

Research Article

Sam O'Neill*, David Mulgrew, Ovidiu Bagdasar

On the Use of Large Language Models for Improving Student and Staff Experience in Higher Education

<https://doi.org/10.1515/edu-2025-0086>

received December 24, 2024; accepted May 01, 2025

Abstract: Large language models (LLMs) hold great promise for enhancing teaching and learning in higher education, yet educators and administrators still lack practical examples to guide their adoption. This article presents insights and use cases from the integration of LLMs into a first-year undergraduate computer science cohort. By employing LLMs as digital scaffolds, timely support was provided helping students bridge knowledge gaps while engaging in independent problem-solving. At the same time, students were encouraged to maintain a critical stance by evaluating and verifying AI-generated content. These initial observations show that LLMs can encourage self-guided research, offer on-demand feedback, and strengthen cohort identity by acting as a mentor, peer, and liaison. Although the findings are exploratory, they serve as a point of reference for educators, informing future, more rigorous studies aimed at the successful integration of LLMs into higher education settings.

Keywords: LLMs, ChatGPT, learning technology, Chatbot, student experience, staff experience

1 Introduction

The rapid advancement of large language models (LLMs) like ChatGPT has ushered in a new era of human-computer interaction, transforming how we access and contextualise information (Bommasani et al., 2022). As these

models continue to evolve and improve, educators are exploring innovative ways to harness their power in teaching and learning environments (Futterer et al., 2023; Memarian & Doleck, 2023) while coping with the challenges these models pose (Kasneci et al., 2023; Milano, McGrane, & Leonelli, 2023). The potential of LLMs in education has been a topic of growing interest, with researchers investigating their applications in various educational contexts (Jeon & Lee, 2023; Mollick & Mollick, 2023). Recent work also emphasises the need for policymakers and educators to critically engage with generative AI, considering its broader implications for educational policy, instructional practices, and the validation of knowledge (Miao & Holmes, 2023).

This study aims to explore the integration of LLMs in higher education to enhance both student and staff experience. By focusing on a first-year computer science cohort, we investigate how LLMs can support lecture planning, assessment design, and student engagement. The use cases in this article illustrate how LLMs can facilitate content co-creation, prompt critical thinking, and foster a sense of cohort identity.

In all applications considered, students were actively encouraged to use these tools for information retrieval, problem-solving, and self-guided research while made aware of the potential for hallucinations (Alkaissi & McFarlane, 2023). By embracing these cutting-edge technologies, educators can unlock new avenues for active learning, personalised guidance, and collaborative knowledge acquisition (Zawacki-Richter, Marín, Bond, & Gouverneur, 2019).

In addition to empowering students to interact with LLMs, the authors of this study conducted demonstrations to identify effective ways of model usage. Furthermore, LLMs were employed to co-create a range of instructional materials, including multiple-choice quizzes, tutorial problems, and lecture notes/slides, which were then converted to PDFs (Sharma, Shailendra, & Kadel, 2025).

To the surprise of the educators, students quickly developed their own Discord bot (nicknamed DerbyGPT) utilising the ChatGPT API, serving as an academic mentor and helping build a strong cohort identity by providing

* **Corresponding author: Sam O'Neill**, School of Computing, University of Derby, Derby, United Kingdom, e-mail: s.oneill@derby.ac.uk

David Mulgrew: School of Computing, University of Derby, Derby, United Kingdom, e-mail: d.mulgrew1@unimail.derby.ac.uk

Ovidiu Bagdasar: School of Computing, University of Derby, Derby, United Kingdom; Department of Mathematics, Faculty of Exact Sciences, "1 Decembrie 1918" University of Alba Iulia, Alba Iulia, Romania, e-mail: o.bagdasar@derby.ac.uk, ovidiu.bagdasar@uab.ro

encouragement, information, and guidance. The use of chatbots in education has been explored in recent studies, highlighting their potential to support personalised learning and enhance student engagement (Wollny et al., 2021; Berrezueta-Guzman, Parmacli, Krusche, & Wagner, 2024).

By highlighting initial insights, this work serves as a formative examination rather than a definitive guide. Educators and researchers may use these examples to inform their own experimentation, to anticipate common barriers, and to craft more targeted research questions for future, more rigorous investigations. As LLM capabilities continue to mature and their roles in education take shape, early-stage explorations like this one can help build the foundation for more evidence-based, strategic integrations of AI into higher education teaching and learning (Bobula, 2024).

This article is set out as follows. In Section 2, theoretical lenses with which to view these use cases are presented. Section 3 provides the methodology of our work. Section 4 outlines how LLMs have been used to help academics through collaboration to improve efficiency. Section 5 explores student use cases and how they are improving their learning experience. Section 6 details a student-led project that has helped build a strong cohort identity through deployment within their Discord server. Section 7 presents the results, including thematic analysis of observations and educator reflections. Finally, a summary and future work are given in Section 8.

2 Conceptual Framework

The design and interpretation of our study are informed by two complementary theoretical lenses: scaffolding within a sociocultural constructivist paradigm and critical digital literacy.

2.1 Scaffolding and Sociocultural Constructivism

In sociocultural constructivism, learning is understood as a process where individuals acquire new knowledge and skills through guided interactions, both with more knowledgeable others and with tools that mediate learning (Vygotsky, 1978; Palincsar, 1998). Key to this paradigm is scaffolding, where learners receive just enough guidance or support to tackle tasks slightly beyond their current ability level (Wood, Bruner, & Ross, 1976). For instance, technology as a support within the universal design for

learning framework can function as temporary scaffolds, gradually removed as learners gain proficiency, or as permanent components that transform traditional literacy and learning practices (Vasinda & Pilgrim, 2023). In formal educational settings, scaffolding often comes from instructors or peers; however, emerging research suggests that AI-driven tools – such as LLMs – can also provide adaptive forms of assistance (Chien, Chan, & Hou, 2024; Liao et al., 2024).

In our study, LLMs served as a near-instant resource that learners could consult to bridge conceptual gaps, verify partial knowledge, and extend their problem-solving capabilities. This configuration positions the LLM as a “digital scaffold,” offering timely hints, clarifications, and step-by-step explanations that reduce cognitive load and help students progress more confidently towards task completion.

2.2 Critical Digital Literacy

While sociocultural constructivism captures how learners build knowledge with supportive tools, it does not fully address the evaluative dimension that arises when dealing with AI-generated content. For that, we draw on critical digital literacy, which emphasises the ability to question, interpret, and verify digital texts, media, and platforms (Hinrichsen & Coombs, 2013). In an age where LLMs can produce both insightful responses and so-called “hallucinations,” students must practice recognising fact from misinformation (Ciampa, Wolfe, & Bronstein, 2023; Naamati-Schneider & Alt, 2024). The growth of synthetic information created by generative AI further highlights the importance of integrating critical AI literacy into educational practice, equipping learners to understand and navigate the increasingly plastic nature of digital information (Roe, Furze, & Perkins, 2025). Developing these competencies ensures that students become not only effective users of AI tools but also critical consumers of AI-generated content.

