Analysis of exploratory behaviour: A step towards modelling of curiosity

Joe Tang
Queensland University of Technology
Brisbane, Queensland, Australia
q2.tang@qut.edu.au

Andrew Gibson
Queensland University of Technology
Brisbane, Queensland, Australia
andrew.gibson@qut.edu.au

Peter Bruza
Queensland University of Technology
Brisbane, Queensland, Australia
p.bruza@qut.edu.au

Abstract

In this research, we analysed exploratory behaviour trace data for students engaging in learning tasks in a technology-enhanced data analytics course as the first step towards modelling curiosity in learning. Curiosity is a complex phenomenon that is not amenable to direct modelling, but it can be understood through related behaviours like exploration, which is critical to effective learning. We analysed trace data from 40 students using visualisation and network analysis techniques, focusing on their interactions with learning tasks within the JupyterLab environment. Our analysis found that providing sufficient exploration time before explicit instruction or answer revelation, and designing learning tasks that embrace errors as opportunities, encouraged behaviours associated with curiosity-driven learning. These findings highlight the importance of designing learning environments that foster curiosity and promote active exploration.

CCS Concepts

• **Applied computing** → **Education**; **Interactive learning environments**; • **Human-centered computing**;

Keywords

exploratory behaviour, curiosity-driven learning, task-centric approach, trace data

ACM Reference Format:

Joe Tang, Andrew Gibson, and Peter Bruza. 2025. Analysis of exploratory behaviour: A step towards modelling of curiosity. In *LAK25: The 15th International Learning Analytics and Knowledge Conference (LAK 2025), March 03–07, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3706468.3706561

1 Introduction

Curiosity-driven exploration is critical to effective learning [26]. However, in the context of increasingly powerful GenAI technologies, there are more temptations than ever for students to take shortcuts to solutions rather than actively exploring problems. In this environment, the importance of encouraging student exploratory behaviour has become increasingly significant.

LAK 2025, March 03–07, 2025, Dublin, Ireland

ACM ISBN 979-8-4007-0701-8/25/03
https://doi.org/10.1145/3706468.3706561

Although not all students will fall into the trap of shortcut-seeking, the risk of more students succumbing to the temptation of easy wins has increased in the current learning environment. On the other hand, curiosity has a strong influence on learning by motivating individuals to seek new information and experiences [11]. This intrinsic desire to explore is crucial for cognitive growth and innovation, particularly in education [7].

Curiosity-driven exploration allows learners to explore learning content without setting formal goals or expecting external rewards [19]. In contrast, more structured learning approaches tend to involve incentives like rewards [25]. For example, Self-Regulated Learning (SRL), involves setting clear objectives and making an effort to achieve them, setting up an achievement reward for effort. Effective learning tends to involve degrees of both structured approaches (e.g. for direction and organisation) and unstructured curiosity-driven exploration (e.g. for diversity and depth).

This paper focuses on the use of learning analytics to expose exploratory behaviour as an observable activity of curiosity. The study represents an initial step towards a larger program of research on modelling curiosity within the context of learning analytics. In this work we identify behavioural patterns that might indicate students' curiosity by examining how they engage with a data analytics learning platform. With an ultimate aim of helping educators and learning designers create learning environments that promote active exploration and knowledge-seeking, the research focuses on the following research question: **What insights can be obtained through the analysis of trace data in terms of student exploratory behaviours?**

The study adopted a task-centric learning analytics approach to investigate how exploratory behaviours in a technology-enhanced learning environment related to student learning processes. The approach emphasised the analysis of learner interactions within the context of specific learning tasks and the technology used to support them [14]. We examined trace data from a graduate-level data analytics course to understand events such as code execution, encountering errors, navigation patterns, and copying and pasting behaviour. Our findings suggested that students who engaged in active exploration demonstrated a deeper understanding of the course material. These students went beyond immediate task requirements, sought additional information and experimented with different approaches. Conversely, students focused solely on completing tasks displayed a more superficial understanding of the learning content. This observation suggests that integrating goal-oriented tasks with more exploratory activities can provide students a more comprehensive understanding of the learning material.

Our study findings hold significance for learning analytics firstly in the value for learning of analysing trace data related to exploratory behaviour, and secondly in the use of these analytics as one possible indicator of curiosity. Further, data analytics and visualisations created during the study hold potential for real-time analytics of exploratory behaviour which may help support the cultivation of curiosity in learning.

2 Literature Review

2.1 Curiosity and Exploratory Behaviour in Learning

The theory surrounding curiosity suggests it is a multi-dimensional construct [7] with two dominant perspectives. The first perspective centres on the concept of a 'knowledge gap', proposing that curiosity arises from a desire to close the gap between 'what one knows and what one wants to know' [18, p. 87]. We refer to this knowledge gap dimension as the epistemic dimension. The second perspective highlights the experiential aspect of curiosity, suggesting that individuals are drawn to novel, complex, or uncertain situations for the sheer enjoyment of exploration and the anticipation of new experiences [17]. We refer to this as the experiential dimension. This research described in this paper is concerned with the epistemic dimension of curiosity.

Curiosity is important for learning and cognitive development. It is defined by a strong desire for new knowledge, motivating individuals to explore and seek information [29]. This intrinsic desire forms the foundation of active learning and exploration, as shown by research in different disciplines, especially education [7, 9]. Instead of passively receiving information, curious people tend to pursue answers actively [29]. This search for knowledge often stems from a knowledge gap or uncertainty [18, 20]. Learners want to make sense of their environment, which motivates them to search for new information [19]. We refer to this process of seeking new information and understanding their environment as curiosity-driven exploration.

While the term 'curiosity-driven exploration' is not often explicitly used in the literature, we use this term to describe instances where learners demonstrate exploration and information-seeking behaviours motivated by a desire to resolve uncertainty or knowledge gaps. This interpretation aligns with the definitions of curiosity and exploratory behaviour discussed in this section.

Curiosity-driven exploration is observed and represented as exploratory behaviour driven by an intrinsic desire to learn and understand more deeply [3, 8]. For example, Jirout and Klahr [8] defined measures of curiosity as 'the threshold of desired uncertainty in the environment which leads to exploratory behavior' (p. 157). Their findings indicated that children who were more curious engaged more in exploratory play, which in turn supports cognitive development and problem-solving skills. These behaviours are not only about filling knowledge gaps but also reflecting their intrinsic desire to learn and understand more deeply. This finding has implications for the design of learning analytics. These systems could pique curiosity and motivate learners to explore more actively by identifying and prioritising knowledge gaps [28].

Learners who engage in curiosity-driven exploration often exhibit a better comprehension of concepts and active participation in learning [1]. This form of exploration is linked to positive learning outcomes because it encourages learners to move beyond immediate task requirements and explore broader contexts [22, 29]. Therefore, learning platforms that incorporate elements fostering curiosity and exploration could lead to more effective learning experiences. For example, incorporating opportunities for learners to generate questions within a digital learning environment could lead to increased curiosity, information-seeking, and better knowledge retention [12]. Furthermore, asking questions, particularly those that are novel, is positively associated with creative performance on divergent thinking tasks [15].

Despite these benefits of curiosity-driven exploration, educators and learning designers may find it beneficial to balance it with goal-oriented learning, especially in structured learning environments. This is because explorations are open-ended and can be unstructured and unpredictable, while structured learning approaches provide a clear vision to keep learners on track. This raises questions about how educators and learning analytics systems can effectively support both approaches. This paper shows how educators may use insights from learning analytics to encourage curiosity-driven exploration in a structured learning environment.

