001
5. DeFalco, J.A., Rowe, J.P., Paquette, L., Georgoulas-Sherry, V., Brawner, K., Mott, B.W., Baker, R.S., Lester, J.C.: Detecting and Addressing Frustration in a Serious Game for Military Training. Int. J. Artif. Intell. Educ. 28, 152–193 (2018). https://doi.org/10.1007/s40593-017-0152-1
6. D'Mello, S.K., Olney, A., Person, N.: Mining collaborative patterns in tutorial dialogues. J. Educ. Data Min. (2010)
7. D'Mello, S.K., Person, N.K., Lehman, B.: Antecedent-Consequent Relationships and Cyclical Patterns between Affective States and Problem Solving Outcomes. In: Artificial Intelligence in Education. pp. 57–64 (2009)
8. Rodrigo, M.M.T., Anglo, E.A., Sugay, J.O., Baker, Rsj.: Use of unsupervised clustering to characterize learner behaviors and affective states while using an intelligent tutoring system. In: International Conference on Computers in Education. pp. 57–64 (2008)
9. McQuiggan, S.W., Robison, J.L., Lester, J.C.: Affective transitions in narrative-centered learning environments. J. Educ. Technol. \& Soc. 13, 40–53 (2010)
10. San Pedro, M.O.Z., Baker, R.S., Heffernan, N.T.: An integrated look at middle school engagement and learning in digital environments as precursors to college attendance. Technol. Knowl. Learn. 22, 243–270 (2017)
11. San Pedro, M.O., Ocumpaugh, J., Baker, R.S., Heffernan, N.T.: Predicting STEM and Non-STEM College Major Enrollment from Middle School Interaction with Mathematics Educational Software. In: Educational Data Mining. pp. 276–279 (2014)
12. San Pedro, M.O.Z., Baker, R.S., Heffernan, N.T., Ocumpaugh, J.L.: Exploring college major choice and middle school student behavior, affect and learning: what happens to students who game the system? In: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge. pp. 36–40 (2015)
13. Almeda, M.V., Baker, R.S.: Predicting Student Participation in STEM Careers: The Role of Affect and Engagement during Middle School. J. Educ. Data Min. 12, 33–47 (2020)
14. Byron, K., Khazanchi, S.: A meta-analytic investigation of the relationship of state and trait anxiety to performance on figural and verbal creative tasks. Personal. Soc. Psychol. Bull. 37, 269–283 (2011)
15. Cisler, J.M., Olatunji, B.O., Feldner, M.T., Forsyth, J.P.: Emotion regulation and the anxiety disorders: An integrative review. J. Psychopathol. Behav. Assess. 32, 68–82 (2010)
16. Zeidner, M.: Test Anxiety in Educational Contexts. Concepts, Findings, and Future Directions. In: Emotion in Education (2007)
17. Tobias, S.: Math Anxiety: An Update. NACADA J. (1990). https://doi.org/10.12930/0271-9517-10.1.47
18. Mallow, J. V., McDermott, L.C.: Science Anxiety: Fear of Science and How to Overcome It. Am. J. Phys. (1988). https://doi.org/10.1119/1.15495
19. Udo, M.K., Ramsey, G.P., Mallow, J. V.: Science anxiety and gender in students taking general education science courses. J. Sci. Educ. Technol. (2004). https://doi.org/10.1007/s10956-004-1465-z

20. Schooler, J.W., Ohlsson, S., Brooks, K.: Thoughts beyond words: When language overshadows insight. J. Exp. Psychol. Gen. 122, 166 (1993)
21. Jiang, Y., Bosch, N., Baker, R.S., Paquette, L., Ocumpaugh, J., Andres, J.M.A.L., Moore, A.L., Biswas, G.: Expert Feature-Engineering vs. Deep Neural Networks: Which is Better for Sensor-Free Affect Detection? In: Artificial Intelligence in Education. pp. 198–211 (2018)
22. Shaffer, D.W.: Quantitative ethnography. Cathcart Press (2017)
23. Kaliisa, R., Misiejuk, K., Irgens, G.A., Misfeldt, M.: Scoping the Emerging Field of Quantitative Ethnography: Opportunities, Challenges and Future Directions. In: International Conference on Quantitative Ethnography. pp. 3–17 (2021)
24. Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D.W., Hu, X., Graesser, A.C.: nCoder+: a semantic tool for improving recall of nCoder coding. In: International Conference on Quantitative Ethnography. pp. 41–54 (2019)
25. Nguyen, H., Ahn, J., Belgrave, A., Lee, J., Cawelti, L., Kim, H.E., Prado, Y., Santagata, R., Villavicencio, A.: Establishing Trustworthiness Through Algorithmic Approaches to Qualitative Research. In: Advances in Quantitative Ethnography: Second International Conference, ICQE 2020, Malibu, CA, USA, February 1-3, 2021, Proceedings. p. 47 (2021)
26. Martin, K., Wang, E.Q., Bain, C., Worsley, M.: Computationally Augmented Ethnography: Emotion Tracking and Learning in Museum Games. In: International Conference on Quantitative Ethnography. pp. 141–153 (2019)
27. Espino, D.P., Wright, T., Brown, V.M., Mbasu, Z., Sweeney, M., Lee, S.B.: Student Emotions in the Shift to Online Learning During the COVID-19 Pandemic. In: International Conference on Quantitative Ethnography. pp. 334–347 (2021)
28. American Psychological Association (APA): Diagnostic and Statistical Manual of Mental Disorders: Depressive Disorders. (2013)
29. Zeidner, M.: Anxiety in education. In: International Handbook of Emotions in Education (2014)
30. von der Embse, N., Jester, D., Roy, D., Post, J.: Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. J. Affect. Disord. 227, 483–493 (2018)
31. Horwitz, E.K.: Language anxiety and achievement. Annu. Rev. Appl. Linguist. 21, 112–126 (2001)
32. Mallow, J. V: Science anxiety: research and action. Handb. Coll. Sci. Teach. 3–14 (2006)
33. Foley, A.E., Herts, J.B., Borgonovi, F., Guerriero, S., Levine, S.C., Beilock, S.L.: The math anxiety-performance link: A global phenomenon. Curr. Dir. Psychol. Sci. 26, 52–58 (2017)
34. Ashcraft, M.H., Kirk, E.P.: The relationships among working memory, math anxiety, and performance. J. Exp. Psychol. Gen. (2001). https://doi.org/10.1037//0096-3445.130.2.224
35. Ashcraft, M.H.: Math anxiety: Personal, educational, and cognitive consequences. Curr. Dir. Psychol. Sci. 11, 181–185 (2002)
36. Plake, B.S., Parker, C.S.: The development and validation of a revised version of the Mathematics Anxiety Rating Scale. Educ. Psychol. Meas. 42, 551–557 (1982)
37. Hong, Z.-R.: Effects of a collaborative science intervention on high achieving students' learning anxiety and attitudes toward science. Int. J. Sci. Educ. 32, 1971–1988 (2010)
38. Ashcraft, M.H., Krause, J.A.: Working memory, math performance, and math anxiety. Psychon. Bull. \& Rev. 14, 243–248 (2007)
39. Suárez-Pellicioni, M., Núñez-Peña, M.I., Colomé, À.: Math anxiety: A review of its cognitive consequences, psychophysiological correlates, and brain bases. Cogn. Affect. \& Behav. Neurosci. 16, 3–22 (2016)
40. Na, Z.: A study of high school students' English learning anxiety. Asian EFL J. 9, 22–