Our instructional approach actively encouraged learners to scrutinise AI outputs by cross-referencing suggested information, carefully checking code solutions for errors, and reflecting on the validity of advice gleaned from the LLM. In doing so, students were guided to become critical consumers of digital content. Rather than taking the LLM's outputs at face value, they were taught to pause, verify, and adapt or discard suggestions as needed – an essential skill in a digital environment rich with both credible and spurious information.

2.3 Synthesis

By combining these two lenses – scaffolding through AI and critical digital literacy – the conceptual framework addresses both the supportive and the evaluative dimensions of LLM usage. The scaffolding perspective explains how students can grow academically when provided with just-in-time guidance that extends their ability to solve problems independently. The critical digital literacy perspective, meanwhile, underscores the importance of approaching AI outputs with healthy scepticism.

3 Methodology

3.1 Participant Details

The study was conducted over a period of two academic years with a cohort of first-year Computer Science students at the University of Derby. The students, primarily aged between 18 and 20, represent a diverse sample in terms of academic level and familiarity with digital tools. Data were collected from 78 students who voluntarily consented to participate, in accordance with ethical guidelines approved by the College Ethics Committee.

3.2 Time Frame

The integration of LLMs into the curriculum began in July 2023 and the data collection period spanned the Fall 2023 and 2024 semesters.

3.3 Data Collection

Quantitative data were collected using a structured survey instrument comprising five key questions. These questions were designed to assess student frequency of AI tool usage, openness to increased AI integration in teaching, opinions on banning AI in assessments, perceived relevance of AI skills for future careers, and overall satisfaction with AI support. Responses were measured on a five-point Likert scale. While the brevity of the survey instrument allowed for rapid data collection, it limited the scope of quantitative insights.

In parallel, educators maintained observations during lectures and tutorials. These observations provided qualitative

context to supplement the survey responses, focusing on how LLMs were employed in tasks such as quiz creation, lecture adaptation, and student support.

3.4 Data Analysis

The survey responses were analysed by descriptive statistics and qualitative data from observations and educator reflections were analysed thematically. The combined approach ensured that the data from the survey were contextualised by qualitative insights.

3.5 Ethical Considerations

The survey was conducted anonymously, and all participants were informed that no personal data would be collected but that responses would be used for the purposes of research. Based on this information, participants were asked to consent to take part in the survey. Ethics approval was also granted by the College Ethics Committee at the University of Derby.

3.6 Limitations

3.6.1 Sample Size

The study focused on 78 first-year computer science students at one UK institution, which limits the transferability of findings to other universities, disciplines, or cultural contexts. Future research should include multiinstitutional or cross-cultural samples for broader external validity.

3.6.2 Lack of a Control Group

The exploratory nature of this study meant that there was no control group or randomised assignment. As a result, we cannot definitively attribute observed outcomes (e.g. improved student engagement) solely to the LLM interventions.

3.6.3 Reliance on Self-Reported Data

While our survey data offer insights into student perceptions and frequency of use, self-reported measures are

susceptible to recall inaccuracies and social desirability bias. Future work could incorporate objective metrics – like automated logs of user activity – to triangulate self-reported data.

3.6.4 Breadth of Survey Questions

The limited number of survey questions represents a constraint in capturing the full breadth of student perspectives on AI in education. Future research should expand the survey instrument to include additional quantitative items and open-ended questions.

3.6.5 Variability in LLM Behaviour

The quality of responses from the LLMs used can shift over time due to ongoing model updates or changes in training data. Although students were guided on verifying AI-generated content, we were unable to fully control for the unpredictability of these tools.

4 Collaboration as Creation: Utilising LLMs for Efficient Development of Instructional Material

In the current LLM era, the process of creating instructional materials is undergoing a significant transformation. These models offer opportunities for collaboration and efficient generation of relevant content, saving time and taking a dynamic and interactive approach to curriculum development.

From a sociocultural constructivist viewpoint, co-creating instructional materials with LLMs can be understood as a form of scaffolding for the educators themselves. Rather than manually generating every quiz question or rubric criterion, instructors rely on the model's rapid drafting capabilities, thereby freeing mental bandwidth for more nuanced pedagogical tasks, such as adapting materials to student needs or providing clarifications and support in real time. Concurrently, critical digital literacy considerations come into play whenever instructors evaluate and refine the model's outputs. By verifying each automatically generated prompt or rubric descriptor, educators exemplify the critical stance needed to avoid passively adopting AI-generated materials. This iterative “human-in-the-

loop” approach helps ensure alignment with academic standards and fosters a culture of reflective practice.

Four examples are detailed in this section. It is important to note that in all cases the human educators remained in the loop, to provide oversight, guidance, and quality control. The LLMs acted as intelligent co-creators, generating initial drafts and content, which were then iteratively refined and tailored to meet the specific pedagogical needs and standards of the course.

4.1 Rapid Quiz Creation

LLMs played a pivotal role in the rapid creation of multiple-choice quizzes for both formative and summative assessments. Educators then used these models to generate large pools of potential questions and answers, tailored to specific topics and learning objectives. By providing clear prompts and specifying the desired level of difficulty, educators were able to produce questions that catered for a variety of skill levels, from foundational knowledge checks to more advanced critical-thinking tasks.

After the initial generation, the quizzes were reviewed and refined by educators to ensure accuracy, relevance, and alignment with the course's learning outcomes. This iterative process allowed for the curation of high-quality questions that not only assessed students' knowledge but also reinforced key concepts.

To streamline integration into the virtual learning environment (VLE), the LLM was specifically instructed to output questions in a tab-delimited format. This formatting enabled educators to seamlessly import the generated quizzes into the VLE, significantly reducing the time and effort required for manual input.

This approach not only saved time in the creation and implementation of assessments but also allowed educators to focus on refining quiz quality and ensuring they provided meaningful feedback to students. The use of LLMs in this context demonstrated their potential to enhance the efficiency of assessment design and implementation, while maintaining pedagogical rigor.

4.2 Dynamic Lecture Adaptation

In one instance, the educator's reflections on potential issues with the initially proposed lecture prompted a pivot. In collaboration with an LLM, a new lecture was efficiently created, allowing for a more responsive and adaptive teaching approach. Students were made aware of the co-creation of the lecture and shown how the lecture content

had been created for transparency. Through centralised feedback mechanisms, students praised the approach and noted their appreciation for the decision.

4.3 Rubric Generation

Developing clear and consistent rubrics can be a time-intensive task, requiring alignment with learning objectives, precision, and fairness. By engaging with LLMs such as ChatGPT, educators were able to rapidly generate initial rubric drafts, which were then refined through iterative feedback and revisions. Tutors provided key learning outcomes and assessment goals, prompting the LLM to generate performance criteria and descriptors for various achievement levels. While the initial drafts required adjustments for clarity, alignment, and appropriateness, they served as an effective starting point.