2.2 Leveraging Learning Analytics to Understand and Encourage Curiosity-Driven Explorations

Recent studies in learning analytics have shown that learning analytics are effective in tracking performance and completion rates. For example, Divjak et al. [5] and Kaliisa et al. [10] demonstrated using learning analytics dashboards to provide insights into student progress and predict outcomes. Similarly, studies explored the application of learning analytics in understanding student engagement [23], predicting student dropout [6], and measuring graduate attribute development [2], all of which contribute to a data-informed understanding of learning outcomes.

Focusing solely on those outcome-oriented metrics can have the unintended effect of obscuring the complex learning processes behind them. For example, exploratory behaviour is a vital part of learning, but that does not necessarily mean that it is visible in traditional measures of 'success', such as test scores or completion rates. A student who spends extended time investigating a concept or experimenting with different approaches, for instance, might not immediately show high performance on a standardised assessment. Nevertheless, this behaviour can lead to a deeper, more flexible understanding of the subject in the long term. Therefore, it is important to develop methods in learning analytics that can capture and analyse these complex learning processes. We claim that exploratory behaviour is one of these, and including it is essential to gain a more comprehensive view of learning.

When capturing learner interactions with digital learning platforms, the trace data can offer a promising avenue for understanding activities associated with learning. Unlike traditional metrics that summarise learning outcomes, trace data can provide a finer grained temporal record of how learners navigate through learning environments, engage with content, and approach learning activities. Researchers can use this information to uncover hidden patterns and deviations indicative of complex learning processes [4, 30, 32].

For example, studies used trace data to examine the sequences of actions learners took while solving problems, revealing the use of different approaches and strategies [21, 30, 32]. Furthermore, trace data analysis was used to understand how learners interact with scaffolding and feedback, providing insights into their help-seeking behaviours and their ability to adapt their learning strategies [30].

In this study, we used trace data to uncover patterns and insights related to exploratory behaviour leading to a better understanding of how curiosity-driven exploration happens in learning environments. Techniques like trace data analysis can help identify and understand exploratory behaviours driven by curiosity, even when curiosity is not explicitly mentioned. For example, Li et al. [16] highlighted the dynamic nature of strategy adoption in learning and found students often deviate from their planned learning strategies. This finding indicates that students' actions are influenced by factors other than predetermined plans: they may be interested in exploring different approaches or unexpected findings during the learning process. Additionally, studies have emphasised the importance of understanding how learners navigate learning environments, engage with content, and approach problem-solving activities [16, 21]. These explorations, even when deviating from structured support or planned strategies, can be an indicator of the learner's curiosity and drive for deeper understanding.

While it is important to acknowledge that these studies did not explicitly focus on 'curiosity-driven exploration', they provided a foundation for examining behaviours that could be seen as indicators of curiosity in learning environments. We used some of the approaches from these previous studies as a first step towards modelling curiosity in this study. Our focus was on the analysis of student exploratory behaviours in a technology-enhanced learning environment, specifically addressing the knowledge gap dimension. By studying these behaviours through trace data analysis, and capturing actions like navigation patterns, resource access, code modifications and question-asking behaviours [24], we can better understand how curiosity manifests in learning environments. This understanding is crucial for creating learning experiences that inspire exploration.

3 Methodology

As this research was an initial step towards the larger goal of modelling curiosity, the study focused on analysis of data from a single postgraduate data analytics course at an Australian university. The course, designed for students with basic programming knowledge, focused on quantitative and computational analytics using Python within the JupyterLab environment[1] (see Figure 1). With its focus on practical coding tasks, the course provided a rich context for examining how learners explore and engage with programming concepts. Specifically, we examined student interactions with intentionally designed 'learning tasks' embedded within the course materials (task examples are included in Figure 1). These tasks required students to apply concepts by filling in missing code snippets, and were designed to encourage exploratory behaviour by presenting manageable challenges that prompt students to seek information and experiment with different approaches. This approach is in line with the principles of task-centric learning analytics model [14], in which technology should be intentionally integrated with the design and implementation of learning tasks to effectively analyse and support student learning.

We obtained an opt-out consent of participation from the university ethics committee, meaning data from students who enrolled in the course would automatically included in the experiment unless they requested to opt out. No students chose to exercise their ethical right to opt out. Aligned with the task-centric learning analytics model, we drew on two primary data sources to gain insights into student learning behaviours: video recordings of the lectures and trace data collected from JupyterLab.

First, lectures were recorded to identify and timestamp learning tasks, synchronising them with student activity traces. Recordings only captured the lecturer's voice and screen broadcast while ensuring student anonymity. To establish a structured framework for analysing student interactions within the context of learning tasks, the lecture recordings were manually annotated to identify and timestamp the start and end points of each distinct learning task presented during the lectures.

Second, trace data was captured using the JupyterLab telemetry plugin[2], recording student activities within the Jupyter environment. Initially, 14 events were captured by the JupyterLab telemetry plugin. However, upon closer examination of the data, it was observed that the 'save_notebook' events were frequently triggered automatically by the system rather than by explicit student action. Given that this type of event did not indicate user-initiated behaviour and could potentially introduce noise into the analysis of curiosity-driven exploration, it was excluded from further investigation. The subsequent analysis focuses on the remaining 13 events, all of which reflect deliberate student interactions within the JupyterLab environment. The data was categorised into broader concepts based on their functionalities in Jupyter: active exploration, navigation within/between notebooks, cell errors, cell execution, and file handling (see Table 1 for event descriptions). These categories also reflected different aspects of student exploration within the JupyterLab environment.

The study employed visual analysis of patterns and flows in time series to examine student interaction patterns during learning tasks. The analysis focused on student interaction patterns during learning tasks, examining the frequency and sequence of events within and across the identified categories. For example, a higher frequency of code testing (cell execution events) coupled with navigation between notebooks could indicate a deeper engagement with the learning material, reflecting a student's effort to connect different concepts and test their understanding. To visualise these patterns, we used timelines to represent individual student activity. These timelines plotted specific events, such as code execution or navigation between files, against time, with the lecturer's actions serving as reference points. Additionally, we used network diagrams to visualise the flow of events within different phases of the learning tasks. In these diagrams, nodes represented distinct event categories (e.g., 'Cell Error', 'File Handling'), and the connections between nodes illustrated the sequence and frequency of these events. The approach aimed to uncover how different exploratory

[1] https://jupyter.org/

[2] The JupyterLab telemetry plugin was developed by the Educational Technology Collective group at the University of Michigan. Source code can be found at https://github.com/educational-technology-collective/etc_jupyterlab_telemetry_library.