34 (2007)

41. González, A., Fernández, M.V.C., Paoloni, P.V.: Hope and anxiety in physics class: Exploring their motivational antecedents and influence on metacognition and performance. J. Res. Sci. Teach. (2017). https://doi.org/10.1002/tea.21377
42. Clarke, S., Horeczko, T., Cotton, D., Bair, A.: Heart rate, anxiety and performance of residents during a simulated critical clinical encounter: a pilot study. BMC Med. Educ. 14, 1–8 (2014)
43. Hutt, S., Ocumpaugh, J., Andres, J.M.A.L., Munshi, A., Bosch, N., Baker, R.S., Zhang, Y., Paquette, L., Slater, S., Biswas, G.: Who's Stopping You? – Using Microanalysis to Explore the Impact of Science Anxiety on Self-Regulated Learning Operations. In: Proceedings of the 43rd Annual Conference of the Cognitive Science Society (2021)
44. Biswas, G., Leelawong, K., Belynne, K., Schwartz, D., Davis, J.: Incorporating Self Regulated Learning Techniques into Learning by Teaching Environments. In: The Twenty Sixth Annual Meeting of the Cognitive Science Society (2004)
45. Ocumpaugh, J., Baker, R.S., Rodrigo, M.M.T.: Baker Rodrigo Ocumpaugh monitoring protocol (BROMP) 2.0 technical and training manual. (2015)
46. Betz, N.E.: Prevalence, distribution, and correlates of math anxiety in college students. J. Couns. Psychol. (1978). https://doi.org/10.1037/0022-0167.25.5.441
47. Johnston-Wilder, S., Brindley, J., Dent, P.: A survey of Mathematics Anxiety and Mathematical Resilience among existing apprentices. (2014)
48. Pajares, F., Urdan, T.: Exploratory factor analysis of the mathematics anxiety scale. Meas. Eval. Couns. Dev. (1996)
49. Mahmood, S., Khatoon, T.: Development and validation of the mathematics anxiety scale for secondary and senior secondary school students. Br. J. Arts Soc. Sci. 2, 169–179 (2011)
50. Barr, D.J., Levy, R., Scheepers, C., Tily, H.J.: Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. 68, 255–278 (2013)
51. Benjamini, Y., Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. 57, 289–300 (1995). https://doi.org/10.2307/2346101
52. McQuiggan, S.W., Lee, S., Lester, J.C.: Early prediction of student frustration. In: International Conference on Affective Computing and Intelligent Interaction. pp. 698–709 (2007)
53. Pekrun, R.: The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. Educ. Psychol. Rev. 18, 315–341 (2006). https://doi.org/10.1007/s10648-006-9029-9
54. Turner III, D.W.: Qualitative interview design: A practical guide for novice investigators. Qual. Rep. 15, 754 (2010)
55. Kirsch, M., Windmann, S.: The role of anxiety in decision-making. Rev. Psychol. 16, 19–28 (2009)
56. Rosenfeld, R.A.: Anxiety and learning. Teach. Sociol. 151–166 (1978)
57. Tsay, C.H.-H., Kofinas, A.K., Trivedi, S.K., Yang, Y.: Overcoming the novelty effect in online gamified learning systems: An empirical evaluation of student engagement and performance. J. Comput. Assist. Learn. 36, 128–146 (2020)
58. Willis, M., Cromby, J.: Bodies, Representations, Situations, Practices: qualitative research on affect, emotion and feeling, (2020)