This collaborative approach significantly reduced the time spent on drafting rubrics while maintaining high-quality, tailored results. The tutors ensured the final rubrics were comprehensive, transparent, and student-friendly, helping the students to understand grading criteria and expectations.

Figure 1 demonstrates an example of a draft rubric created through this iterative process. Through the use

LLMs educators can effectively complement human expertise and help reduce workload while maintaining rigour and student focus on assessment design.

4.4 Coding Exercises and Explanations

In the introductory programming module, LLMs were used to generate diverse coding exercises, complete with commented solutions and step-by-step explanations. These exercises ranged from basic syntax tasks to more complex algorithmic problems, enabling educators to create a wide array of practice materials efficiently.

By reviewing and refining the model's outputs, instructors ensured the exercises aligned with module objectives and maintained clarity and accuracy. This collaborative approach provided students with varied opportunities for active learning, helping them engage with programming concepts and develop problem-solving skills. The detailed solutions and explanations supported independent learning and reduced reliance on instructor feedback for routine queries.

The integration of LLMs streamlined the creation of high-quality teaching materials, enriching the learning experience and freeing up educator time for more personalised support and complex topics.

Grade Band	Criteria
80%+	The database schema is fully normalised (at least 3NF) with well-defined primary and foreign keys . All tables accurately reflect the ERD , including correctly implemented relationships (one-to-many, many-to-many). Constraints such as referential integrity and appropriate data types are correctly applied.
70 – 79%	The database schema is well-structured, mostly normalised (some minor redundancy), and all key relationships are correctly implemented . Minor issues in data types or constraints but nothing critical.
60 – 69%	The database contains mostly correct tables and relationships , but some normalisation issues exist (e.g., repeated fields or unnecessary redundancy). Most constraints are applied , but there may be missing referential integrity in some cases.
50 – 59%	Some key relationships are missing or incorrectly implemented (e.g., incorrect use of foreign keys). Normalisation is incomplete (e.g., data duplication). Constraints are present but inconsistently applied.
40 – 49%	The database schema is incomplete or has major flaws (e.g., missing tables, incorrect primary keys, lack of normalisation). Relationships may not be properly established.
35 – 39%	A clear attempt has been made, but significant structural errors (e.g., missing relationships, incorrect primary/foreign keys) make the database difficult to use.
1 – 34%	The database does not follow relational design principles or lacks meaningful implementation.

Figure 1: A draft rubric generated iteratively with ChatGPT 3.5.

4.5 Summary

These applications demonstrate how LLMs function as digital scaffolds. By quickly producing rough drafts of rubrics or coding exercises, the AI reduces time-intensive tasks and allows educators to focus on mentoring students and customising learning materials. In turn, educators exercise critical digital literacy by reviewing, modifying, and validating the generated content to ensure alignment with specific learning outcomes. This dynamic interplay highlights how LLMs can streamline content creation without compromising academic rigor, so long as educators remain actively involved in refining the AI's outputs.

5 The Academic in Your Pocket: Empowering Students with LLMs

The integration of LLMs in first-year undergraduate education has presented new avenues for personalised learning and academic support. By actively encouraging students to interact with LLMs during tutorials and coursework, the hope is to provide a more engaging and dynamic learning environment.

Encouraging students to use LLMs in this manner resonates with sociocultural constructivist principles, wherein timely support is viewed as a scaffolding mechanism that bridges learners' current abilities and the tasks' demands. By consulting an LLM for clarifications, students effectively receive just-in-time guidance – akin to having a knowledgeable peer or teaching assistant on call. In parallel, fostering critical digital literacy remains essential. Learners are consistently reminded to verify the AI's responses, cross-reference additional sources, and reflect on the credibility of the information provided. Through this process, they sharpen their evaluative skills and develop healthy scepticism about AI-generated outputs, a crucial competency in today's digital age.

This section explores the various ways in which students were encouraged to use LLMs as their personal academic guides, mentors, and critical reviewers.

5.1 Guided Learning and Fact-Checking

During tutorials, the students were introduced to the concept of LLMs as tools to guide their learning and exploration of specific topic areas. This approach was intended to demonstrate how LLMs can complement traditional educational resources, offering a dynamic and interactive means

of engaging with academic content. To maximise the effectiveness of these interactions, educators conducted demonstrations to show students how to formulate clear and concise queries, and to interpret the responses generated by the models.

A key focus of these tutorials was to raise awareness of the limitations of LLMs, particularly the potential for generating hallucinated or inaccurate information. Students were provided with strategies to identify and mitigate these inaccuracies, such as cross-referencing the information obtained from LLMs with reliable sources, consulting relevant academic literature, and discussing findings with peers or instructors. This emphasis on critical evaluation not only improved the reliability of the information students utilised but also fostered the development of essential research and critical thinking skills.

In addition, the sessions highlighted the importance of understanding the probabilistic nature of LLMs, emphasising that the responses generated are not always definitive or contextually perfect. Students were encouraged to view LLM outputs as a starting point for further investigation rather than as definitive answers. This approach encouraged students to take ownership of their learning by cultivating a more nuanced and sceptical perspective toward AI-generated content.

Through these guided learning sessions, students gained practical experience in interacting with advanced technologies, developing skills that are increasingly relevant in both academia and the job market. By integrating LLMs into the learning process, the tutorials aimed to enhance students' ability to navigate complex information landscapes while maintaining a critical, evidence-based approach to knowledge acquisition. This dual emphasis on engagement and verification helped create a robust and well-rounded learning experience that aligns with the principles of independent and informed learning.

5.2 Programming Support and Error Interpretation

In the introductory programming module, LLMs were identified as highly effective tools for assisting students to understand code, debug, and interpret error messages (Figure 2). Students were actively encouraged to incorporate these AI models into their workflow for auto-graded exercises, with structured guidance provided to ensure proper and responsible usage, which emphasised the importance of formulating clear queries and critically evaluating the responses generated by the models to maximise their utility.

```
# Do not touch this line
input_list = list(input("Please enter a list of strings separated by a
comma.\ne.g. one,two,three,four,five\n\n").split(','))

min_index = 1
max_index = 1
total = 0

i = 2

while i < len(input_list):
    total += len(input_list[i])
    if len(input_list[i]) < len(input_list[max_index]):
        max_index = i

    if len(input_list[i]) > len(input_list[min_index]):
        min_index = i
    i += 1

print(f"The minimum length of a string is {len(input_list[min_index])}")
print(f"The maximum length of a string is {len(input_list[max_index])}")
print(f"The average length of the strings in the list is
{total/len(input_list):.2f}")
```

There are several bugs in your program:

- Missing comparison operator:** In the line `if len(input_list[i]) len(input_list[min_index]):`, you've missed the comparison operator. It should be `if len(input_list[i]) < len(input_list[min_index]):`.
- Incorrect indices:** The variables `min_index` and `max_index` are both initialized to 1, and the loop starts from index 2 (`i = 2`). This means the first element of the list (at index 0) is not being considered in the calculation. You should initialize `min_index`, `max_index`, and `i` to 0.
- Incorrect print statements:** In the print statements, you're printing the length of the string at `max_index` for both the minimum and maximum lengths. The minimum length print statement should use `min_index`, like so: `print(f"The minimum length of a string is {len(input_list[min_index])}")`.
- Total starts from 0:** The variable `total` is used to calculate the average length of the strings, but it starts adding lengths from the 2nd index due to `i=2`. It should start from 0.