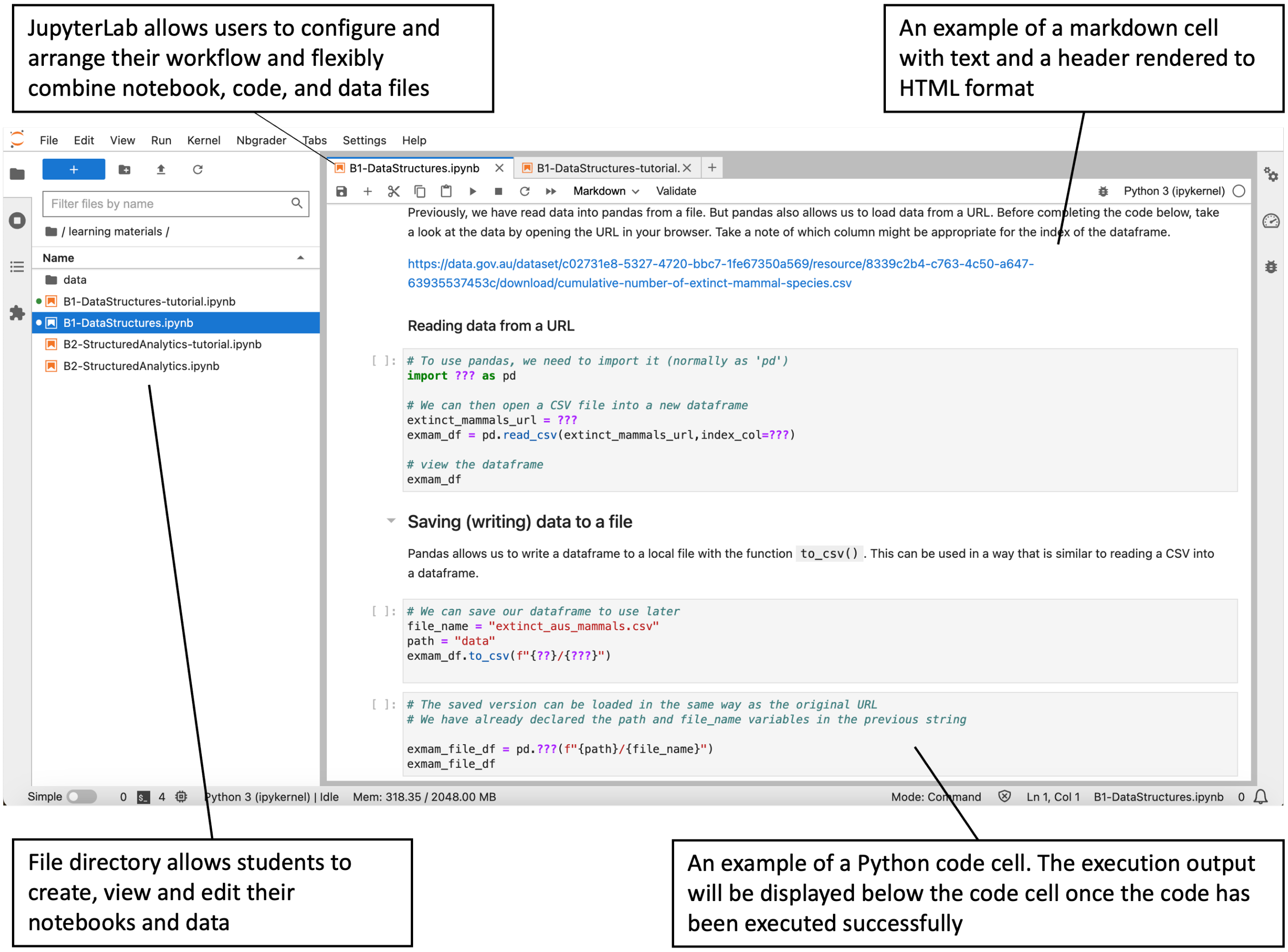

Figure 1: JupyterLab interface used in the course, demonstrating how learning tasks are embedded within Python code cells. Learning tasks for the experiments are displayed in the Python code cells.

behaviours manifested and whether they correlated with deeper engagement in learning.

We investigated how students interacted with learning tasks with a focus specifically on Weeks 2 and 3 of the lectures. We chose this timeframe strategically for several reasons. First, this period allowed for a concentrated analysis of student interactions with learning tasks, aligning with the study's task-centric methodology. Second, it minimised the potential influence of upcoming assignment deadlines on student behaviours. As further analysis revealed that students tended to shift their focus towards assignments as deadlines approached, potentially reducing their engagement with regular learning activities. Weeks 2 and 3 were positioned at the beginning of the semester with the first assignment still weeks away, mitigating this factor. Further, we observed that students tended to be more engaged with in-class activities at the beginning of a semester. Analysing this period, therefore, provided a clearer representation of students' natural interaction with learning tasks, minimising external pressures. Overall, the data encompassed interactions from 40 students, with 33 attending Week 2's lecture, 34 attending Week 3's lecture, and 27 students present for both. This resulted in a rich dataset of 16 learning tasks (8 per week).

We selected two consecutive tasks from Week 3 to investigate the influence of dedicated exploration time on students' behaviours in-depth. We made this selection for the following reasons. First, these tasks exemplified the observed variations in exploration time and allowed for a focused analysis of behaviour within a limited timeframe. Second, analysing consecutive tasks minimised potential variations in student behaviour due to changes in topics or task complexity. This ensured the analysis remained its focus on the impact of exploration time. Last, this focused selection facilitates a more manageable, in-depth analysis within a well-defined context. Although the in-depth analysis focused on these two tasks, the broader network analysis visualised event flows within different phases. The network analysis incorporated data from all 16

Table 1: Categorisation of JupyterLab telemetry events by relevance to exploratory behaviour.

Category	Description	Events
Active Exploration	Reflects active engagement in seeking and experimenting with information, such as copying/pasting code, adding/removing cells, and testing code functionality.	*clipboard_copy* *clipboard_cut* *clipboard_paste* *add_cell* *remove_cell*
Navigation within Notebook	Indicates movement within a single notebook, suggesting engagement with existing content and potential revisiting of previous information.	*scroll* *active_cell_changed*
Navigation between Notebooks	Suggests seeking information beyond the immediate task, potentially referring back to previous content or exploring related concepts.	*notebook_visible* *notebook_hidden*
Cell Error	Represents encountering errors during code execution, offering insights into how students respond to challenges and whether they persist in finding solutions.	*cell_errored*
Cell Execution	Indicates successful running of code, reflecting active participation and experimentation with the material.	*cell_executed*
File Handling	Includes actions like opening, closing, and saving notebooks, providing a general overview of engagement but considered less directly relevant to exploratory behaviour in this study's context.	*open_notebook* *close_notebook*

tasks across two weeks, ensuring a comprehensive examination of exploratory behaviours within the course.

4 Analysis and Results

Annotations on videos recordings of lectures revealed four general phases were present in the learning tasks. Thus we segemented each task according to the phases reflecting the pedagogical flow of the lecture:

- **(TP1)** Lecturer Introducing Task: This phase encompassed the lecturer's initial explanation of the task, including its objectives, relevance to the course content, and any specific instructions for completion.
- **(TP2)** Students Working on Task: This phase represented the time allocated for students to independently work on the task within the JupyterLab environment. This phase was only present in tasks where explicit time was allotted for individual or group work.
- **(TP3)** Lecturer Filling the Code: This phase captured the lecturer's demonstration of the solution by populating the missing code snippets within the JupyterLab environment.
- **(TP4)** Lecturer Explaining the Code and Outputs: The final phase involved the lecturer providing a detailed explanation of the code, its underlying logic, and the resulting outputs or visualisations.

Note that not every learning task included all four phases due to variations in the lecturer's delivery style and the nature of the tasks themselves. However, TP1, TP3, and TP4, representing the core components of task introduction, solution demonstration, and explanation, were generally present in most tasks. TP2, reflecting dedicated time for student exploration, was only incorporated into a few tasks.

Table 2 reveals variations in student engagement and task structure across Weeks 2 and 3, reflecting the dynamic nature of a real-world teaching and learning setting. There was a strong correlation between the mean count of events and task duration, indicating that longer tasks naturally allowed for more student interactions. Variations also existed in the duration of each task phase (TP1-TP4) within a single task, with some tasks, like Tasks 3_5 and 3_6, building upon each other, leading to the absence of an introductory phase (TP1) for Task 3_6. Additionally, the presence and duration of dedicated exploration time (TP2) varied significantly, with some tasks lacking this phase altogether (indicated by a bold 0 in 'TP2 duration (seconds)' in Table 2). For instance, Week 3 started with two longer tasks, leading to time constraints later in the lecture and consequently reduced exploration time for subsequent tasks. This inconsistency in task structure, particularly the presence or absence of TP2, necessitated a nuanced analysis beyond that of simple metrics like total task duration or event counts, to provide a more effective understanding of the student behaviour. This study explored the extent to which the inclusion and duration of TP2 contributed to curiosity-driven learning.

Figure 2 illustrates the overall student activity patterns throughout the entire session, highlighting the peaks of engagement coinciding with the introduction and explanation of learning tasks.

Table 2: Summary of learning task event statistics and phase durations in Weeks 2 and 3 of the lectures. Not all learning tasks included all four phases. Dedicated Exploration Time (TP2) varied significantly, with some tasks lacking this phase altogether (indicated by a bold 0), particularly in Week 3 due to time constraints.

Task	2_1	2_2	2_3	2_4	2_5	2_6	2_7	2_8	3_1	3_2	3_3	3_4	3_5	3_6	3_7	3_8
no. of participants	33	33	33	33	33	33	33	33	34	34	34	34	34	34	34	34
mean count of events	56.70	28.18	13.67	23.09	49.24	50.48	11.42	18.79	76.91	52.38	17.65	26.65	50.47	7.91	14.62	39.32
std. count of events	37.10	23.25	8.10	20.04	18.89	38.95	9.55	14.80	41.91	42.82	14.35	13.96	30.91	6.45	9.46	34.38
median count of events	50	22	13	20	52	42	9	18	62.5	44.5	14.5	24.5	43.5	6.5	13	37
task duration (seconds)	602	278	125	205	473	735	129	208	738	577	190	261	595	98	251	356
TP1 duration (seconds)	76	88	25	155	56	125	48	163	70	112	32	112	101	0	94	88
TP2 duration (seconds)	113	56	31	**0**	160	65	**0**	**0**	107	54	**0**	**0**	**0**	**0**	**0**	**0**
TP3 duration (seconds)	228	82	57	19	163	511	37	0	372	14	63	84	19	25	63	9
TP4 duration (seconds)	185	52	12	31	94	34	44	45	189	397	95	65	475	73	94	259

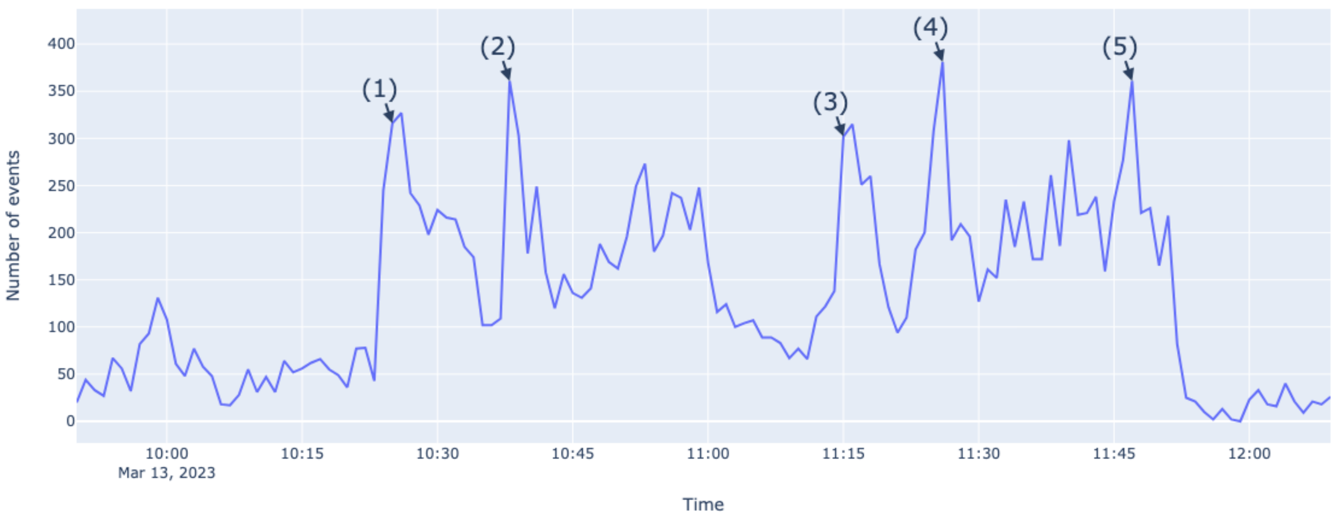

Figure 2: Overall student activity patterns in JupyterLab during Week 3's lecture. This line plot shows the total number of student interaction events recorded in JupyterLab over the duration of the learning session. The peaks of intensive activities closely aligned with working with actions in learning tasks. The five most prominent peaks in the plot correspond to the following: (1) Student started working on Task 1; (2) Student started working on Task 2; (3) Lecturer provided solution directly for Task 5; (4) Lecturer demonstrated a new function during Task 6; (5) Lecturer demonstrated a function during Task 8.

These peaks suggest that students tended to interact more actively with the learning material when presented with a defined challenge or a clear opportunity to confirm their understanding. This observation aligns with the concept of curiosity-driven exploration, where learners are driven by a desire to know and seek new information to make sense of their environment [19].

Figure 3 provides a closer look at individual student activity patterns during these two tasks, using the lecturer's actions (Figure 3a) as reference points. This detailed visualisation, facilitated by the trace data, allows for the identification of three distinct categories of student engagement:

- **Engagement-Oriented Behaviour**: These students were highly active within the JupyterLab environment, displaying a genuine interest in the learning process and actively seeking new information and experimenting with different approaches. They frequently employed a trial-and-error approach to problem-solving, readily viewing errors as opportunities for learning. They often went beyond the immediate task requirements, exploring additional resources and experimenting with the code. This pattern, characterised by active information-seeking and experimentation with code, aligns with the concept of curiosity-driven exploration, where learners are intrinsically motivated to seek a deeper understanding. Students in Figure 3b and Figure 3c exemplified this behaviour.
- **Answer-Seeking Behaviour**: In contrast to the highly engaged group, these students prioritised obtaining the correct answer over deep exploration or seeking additional information. While they demonstrated some activity within the JupyterLab environment, their primary goal was to achieve the desired outcome rather than engage deeply with the