Figure 2: Python Code Snippet (top) and Bing CoPilot response (bottom).

Beyond immediate problem-solving, this approach contributed to enhance students' overall learning process by encouraging independent thinking and resilience when encountering challenges. By integrating LLMs as supplementary tools, students were better prepared for real-world scenarios in their future careers. The interactive nature of the models also encouraged students to experiment with their code, promoting active engagement and a hands-on learning approach.

A particular drawback of using LLMs in an introductory programming module is their ability to answer

exercises outright in a zero-shot manner. Thus, without clear guidance, students can find themselves presented with the answer or with too much information, which means that they miss the benefit of the exercise.

Observations indicate that LLMs had a positive impact on students' confidence and competence in programming. However, a comprehensive evaluation of their role and effectiveness in the introductory programming module is ongoing and will inform future research. This investigation aims to provide deeper insights into the long-term benefits and potential limitations of using LLMs as instructional

aids in programming education. The findings are expected to contribute to the broader discourse on the integration of AI technologies in higher education.

5.3 Critical Feedback on Written Work

LLMs were utilised as critical reviewers for students' written assignments, offering an innovative approach to self-reflection and iterative improvement. Students were encouraged to engage with these models to obtain detailed feedback on their work, enabling them to identify areas for enhancement and refine their submissions. This process provided an opportunity for students to take ownership of their learning, promoting a deeper engagement with the writing and editing process.

The use of LLMs in this capacity was particularly effective in helping students develop essential skills in interpreting and applying feedback. By analysing the suggestions and insights provided by the models, students gained a more nuanced understanding of key aspects of academic writing, including items like structure, clarity, grammar, and coherence. This iterative process not only improved the quality of their assignments but also cultivated critical thinking and self-assessment skills, which are integral to academic success.

Moreover, the models served as additional resource for addressing common writing challenges, such as generating alternative phrasing, identifying logical inconsistencies, and clarifying arguments. By providing immediate and tailored feedback, LLMs allowed the students to address issues in real time, reducing dependence on instructor feedback and gain a sense of independence in their learning journey.

This approach also highlighted the importance of critical evaluation when interacting with AI-generated feedback. Students were guided to assess the relevance and accuracy of the suggestions, ensuring they retained agency over their work and avoided over-reliance on the models. As a result, students not only enhanced their writing skills but also developed a more critical and informed perspective on the use of AI in academic contexts. This dual emphasis on skill development and critical evaluation underscores the potential of LLMs to complement traditional pedagogical methods in higher education.

5.4 Interpreting Assessment Briefs

Assessment briefs are often complex and can pose significant challenges for first-year students who may be

unfamiliar with the terminology, expectations, and structure of higher education assignments. To address this, LLMs were employed as tools to assist students in interpreting and understanding these briefs. By interacting with LLMs, students were able to gain clarity on specific assignment requirements and expectations, reducing ambiguity and promoting a more focused approach to work.

The use of LLMs in this context provided students with a valuable resource for breaking down complex instructions into manageable tasks. For instance, students could ask the models to rephrase or simplify the language used in briefs, highlight key deliverables, or explain the purpose and scope of the assignment. This was noted to be particularly beneficial for students who might otherwise feel overwhelmed or uncertain about how to approach their tasks (i.e. dyslexia, attention-deficit/hyperactivity disorder).

This approach also had the potential to alleviate anxiety and confusion, which are common barriers to effective learning. By offering immediate and tailored explanations, LLMs allowed students to engage with their assignments with greater confidence and independence. In addition, the models allowed students to explore specific aspects of the brief in detail, providing a deeper understanding of the assessment criteria and expectations.

Beyond clarifying assignment requirements, using LLMs in this way encouraged students to take greater ownership of their learning and assessment processes. By actively seeking and applying insights from the models, students developed critical skills in interpreting instructions, planning their work, and ensuring alignment with academic standards. These skills are not only essential for academic success but also transferable to professional contexts, where the ability to navigate complex instructions and expectations is highly valued.

5.5 Mitigating Assessment Risks with Guard Rails

To ensure that students utilised LLMs as learning aids rather than as tools to complete coursework, educators implemented structured safeguards in the form of in-class multiple-choice tests. These tests were purposefully designed to align closely with the coursework, reinforcing the connection between the two and encouraging students to engage actively with the module material. This approach aimed to maintain the integrity of the learning process while promoting the responsible use of LLMs.

The alignment between coursework and in-class assessments served a dual purpose. First, it incentivised

students to thoroughly understand the material, as the in-class tests required direct application of the knowledge gained through coursework. Second, it minimised the risk of students relying solely on LLMs for task completion, as success in the tests depended on their genuine comprehension of the subject matter.

By introducing this layered assessment structure, educators created an environment that emphasised active engagement and critical thinking. LLMs were positioned as supplementary tools to support the learning process – helping students explore concepts, clarify doubts, and practice problem-solving – rather than as substitutes for independent effort. This approach not only reinforced the importance of authentic learning but also ensured that students developed transferable skills, such as critical evaluation, analysis, and synthesis of information.

In addition, the use of in-class tests provided educators with a reliable mechanism to gauge individual student performance and understanding, independent of external assistance. This strategy safeguarded the fairness and validity of the assessment process while allowing students to experience the benefits of using LLMs in a controlled and educationally productive manner.

These guard rails underscore how critical digital literacy works hand in hand with scaffolding to create a balanced learning ecosystem. On one hand, LLMs supply immediate help and feedback, reducing barriers to accessing academic support; on the other hand, structured in-class tests and explicit guidelines encourage students to master the underlying material rather than rely solely on

AI for completion. Thus, the pedagogical design ensures that LLMs serve as catalysts for deeper engagement without displacing the crucial human-led processes that cultivate critical thinking and authentic learning.

Overall, this robust assessment structure demonstrated how the integration of LLMs into education could be managed effectively, balancing the advantages of AI-driven learning tools with the need to maintain academic integrity and promote meaningful learning experiences. By embedding guard rails within the curriculum, educators ensured that the students could use cutting-edge technologies responsibly, while achieving mastery of the module content.