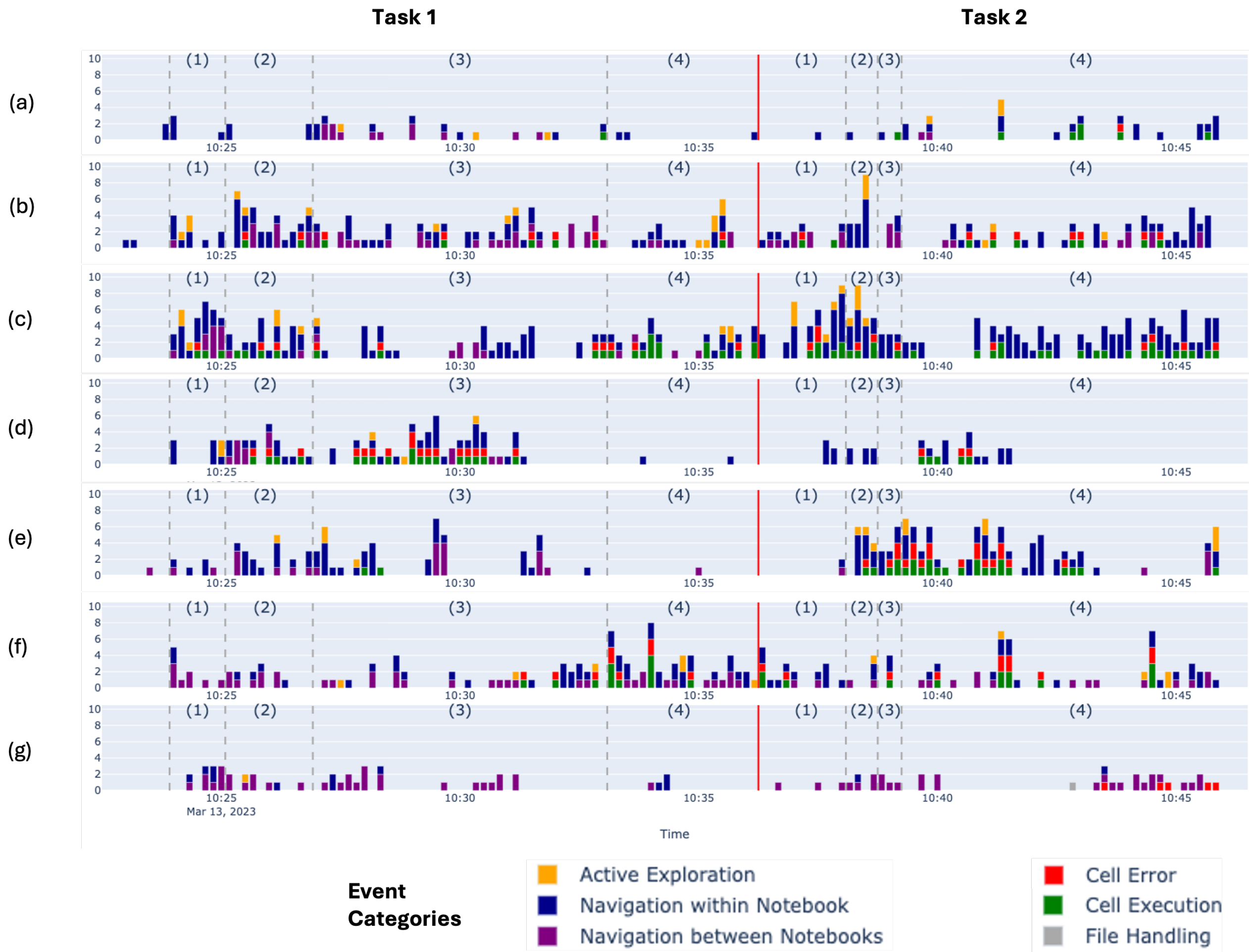

Figure 3: Individual activities from the lecturer (*a*) and sampled students (*b*-*g*) for the learning tasks. Below are the descriptions of time intervals for each learning task: (1) TP1: Lecturer introduced the task; (2) TP2: Students worked on the task; (3) TP3: Lecturer filled in the code; (4) TP4: Lecturer explained the code and outputs. The red vertical line represented the point at which the lecturer completed explaining Task 1 and began commenting on the context leading to Task 2.

learning process. Their exploration was limited to understanding the task requirements and achieving the desired outcome. This pattern aligns with a more goal-oriented approach, where learners focus on task completion. This category can be further subdivided into two distinct patterns: (1) Students who became inactive after successfully completing a task, illustrated by the student in Figure 3e, and (2) Students who primarily engaged in passive navigation of the learning material, becoming active only when encountering errors or when the correct answer was revealed. The student shown in Figure 3f demonstrated this pattern.

- **Disengagement-Oriented Behaviour**: Characterised by minimal participation and a lack of active engagement, this behaviour pattern was evident in students like the student in Figure 3g, whose activities were limited to navigating between notebooks. These students showed minimal interaction with the learning tasks and might not be actively seeking information or exploring the material.

The observed differences in student behaviours, particularly the presence of Engagement-Oriented Behaviour, highlighted potential avenues for investigating how specific task design elements, namely dedicated exploration time (TP2), influence these patterns. To uncover these subtle influences, a network analysis of all the 16 learning tasks was conducted, visualising the flow of events within different phases of the learning tasks (Figure 4).

As expected, 'Navigation within Notebooks' emerged as the most prevalent activity, since scrolling and moving between cells are fundamental actions associated with various activities. However, a perhaps more interesting observation was the notable increase between the events of 'Active Exploration' and 'Navigation within Notebook' (shown as red edges in the event flow of TP3 and TP4 in Figure 4) after the dedicated exploration time (TP2).

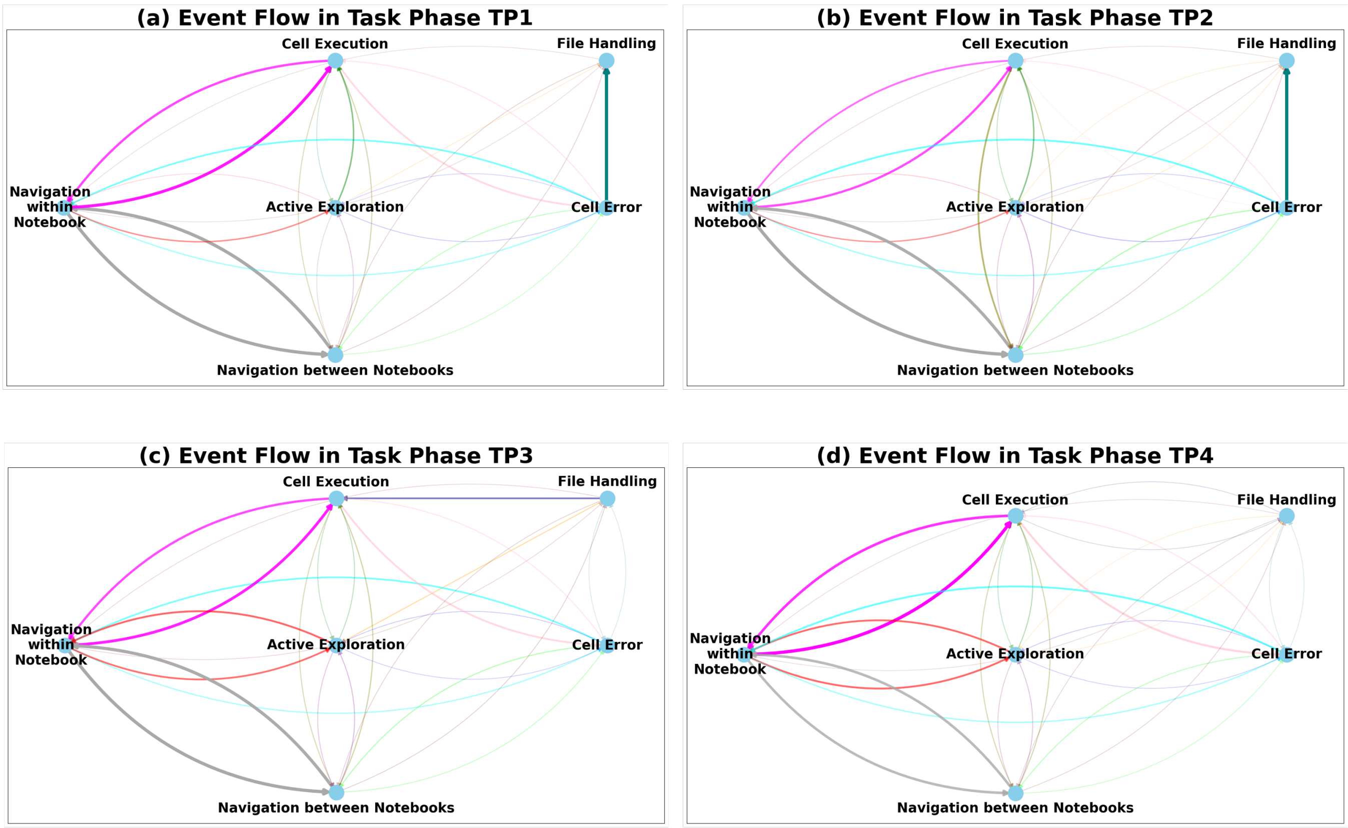

Figure 4: Network plots illustrate the event flow in different task phases. Nodes represent event categories, and edges indicate transitions between events. Edges are weighted based on the frequency of transitions between event categories, with the weight w_{ij} representing the number of times a transition occurs from event category i to event category j. These weights adjust the thickness of the edges, providing a visual representation of transition frequency, where thicker edges indicate more frequent transitions. The weight w_{ij} for an edge from node i to node j is calculated as $w_{ij} = \sum_{k=1}^{N-1} \delta(i, j, k)$, where N is the total number of events, and $\delta(i, j, k)$ is an indicator function that equals 1 if there is a transition from event i to event j at position k in the sequence, and 0 otherwise. Different colours are assigned to unique pairs of nodes to ensure clarity and visual distinction, helping to easily identify transitions between specific event categories. The transparency of the edges is standardised based on the weights, adding another layer of visual differentiation. The edge thickness t_{ij} is scaled from the weights using the normalisation $t_{ij} = \left(\frac{w_{ij} - w_{\min}}{w_{\max} - w_{\min}}\right) \times (t_{\max} - t_{\min}) + t_{\min}$, where $w_{\min}$ and $w_{\max}$ are the minimum and maximum weights, respectively, and $t_{\min}$ and $t_{\max}$ are the minimum and maximum thickness values used for scaling.

A noticeable change in event flow patterns was observed before and after the answer reveal, especially related to ‘Cell Error’ and ‘File Handling’ events. Before the answer was revealed, a one-way flow (shown as teal edges in the event flow of TP1 and TP2 in Figure 4) from ‘Cell Error’ to ‘File Handling’ indicated that students encountering errors were actively seeking solutions by opening and closing other notebooks, possibly referring to other learning materials or creating new notebooks for testing.

Another distinct unidirectional flow from ‘File Handling’ to ‘Active Exploration’ in TP3 (the phase immediately after the exploration time, shown as gold edges in the event flow of TP3 in Figure 4) suggests that some students might have sought answers from other notebooks and then copied and pasted them into their working notebooks.

5 Discussion

While this study did not directly measure curiosity — a multi-dimensional construct not easily captured through direct measurement — it explored student interactions with learning tasks in a postgraduate data analytics course using a task-centric learning analytics approach. By analysing trace data captured within the JupyterLab environment, we identified distinct behaviour patterns that provide insights into how different students engage with learning tasks in a technology-enhanced environment. Network analysis of these interactions also revealed patterns suggestive of knowledge-gap driven exploration. These observed behaviours, particularly those categorised as ‘Engagement-Oriented Behaviour’, serve as a foundation for understanding how specific actions within a learning environment relate to the knowledge-gap dimension of curiosity, and this represents an intial step towards the larger goal of modelling the complex phenomenon of curiosity.

Three key insights emerged from this analysis with implications valuable for both learning analytics research and pedagogical practice.

5.1 Insight 1: Providing sufficient exploration time before explicit instruction or answer revelation

Providing learners with sufficient exploration time before explicit instruction or answer revelation can encourage behaviours indicative of curiosity-driven learning. This insight is supported by observations from the network analysis, especially the notable increase in 'Active Exploration' and 'Navigation within Notebook' activities after the dedicated exploration time and the lecturer's explanation (Figure 4). This pattern suggests that students may benefit from additional time for independent exploration before the lecturer provided answers, as the allocated exploration time during lectures was often insufficient (as you can see from Table 2, the exploration time in TP2 varied). Further, the observation aligns with the study's findings on individual student activity patterns, where students exhibiting 'Engagement-Oriented Behaviour' demonstrated higher levels of activity and a greater tendency for experimentation when given autonomy in their learning process.

However, our analysis also revealed instances where students, potentially constrained by limited exploration time, resorted to answer-seeking behaviours, such as copying and pasting solutions from other notebooks. This behaviour, observed in the one-way flow from 'File Handling' to 'Active Exploration' in TP3 (the phase immediately after exploration time, see Figure 4c), indicates the importance of providing adequate time for exploration. the potential benefit of providing a longer exploration time to encourage genuine exploration and learning from mistakes.

To encourage curiosity-driven learning and deeper engagement with the material, educators could prioritise providing sufficient exploration time before explicit instruction or answer revelation. This can be achieved through some practical strategies. For example, design learning tasks that incorporate dedicated exploration phases and provide adequate time allocation for exploration. Insufficient exploration time can lead to answer-seeking behaviours like waiting for the answers from the lecturer. Educators could consider the time required for students to genuinely engage with the task and allocate sufficient time for independent exploration. They might also use real-time learning analytics dashboards to monitor student engagement during exploration time which would allow them to adjust the allocated time as needed. In addition, it could be helpful to provide scaffolding and support during exploration time. While independent exploration is crucial, it is also important to provide learners with appropriate scaffolding and support to ensure that the exploration is worthwhile. This may include providing access to relevant resources, offering prompts to guide exploration, and facilitating peer learning opportunities. The use of technology to provide real-time insights into exploratory behaviours could enable educators to offer timely interventions and support as needed.

5.2 Insight 2: Designing learning tasks and environments that embrace errors as learning opportunities

Designing learning tasks and environments that embrace errors as opportunities for learning can encourage deeper engagement and foster curiosity-driven exploration. This insight underscores that errors also play a role in prompting students to actively seek solutions and develop a more profound understanding of the learning material.

The network analysis revealed a clear pattern of error-driven exploration. Before the answer was revealed, a unidirectional flow from 'Cell Error' to 'File Handling' indicated that students encountering errors actively sought solutions. This behaviour, observed in TP1 and TP2 (before the answer reveal, see Figure 4a & 4b), involved actions like opening and closing notebooks, suggesting that students were referring to previous examples or creating new notebooks for testing. This pattern aligns with the concept of curiosity stemming from a desire to resolve uncertainty or knowledge gaps, as highlighted by Loewenstein [18], Noordewier and van Dijk [20]. Instead of viewing errors as setbacks, these students saw them as opportunities to investigate further and experiment with different approaches.

The trial-and-error approach employed by students exhibiting Engagement-Oriented Behaviour (Figure 3) further supports this notion. These students, characterised by their active exploration and experimentation, readily viewed errors as opportunities for learning. This suggests that incorporating elements of productive struggle and encouraging learners to persevere through challenges can be beneficial in fostering curiosity-driven learning. By providing a supportive environment where errors are embraced and seen as valuable learning experiences, educators can encourage students to embrace challenges and develop resilience in their learning journey.

To create such learning environments, we foreground some well-known actions reinforced by this study. First, design learning tasks that incorporate productive struggle. Instead of presenting straightforward problems, educators can design tasks that include manageable challenges. This encourages students to apply their knowledge in new contexts, experiment with different approaches, and develop problem-solving skills. Second, provide opportunities for experimentation and iteration. Encourage students to try different solutions, even if it leads to errors. Emphasise that the learning process often involves making mistakes and refining solutions based on those errors. Third, offer constructive feedback that focuses on the learning process. When students encounter errors, provide feedback that not only addresses the error itself but also guides them in understanding the underlying concepts and encourages them to reflect on their approach. Last, facilitate peer learning and collaboration. Encourage students to share their thought processes, discuss their errors, and learn from each other's mistakes. This can create a more supportive learning environment where errors are embraced.

5.3 Insight 3: Recognising and encouraging different engagement patterns promotes learners' understanding

Not all learners engage with learning tasks in the same way, indicating the need for educators to avoid 'one-size-fits-all' teaching methods where possible. Some learners are driven by a desire for exploration and experimentation (Engagement-Oriented Behaviour), while others prioritise task completion and seeking correct answers (Answer-Seeking Behaviour). Further, some learners may exhibit patterns of disengagement, denoted by minimal interaction with the learning material. Although Answer-Seeking Behaviour can lead to successful task completion, it might also result in a more superficial understanding of learning material compared to Engagement-Oriented Behaviour. Active exploration and experimentation, characteristic of Engagement-Oriented Behaviour, likely leads to a deeper understanding of the course material, aligning with findings from Oudeyer et al. [22], Shin and Kim [29].