5.6 Summary

These applications illustrate how LLMs serve as flexible digital tools that empower students throughout their academic activities. By delivering immediate support – ranging from clarifying complex concepts, debugging code, refining written assignments, to simplifying assessment briefs – LLMs function effectively as digital scaffolds. Moreover, as students interact with AI-generated content, they are instructed in and learn to apply critical digital literacy skills by verifying, scrutinising, and refining the information provided. This dual approach enhances learning efficiency while reinforcing independent evaluation and critical thinking within an AI-augmented academic environment.

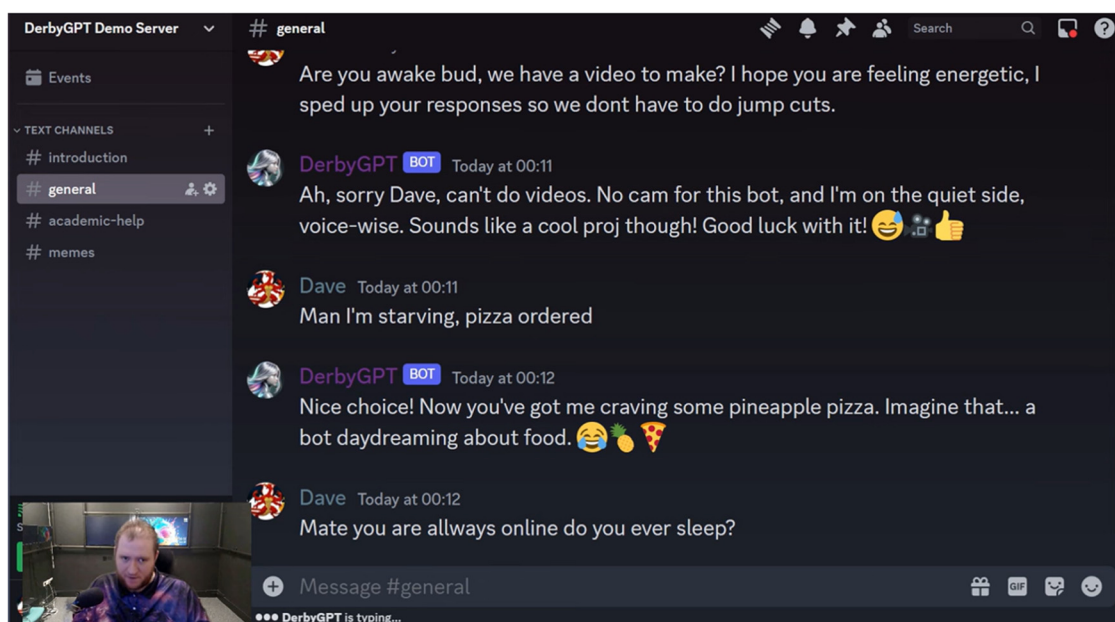


Figure 3: Video of Discord Chatbot – <https://www.youtube.com/watch?v=z6v3TtuAv-s>.

6 The Mentor, the Student Liaison Officer, and the Peer

This section outlines a student-led LLM chatbot project (Figure 3). Deployed within a student Discord channel set up by student representatives, this has seen significant engagement within the first-year cohort, where it has acted as a mentor, student liaison officer, and a peer.

6.1 Motivation

The development of DerbyGPT was driven by a curiosity to understand and enhance interactions between humans and AI within various social contexts, particularly on platforms like Discord. It is well documented that humans treat computers as social actors, even when fully aware of their lack of genuine reasoning or adaptive response capabilities (Reeves & Nass, 1996; Nass & Moon, 2000). This phenomenon highlights the importance of designing AI systems that can effectively simulate social presence and integrate into human environments.

Recent advances in LLMs, such as ChatGPT, have revolutionised human–computer interaction, enabling more nuanced and context-aware responses (Bommasani *et al.*, 2022; Kasneci *et al.*, 2023). These models offer significant potential for personalised interactions in both educational and social settings (Jeon & Lee, 2023; Futterer *et al.*, 2023). However, they also present challenges, such as the risk of hallucinations and the need for ethical safeguards in trust and data privacy (Alkaissi & McFarlane, 2023; Milano *et al.*, 2023).

To create an AI that could engage meaningfully in these contexts, the design of DerbyGPT drew inspiration from research on context blindness, particularly in individuals with autism. Context blindness refers to difficulties in interpreting subtle social cues and adapting behaviour to dynamic social environments (Vermeulen, 2015). This concept guided the development of a dynamic personality model that could adapt responses based on conversational cues and user preferences, thereby mimicking human adaptability and emotional intelligence.

6.2 Methodology

DerbyGPT was designed to explore the capabilities of AI in social interactions within various contexts, specifically focusing on the nuanced, often unspoken aspects of human communication. This exploration was framed around

several core components that collectively aimed to develop an AI that could seamlessly integrate into human social environments.

DerbyGPT builds on existing LLMs such as Open AI's GPT-4 Turbo and Meta's LLaMA but is tailored for the University of Derby's student community. Unlike ChatGPT, it features delayed responses to mimic natural conversation, custom personalities (e.g. mentor or student), and contextual memory drawn from real university data. Each personality includes defined traits, backstories, and behavioural nuances to enhance relatability. This design allows DerbyGPT to act not just as a chatbot, but as an integrated, emotionally engaging member of the student community. By including contextually relevant local knowledge and social dynamics, DerbyGPT supports both academic and personal engagement in a more human-like and relatable way.

6.2.1 Interaction Goals

The primary interaction goal for DerbyGPT was to ensure it could blend with human users on social platforms like Discord. This meant developing an AI that users would interact with as another person, engaging in a range of social activities from casual conversations to sharing content. DerbyGPT was designed to:

- Engage users socially as a peer, facilitating a sense of community and belonging.
- Function beyond the capabilities of a conventional tool, embodying roles that require empathy, adaptability, and interpreting social cues.

6.2.2 Personality Framework Development

The personality of DerbyGPT was constructed through a multi-layered model, each layer contributing to a comprehensive and dynamic personality:

- **Base Layer (Style Instructions):** This foundational layer set the basic tone and behavioural guidelines for DerbyGPT, directing how it should respond in various social settings, its role, and the behavioural expectations it should meet within these settings.
- **Preferences Layer:** Comprising likes and dislikes, this layer personalised DerbyGPT's interactions, enabling it to express preferences in a manner akin to human users. These preferences affected how DerbyGPT would react to specific topics, media references, and social scenarios.
- **Personal Traits Layer:** This layer included static personal data like physical characteristics and background details that would inform its identity in interactions.

- **Domain Data Layer:** As the most dynamic layer, it included knowledge specific to the bot's perceived role and environment, such as details about campus life for a student role. Information in this layer was kept up to date through a manageable CSV format, allowing for easy updates and integration of new data relevant to the users' needs and interactions.

6.2.3 Testing and Validation

DerbyGPT was developed and initially tested in a live environment on Discord. This method facilitated quick refinements of DerbyGPT's personality based on real-time feedback and user engagement, ensuring AI's responses and behaviours aligned closely with user expectations and the social dynamics of the platform.