Based on our analysis, we advocate for a more inclusive learning environment that accommodates various learning approaches to encourage deeper understanding. Educators and learning designers can achieve this by implementing diverse learning activities and assessments that cater to a range of learning preferences. This may include providing opportunities for exploration and focused problem-solving, such as offering optional challenges for students who lean towards Engagement-Oriented Behaviour, while ensuring that structured activities reinforce core concepts for those who prefer Answer-Seeking Behaviour.

5.4 Implications for Learning Analytics

We presented several implications for the learning analytics community, highlighting the potential of data analysis and visualisation to understand and support curiosity-driven exploration. Findings from our analysis underscore the need for learning analytics systems that go beyond measuring traditional learning outcomes, such as performance and completion rates. Instead, these systems should focus on capturing and analysing more complex learning processes like curiosity-driven exploration. By shifting the focus from product-oriented metrics to process-oriented insights, learning analytics can provide a more comprehensive and nuanced view of student learning.

In this study, we demonstrated that trace data offers a valuable source of information for understanding exploratory behaviours. By analysing patterns in these data, learning analytics systems can identify behaviours indicative of curiosity, such as active exploration, navigation patterns, and responses to encountering errors. These insights can then be visualised in dashboards or other user-friendly formats, providing educators with a real-time understanding of how students are engaging with the learning material. From an educator's perspective, approaches that allow for real-time visualisation and understanding of student data are needed. The ability to identify students struggling with specific concepts or exhibiting patterns of disengagement allows for timely interventions that can help keep learners on track and encourage deeper engagement with the material [13, 27]. Although some learning analytics systems, such as the monitoring tools for synchronous online activities, are designed to be real-time, most analyses are conducted retrospectively [27]. This retrospective approach often results in reactive interventions, occurring only after a student has already disengaged or fallen behind [31].

Our findings suggest the need for learning analytics systems that can identify and prioritise knowledge gaps, potentially piquing curiosity and motivating learners to explore more actively. This aligns with the task-centric learning analytics approach [14], which emphasises aligning learning analytics tools with specific learning tasks. By integrating learning analytics systems with task designs, educators can gain a more contextualised understanding of student behaviours and tailor interventions accordingly. For example, real-time feedback on student exploration patterns during dedicated exploration phases (TP2) can help educators adjust the allocated time, provide additional scaffolding, or offer personalised guidance to foster curiosity-driven learning.

The prevalence of 'Answer-Seeking Behaviour' observed in the study presents a challenge for educators aiming to foster curiosity-driven learning. This behaviour pattern, characterised by a focus on obtaining the correct answer rather than engaging deeply with the learning process, suggests that some students might prioritise task completion over developing a comprehensive understanding of the material. Addressing this challenge requires a shift in pedagogical approaches, moving away from simply providing answers to creating learning experiences that encourage active problem-solving, critical thinking, and a genuine desire to learn. Implementing strategies such as collaborative learning activities, designing tasks that require application of knowledge rather than rote memorisation, and clearly communicating the value of exploration and learning from mistakes can help cultivate a more curiosity-driven learning environment. Kitto and Gibson [13] suggested that a shift is needed in learning analytics from solely focusing on optimising learning to understanding and improving the complex socio-technical system in which it operates. There is a need for educational systems that provide real-time insights into exploratory behaviours, allowing educators to facilitate a more supportive and responsive learning environment that encourages curiosity and deeper engagement.

5.5 Future Directions

Future research is needed to explore more diverse data sources addressing the limitations of analysing curiosity based solely on categorised behaviours. While we demonstrated distinct behavioural patterns from categorising student interaction events within the JupyterLab environment, it is necessary to recognise that these categories, while providing a useful framework for analysis, might not fully capture the complex and multifaceted nature of curiosity. For example, a student repeatedly copying and pasting code from LLMs might generate a high volume of 'Active Exploration' events without genuinely engaging in curiosity-driven learning. Researchers could administer pre- and post-task questionnaires to assess changes in curiosity levels or incorporate self-reported curiosity measures during task engagement. Additionally, gathering self-reported data about students' confidence levels could provide further insights into their experiences and help triangulate findings from trace data. Combining these measures with behavioural data would provide a more comprehensive understanding of the relationship between curiosity and exploratory behaviours in technology-enhanced learning environments.

Future research could also explore the impact of external factors, such as students' prior knowledge, learning preferences, or the learning environment, on the relationship between curiosity and exploratory behaviours. This study focused on a specific postgraduate data analytics course, so further research is needed to explore the generalisability of these findings to other learning contexts and disciplines. Investigating whether similar patterns of exploratory behaviour are observed in different academic subjects, age groups, and learning environments would strengthen the study's external validity and provide valuable insights for a broader range of educators and instructional designers.

6 Towards Modelling of Curiosity

This research is a small step towards modelling curiosity. It is important to acknowledge that curiosity is a multidimensional concept that cannot be completely defined by observable behaviours. This paper represented an initial step towards understanding and modelling curiosity by focusing on analysing students' exploratory behaviours in response to learning tasks, primarily addressing the knowledge gap dimension of curiosity.

The next step involves incorporating additional sources of information to address the experiential dimension of curiosity and the relationship between the dimensions. This will provide a more comprehensive understanding of the interplay of curiosity dimensions, including epistemic (knowledge gap) and experiential dimensions, that contribute to an individual's curiosity. Such a model could incorporate elements like the learner's prior knowledge and experiences, their emotional state and motivation levels, the social dynamics of the learning environment and the specific design of the learning tasks and technologies. By examining how these factors interact and influence curiosity-driven behaviours, researchers could gain a deeper understanding of how to effectively foster and sustain curiosity in educational settings. This holistic approach to modelling curiosity will be valuable to the field of learning analytics.

References

[1] Rania Abdelghani, Pierre-Yves Oudeyer, Edith Law, Catherine de Vulpillières, and Hélène Sauzéon. 2022. Conversational agents for fostering curiosity-driven learning in children. *International Journal of Human Computer Studies* 167 (2022). https://doi.org/10.1016/j.ijhcs.2022.102887

[2] Abhinava Barthakur, Jelena Jovanovic, Andrew Zamecnik, Vitomir Kovanovic, Gongjun Xu, and Shane Dawson. 2024. Towards Comprehensive Monitoring of Graduate Attribute Development: A Learning Analytics Approach in Higher Education. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 78–89. https://doi.org/10.1145/3636555.3636852

[3] D. E. Berlyne. 1966. Curiosity and exploration. *Science* 153, 3731 (July 1966), 25–33. https://doi.org/10.1126/science.153.3731.25

[4] Conrad Borchers, Jiayi Zhang, Ryan S. Baker, and Vincent Aleven. 2024. Using Think-Aloud Data to Understand Relations between Self-Regulation Cycle Characteristics and Student Performance in Intelligent Tutoring Systems. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 529–539. https://doi.org/10.1145/3636555.3636911

[5] Blazenka Divjak, Barbi Svetec, and Damir Horvat. 2023. Learning analytics dashboards: What do students actually ask for?. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK2023)*. Association for Computing Machinery, New York, NY, USA, 44–56. https://doi.org/10.1145/3576050.3576141

[6] Dominik Glandorf, Hye Rin Lee, Gabe Avakian Orona, Marina Pumptow, Renzhe Yu, and Christian Fischer. 2024. Temporal and Between-Group Variability in College Dropout Prediction. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 486–497. https://doi.org/10.1145/3636555.3636906

[7] Emily M. Grossnickle. 2016. Disentangling curiosity: Dimensionality, definitions, and distinctions from interest in educational contexts. *Educational Psychology Review* 28, 1 (March 2016), 23–60. https://doi.org/10.1007/s10648-014-9294-y

[8] Jamie Jirout and David Klahr. 2012. Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review* 32, 2 (2012), 125–160.