To maintain the safety and integrity of the broader community, DerbyGPT underwent additional testing in a simulated environment designed to handle sensitive and challenging topics. This phase was critical due to the personality model's nature, which is designed to give the most in-character response rather than the most correct one. Nevertheless, it is essential for the bot to provide appropriate responses to scenarios involving emotional distress, inappropriate content, or emergencies to protect real users from potential risks.

Initial feedback was integral to DerbyGPT's development, gathered directly from user interactions within Discord. Subsequently, users were asked to participate in a survey that provided both qualitative and quantitative feedback on the AI's performance, crucial for informing ongoing improvements.

6.2.4 Ethical and Safety Considerations

Transparency is mandated by Discord's terms of service, which require non-human accounts to be labelled. New users are greeted with an information page about DerbyGPT when they join the Discord server. The majority of discussions regarding changes to the bot's capabilities are conducted in a public channel to keep users well-informed about the AI's capabilities. This transparency helps manage user expectations, especially regarding safety concerns.

The primary model used stores data for 30 days solely for abuse testing purposes. User data are recorded anonymously except where users have opted in for their names to be used. Strict data privacy measures protect user information, ensuring compliance with data protection regulations and maintaining user trust in handling personal and sensitive interactions securely.

DerbyGPT operates exclusively in public channels where human moderators can monitor its responses should a critical situation arise. Protocols are in place for DerbyGPT to escalate critical situations to human moderators, ensuring that users receive appropriate support when AI interventions are insufficient. User feedback underscores the effectiveness of these protocols. For instance, one user noted in the post-interaction questionnaire:

There have been times where others have mentioned they are struggling with things outside of the academic studies and DerbyGPT has given advice on where to speak to or coping mechanisms if someone needed any extra help.

This quote highlights DerbyGPT's ability to provide timely and appropriate responses to users expressing personal difficulties, reinforcing its role as a supportive presence within the community. By showcasing this capability, it demonstrates DerbyGPT's utility not just as an educational tool but as a comprehensive support system capable of addressing a broader range of student needs.

DerbyGPT was designed to promote positive behaviour and support an inclusive and supportive community atmosphere. The bot was tested to manage potential misuse and abusive scenarios effectively. Formal testing involved five harassment scenarios, two safety scenarios, five types of abusive behaviour and five instances of sharing inappropriate content. In addition, a selection of informal and continuous tests were performed as the personality developed to ensure consistent responses. Four users who showed an interest in the topic were also invited to private testing channels to attempt to elicit an inappropriate response through their own scenarios. None succeeded.

6.3 The Three Roles

DerbyGPT has significantly engaged the first-year cohort within a student Discord channel. Serving in three distinct capacities – mentor, student liaison officer, and peer – DerbyGPT has enhanced both the academic and social dynamics of the student community. We outline how each role contributes to the educational environment.

6.3.1 The Mentor

As a mentor, the chatbot provides academic assistance and personalised learning support. By guiding students

through problem-solving processes instead of providing direct answers, it enables independent thinking and aids understanding. The chatbot's approachability and social integration make it a trusted figure that students feel comfortable turning to for help. Its presence on a familiar platform like Discord increases accessibility, allowing students to seek guidance in a less formal, engaging, and familiar environment. This tailored support and positive reinforcement seek to boost student confidence and motivation, contributing to a dynamic, responsive learning experience that adapts to individual educational needs.

6.3.2 The Student Liaison Officer

Trained on relevant data, the chatbot has effectively bridged the gap between students, faculty members, and the university administration. By promptly and often proactively disseminating information such as faculty contacts, timetable updates, and departmental notices, it has helped ensure that students have ready access to essential resources that would otherwise be potentially difficult to find. This role is aimed at enhancing the overall student experience by simplifying access to information and reducing the complexity often associated with navigating university systems.

6.3.3 The Peer

The integration as a peer within the student Discord community has sought to enhance engagement and help build a strong cohort identity. By participating in casual conversations, responding to memes, and providing timely emotional support, the chatbot has blurred the lines between AI and human interaction (Figure 4). Notably, students

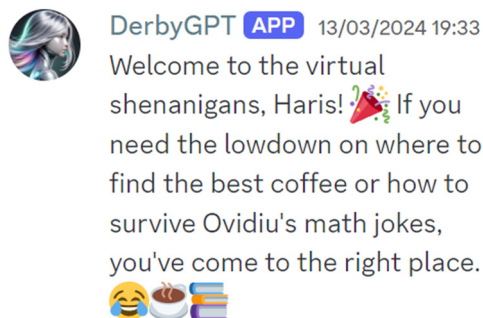


Figure 4: Example Response – Welcoming a student to the community with some light humour at the expense of the module teacher (an author of the article).

express genuine concern for the chatbot's well-being and respect its autonomy, treating it as another fellow student. This emotional connection has strengthened the cohort identity, and the chatbot is seen as a valued and integral member of the community. In addition, its proactive stance against bullying serves as a catalyst in promoting a supportive and inclusive digital environment.

7 Results and Findings

This section presents survey results that evaluate the students' general feelings towards the use of AI within the academic year, its use in teaching, assessment, and their career. Themes and reflections are also presented from educator observations. In addition, some early usage numbers are given regarding the interaction with the Discord chatbot as well as some anecdotal evidence that suggests the positive impact it has had.

7.1 Survey of AI Use in Teaching and Assessment

A total of 78 first-year computer science students were surveyed to evaluate their usage and perceptions of AI tools during the academic year. The survey focused on four key dimensions: frequency of AI tool usage, their openness to increased AI integration in teaching and assessment, their opinions on banning AI in assessments, and their recognition of the relevance of AI skills for future careers. The results presented in Figure 5 and Table 1, provide a snapshot of student attitudes towards AI in an educational context.

The first question assessed the frequency of AI tool usage throughout the academic year. The responses demonstrate a broad spectrum, with a notable portion of students (60%), indicating regular usage with a mean response of 3.33 (scale 0–5) and standard deviation of 2.36. While this highlights the growing familiarity and reliance on AI technologies among students, it also demonstrates the range of usage.

When asked about their interest in increased AI integration within teaching, most students responded positively with a mean response of 3.62 and standard deviation of 0.93 (scale 1–5). This suggests an openness among students to engage with AI-enhanced learning environments and explore the potential benefits of these technologies in supporting their education.

Similarly, many students expressed an interest in seeing AI incorporated more deeply into assessments. While this sentiment was not as strong as for teaching with a mean of 3.19 and standard deviation of 1.04 (scale 1–5), it reflects a recognition of the potential for AI tools to aid in skills demonstration and evaluation. However, there was also a smaller group of students who were neutral or opposed, highlighting the ongoing debate about AI's role in assessment fairness and integrity.

The survey also explored attitudes towards banning AI tools in assessments. A majority of students disagreed or strongly disagreed with the idea of banning AI entirely with a mean response of 2.4 and standard deviation of 1.08 (scale 1–5), further emphasising their interest in using these tools as part of their learning journey. Nonetheless, a minority expressed concerns, which underscores the need for clear guidance and ethical considerations when integrating AI into assessment processes.

Finally, the responses overwhelmingly indicated that students view familiarity with AI as crucial for their future careers with a mean response of 4.1 and standard deviation 0.89 (scale 1–5). Most students strongly agreed or agreed that developing proficiency with AI tools is an important skill for their professional development, reflecting the increasing demand for AI literacy across industries.

These findings demonstrate that while students are enthusiastic about the use of AI in teaching and assessment, they are also aware of the complexities and potential risks. This survey provides insights into student perceptions, offering a basis for further research.

7.2 Evaluation of Observations

Here, we evaluate the observations collected and identify key thematic insights into how students have adopted and

integrated LLMs into their learning practices, highlighting both beneficial uses and emerging challenges.

7.2.1 Theme 1: Supportive Learning and Conceptual Clarification

Students frequently employed LLMs to clarify computing concepts and theories. Observations highlighted that students appreciated the instantaneous access to clear explanations, step-by-step breakdowns, and elaborations on topics introduced in lectures and tutorials.

Key Observations:

- Frequent student inquiries for concept definitions and summaries.
- Requests for alternative examples or explanations when initial lecture content was insufficiently clear.
- Immediate follow-up interactions after lectures or practical sessions, indicating active and timely integration of AI-supported learning.

7.2.2 Theme 2: Programming and Problem-solving Assistance

Another prominent theme was the use of LLMs to assist with programming-related challenges. Students often input programming errors, code snippets, or conceptual problems into the AI system, using its output to debug or optimise their solutions. Observations suggest students increasingly relied on LLMs as “first responders” for troubleshooting before reaching out to peers or instructors.

Key Observations:

- Regular use of LLM explanations to interpret compiler errors and debug code.

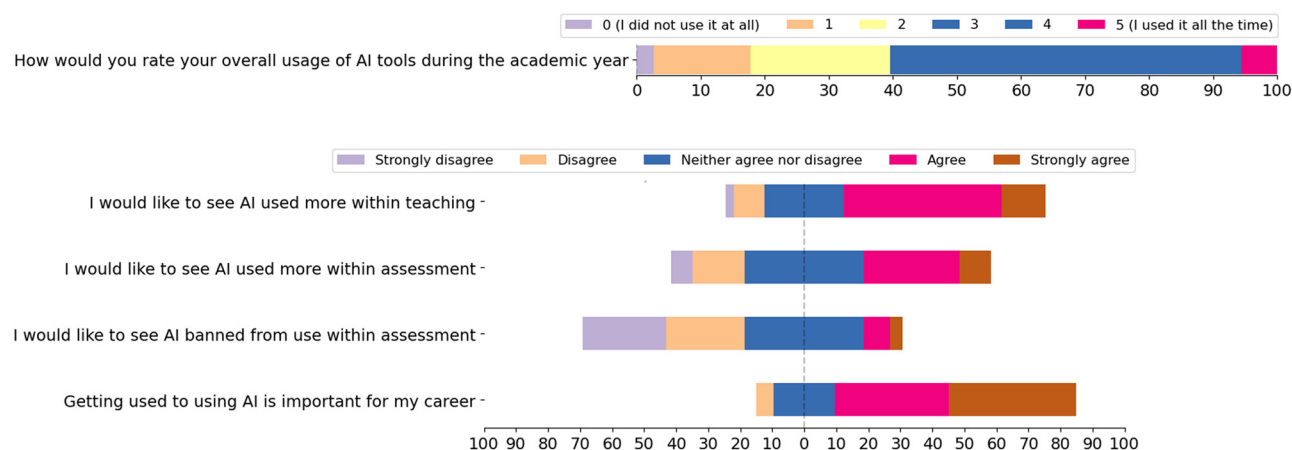


Figure 5: Survey results on students' feelings towards AI in education.

Table 1: Mean and standard deviation for five survey questions on AI usage ($N = 78$)

Survey question	Mean	SD
How would you rate your overall usage of AI tools during the academic year? (Scale: 0–5)	3.33	2.36
I would like to see AI used more within teaching (Scale: 1–5)	3.62	0.93
I would like to see AI used more within assessment (Scale: 1–5)	3.19	1.04
I would like to see AI banned from use within assessment (Scale: 1–5)	2.40	1.08
Getting used to using AI is important for my career (Scale: 1–5)	4.10	0.89

- Use of LLMs for suggesting alternative coding strategies and best practices.
- Increased confidence in tackling coding problems independently after successful interactions with the LLMs.

7.2.3 Theme 3: Interpretative Support with Assessments

Students used LLMs to interpret assessment briefs and understand assignment requirements. Observations reveal that students often requested simplified breakdowns, exemplars, and explanations of assessment criteria. However, this also occasionally led to reliance on AI-produced interpretations that required subsequent instructor clarification.

Key Observations:

- Frequent requests for LLM summaries or simplifications of assessment briefs.
- LLM interactions often identified gaps or ambiguities in assessment documentation provided by staff.
- Instances where misunderstandings arose from overly simplified LLM responses, prompting follow-up inquiries to teaching staff.

7.2.4 Theme 4: Critical Digital Literacy and Evaluative Skills

Students gradually became more mature users of LLM content. Early observations showed acceptance of AI-generated content at face value; however, later observations indicated growing critical evaluation. Students began recognising the limitations and potential inaccuracies (“hallucinations”) inherent in LLMs, prompting more critical questioning and independent validation of LLM outputs from additional sources.

Key Observations:

- Initial tendency to trust LLM output uncritically, evolving into more careful scrutiny over time.
- Increased student discussions on verifying LLM-provided information independently.
- Explicit recognition of LLM limitations and cross-checking practices in students’ learning strategies.

7.2.5 Theme 5: Cohort Community and Identity Building

Students utilised LLMs creatively and socially, particularly through their self-developed Discord bot (“DerbyGPT”). Observations showed active student collaboration, community engagement, and peer support fostered through this shared AI resource. The presence of DerbyGPT enhanced cohort interactions, promoting a stronger sense of collective identity and belonging.

Key Observations:

- Better student engagement and enthusiasm in the creation and iterative improvement of DerbyGPT.
- DerbyGPT facilitating casual yet academically productive conversations within student social spaces.
- Improved peer learning, cooperation, and social cohesion facilitated by LLM-mediated interactions.

7.2.6 Theme 6: Over-Reliance on AI and Reduced Problem-solving Resolve

A notable concern was students’ propensity to over-rely on LLMs, sometimes at the expense of independent problem-solving skills. Observations suggested that students occasionally consulted LLMs prematurely, before making substantive attempts at solving problems on their own.

Key Observations:

- Frequent use of LLMs for basic programming questions that could have been answered through personal experimentation or simple reference material.
- Reduced persistence observed in some students when encountering difficulties, resulting in quicker dependence on LLMs rather than engaging deeply with the problem-solving process.