[9] Jamie J. Jirout and Shoronda E. Matthews. 2022. Developing Intellectual Character: An Educational Perspective on How Uncertainty-Driven Curiosity Can Support Learning. In *Uncertainty: A Catalyst for Creativity, Learning and Development*, Ronald A. Beghetto and Garrett J. Jaeger (Eds.). Springer International Publishing, Cham, 253–268. https://doi.org/10.1007/978-3-030-98729-9_14

[10] Rogers Kaliisa, Kamila Misiejuk, Sonsoles López-Pernas, Mohammad Khalil, and Mohammed Saqr. 2024. Have Learning Analytics Dashboards Lived Up to the Hype? A Systematic Review of Impact on Students' Achievement, Motivation, Participation and Attitude. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 295–304. https://doi.org/10.1145/3636555.3636884

[11] Min Jeong Kang, Ming Hsu, Ian M. Krajbich, George Loewenstein, Samuel M. McClure, Joseph Tao-yi Wang, and Colin F. Camerer. 2009. The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science* 20, 8 (Aug. 2009), 963–973. https://doi.org/10.1111/j.1467-9280.2009.02402.x

[12] Kara Kedrick, Paul Schrater, and Wilma Koutstaal. 2023. The Multifaceted Role of Self-Generated Question Asking in Curiosity-Driven Learning. *Cognitive Science* 47, 4 (2023), e13253. https://doi.org/10.1111/cogs.13253 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.13253.

[13] Kirsty Kitto and Andrew Gibson. 2024. Places to intervene in complex learning systems. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 929–935. https://doi.org/10.1145/3636555.3636938

[14] Simon Knight, Andrew Gibson, and Antonette Shibani. 2020. Implementing learning analytics for learning impact: Taking tools to task. *The Internet and Higher Education* 45 (April 2020), 100729. https://doi.org/10.1016/j.iheduc.2020.100729

[15] Wilma Koutstaal, Kara Kedrick, and Joshua Gonzalez-Brito. 2022. Capturing, clarifying, and consolidating the curiosity-creativity connection. *Scientific Reports* 12, 1 (Sept. 2022), 15300. https://doi.org/10.1038/s41598-022-19694-4 Number: 1 Publisher: Nature Publishing Group.

[16] Tongguang Li, Yizhou Fan, Namrata Srivastava, Zijie Zeng, Xinyu Li, Hassan Khosravi, Yi-Shan Tsai, Zachari Swiecki, and Dragan Gašević. 2024. Analytics of Planning Behaviours in Self-Regulated Learning: Links with Strategy Use and Prior Knowledge. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 438–449. https://doi.org/10.1145/3636555.3636900

[17] Jordan A. Litman. 2005. Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition and Emotion* 19, 6 (Sept. 2005), 793–814. https://doi.org/10.1080/02699930541000101

[18] George Loewenstein. 1994. The psychology of curiosity: A review and reinterpretation. *Psychological bulletin* 116, 1 (1994), 75.

[19] Kou Murayama. 2022. A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic–extrinsic rewards. *Psychological Review* 129, 1 (Jan. 2022), 175–198. https://doi.org/10.1037/rev0000349

[20] Marret K Noordewier and Eric van Dijk. 2020. Deprivation and discovery motives determine how it feels to be curious. *Current Opinion in Behavioral Sciences* 35 (Oct. 2020), 71–76. https://doi.org/10.1016/j.cobeha.2020.07.017

[21] Ikenna Osakwe, Guanliang Chen, Yizhou Fan, Mladen Rakovic, Shaveen Singh, Inge Molenaar, and Dragan Gašević. 2024. Measurement of Self-regulated Learning: Strategies for mapping trace data to learning processes and downstream analysis implications. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 563–575. https://doi.org/10.1145/3636555.3636915

[22] Pierre-Yves Oudeyer, Jacqueline Gottlieb, and Manuel Lopes. 2016. Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. In *Progress in Brain Research*, Bettina Studer and Stefan Knecht (Eds.). Motivation, Vol. 229. Elsevier, 257–284. https://doi.org/10.1016/bs.pbr.2016.05.005

[23] Jae H. Paik, Igor Himelfarb, Seung Hee Yoo, KyoungMi Yoo, and Hoyong Ha. 2024. The relationships among school engagement, students' emotions, and academic performance in an elementary online learning. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 219–230. https://doi.org/10.1145/3636555.3636871

[24] Zilong Pan and Min Liu. 2022. The effects of learning analytics hint system in supporting students problem-solving. In *LAK22: 12th International Learning Analytics and Knowledge Conference (LAK22)*. Association for Computing Machinery, New York, NY, USA, 77–86. https://doi.org/10.1145/3506860.3506871

[25] Ernesto Panadero. 2017. A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology* 8 (April 2017). https://doi.org/10.3389/fpsyg.2017.00422 Publisher: Frontiers.

[26] Francesco Poli, Jill X. O'Reilly, Rogier B. Mars, and Sabine Hunnius. 2024. Curiosity and the dynamics of optimal exploration. *Trends in Cognitive Sciences* 0, 0 (Feb. 2024). https://doi.org/10.1016/j.tics.2024.02.001 Publisher: Elsevier.

[27] Stanislav Pozdniakov, Roberto Martinez-Maldonado, Yi-Shan Tsai, Mutlu Cukurova, Tom Bartindale, Peter Chen, Harrison Marshall, Dan Richardson, and Dragan Gasevic. 2022. The Question-driven Dashboard: How Can We Design Analytics Interfaces Aligned to Teachers' Inquiry?. In *LAK22: 12th International Learning Analytics and Knowledge Conference (LAK22)*. Association for Computing Machinery, New York, NY, USA, 175–185. https://doi.org/10.1145/3506860.3506885

[28] Clara Schumacher and Dirk Ifenthaler. 2018. The importance of students' motivational dispositions for designing learning analytics. *Journal of Computing in Higher Education* 30, 3 (Dec. 2018), 599–619. https://doi.org/10.1007/s12528-018-9188-y

[29] Dajung Diane Shin and Sung-il Kim. 2019. Homo curious: Curious or interested? *Educational Psychology Review* 31, 4 (Dec. 2019), 853–874. https://doi.org/10.1007/s10648-019-09497-x

[30] Namrata Srivastava, Yizhou Fan, Mladen Rakovic, Shaveen Singh, Jelena Jovanovic, Joep van der Graaf, Lyn Lim, Surya Surendrannair, Jonathan Kilgour, Inge Molenaar, Maria Bannert, Johanna Moore, and Dragan Gasevic. 2022. Effects of Internal and External Conditions on Strategies of Self-regulated Learning: A Learning Analytics Study. In *LAK22: 12th International Learning Analytics and Knowledge Conference (LAK22)*. Association for Computing Machinery, New York, NY, USA, 392–403. https://doi.org/10.1145/3506860.3506972

[31] Kirk Vanacore, Ashish Gurung, Adam Sales, and Neil T. Heffernan. 2024. The Effect of Assistance on Gamers: Assessing The Impact of On-Demand Hints & Feedback Availability on Learning for Students Who Game the System. In *Proceedings of the 14th Learning Analytics and Knowledge Conference (LAK '24)*. Association for Computing Machinery, New York, NY, USA, 462–472. https://doi.org/10.1145/3636555.3636904

[32] Tom Zhang, Michelle Taub, and Zhongzhou Chen. 2022. A Multi-Level Trace Clustering Analysis Scheme for Measuring Students' Self-Regulated Learning Behavior in a Mastery-Based Online Learning Environment. In *LAK22: 12th International Learning Analytics and Knowledge Conference (LAK22)*. Association for Computing Machinery, New York, NY, USA, 197–207. https://doi.org/10.1145/3506860.3506887