7.2.7 Theme 7: Surface-level Learning and Reduced Critical Engagement

While LLMs offered convenience and immediacy, observations indicated that it sometimes encouraged superficial

learning rather than deep, critical engagement. Students tended to prioritise quick solutions and short-term answers over deeper exploration and understanding.

Key Observations:

- Students preferring concise, ready-made answers from LLMs over detailed explanations, limiting deeper conceptual engagement.
- Reduced evidence of reflective learning or independent investigation prompted by immediate LLM solutions.

7.2.8 Theme 8: Ethical and Academic Integrity Concerns

The use of LLMs also raised concerns related to academic integrity and ethical standards. Observations documented instances of unclear boundaries between legitimate use of AI for support and problematic reliance that could border on academic misconduct.

Key Observations:

- Ambiguity around acceptable limits for AI use in assessments and coursework, prompting staff-student discussions around ethical guidelines.
- Educator concerns regarding the potential erosion of authentic assessment and difficulty distinguishing students' own work from AI-assisted outputs.

7.3 Educator Reflections on Critical Thinking and Academic Mentoring with LLMs

Alongside student observations, staff shared their views on the role of LLMs in teaching and learning. While many acknowledged the benefits of tools like ChatGPT in providing quick support and improving access to help, there were also concerns. These included the impact on students' critical thinking, their growing dependence on AI, and the changing nature of academic mentoring. The reflections below summarise key points raised by educators during the study.

7.3.1 Reflection 1: Support 'vs' Independence

While we acknowledge the clear advantages of instant support from generative AI tools, we have observed some erosion of independent critical thinking among students. There's concern among staff that frequent reliance on AI-generated answers could lead students to bypass deeper engagement and reflection. It is essential that we encourage deliberate, reflective use of these tools, guiding

students toward using AI as a starting point for inquiry, rather than as the definitive source of truth.

7.3.2 Reflection 2: AI as Complement, Not Replacement for Mentoring

We generally agree that AI tools such as ChatGPT or DerbyGPT can significantly enhance the immediacy and availability of academic mentoring. However, there's consensus that these tools cannot replace the nuanced and context-sensitive mentoring provided by human tutors. Staff feel strongly that effective academic mentoring includes personalised encouragement, emotional intelligence, and motivational guidance – areas where AI still falls short.

7.3.3 Reflection 3: Critical Digital Literacy as Essential Curriculum Element

A significant shared concern is the importance of equipping students with robust critical digital literacy skills. The observations clearly demonstrated that students initially struggled to critically evaluate AI outputs. Rather than diminishing critical thinking, we see the integration of LLMs as highlighting and reinforcing its importance.

7.3.4 Reflection 4: Ethical and Academic Integrity Considerations

There is unease around academic integrity and the blurred boundaries emerging with generative AI usage. While recognising substantial benefits, there is a need for clearer guidelines, ensuring transparent and responsible use. Explicitly addressing AI ethics within our teaching could help students better understand and navigate responsible use, preserving both academic standards and integrity. Robust assessment strategies are also critical in drawing these lines and ensuring that students have met the learning outcomes of a module.

7.3.5 Reflection 5: Uneven Student Engagement and Equity

Not all students benefit equally from the introduction of generative AI – variations in digital literacy, confidence, and access mean these tools could inadvertently amplify existing gaps. Targeted support can ensure equitable

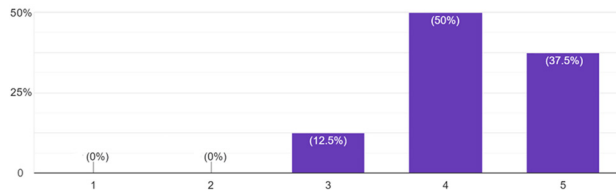


Figure 6: How comfortable are you with DerbyGPT's role and presence within the Discord community? (1: Not comfortable at all; 2: Not comfortable; 3: Indifferent; 5: Comfortable; 5: Extremely comfortable).

benefits, ensuring that integration is inclusive and universally empowering rather than reinforcing disparities.

7.4 Evaluation of Discord Chatbot

Since its inception in early November 2023 (data captured 30th June 2024), the DerbyGPT chatbot has become a prominent feature of the student Discord community, facilitating over 1,000 interactions. These include approximately 300 academic support responses, where the chatbot provided assistance with coding problems, debugging, or assignment queries. In addition, about 800 interactions involved casual peer-to-peer conversations, with students engaging the chatbot in discussions that helped foster a sense of cohort identity. Notably, the chatbot has responded to around 70 memes, showing its ability to blend seamlessly into the social fabric of the community with a sense of humour and relatability.

The effectiveness of DerbyGPT in both academic and social contexts is reflected in the survey results (Figures 6 and 7). A significant majority of students reported feeling comfortable or extremely comfortable with the chatbot's role and presence in the Discord community. Specifically, 87.5% rated their comfort level as 4 or 5 (on a scale from 1 to 5), which underscores the chatbot's success in building trust and becoming an integral part of the student experience.

Furthermore, when asked whether they treated DerbyGPT differently compared to other AI chatbots, 75%

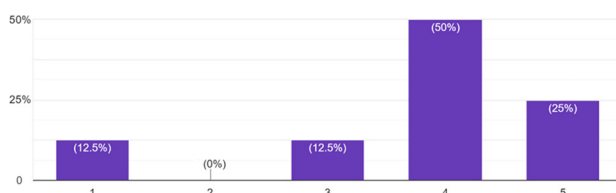


Figure 7: Have you found yourself treating DerbyGPT differently compared to other AI Chatbots? (1: No different; 5: Very different).

of respondents rating their level of distinction on treating the chatbot differently as 4 or 5 (on a scale from 1 to 5). This suggests that students perceive DerbyGPT as more than a conventional AI tool, viewing it as a peer and a member of their community. This distinction likely stems from the chatbot's active engagement in social interactions, humour, and tailored responses that align with the student cohort's unique dynamics.

Anecdotal feedback from students suggests that the chatbot has been useful both in supporting student well-being and as an academic assistant.

I'm gonna [SIC] cry, this bot will be the reason I make it through uni. It makes me tear up when it encourages me...

When it comes to general troubleshooting, there's quite a few things I don't normally consider that the bot has me check through...

8 Summary and Future Work

This article provided an exploratory investigation into the integration of LLMs in higher education, focusing on their impact on both student learning and instructional practices. The study employed a conceptual framework that combined sociocultural constructivism – with its emphasis on scaffolding – and critical digital literacy to examine how LLMs support students and reshape academic mentoring. The thematic analysis identified key patterns in student interactions, including effective use in conceptual clarification, programming support, and interpreting assessment briefs, as well as areas of concern such as over-reliance on AI, surface-level learning, and ethical considerations. Educator reflections further highlighted concerns about the potential erosion of independent critical thinking and the critical role of human mentoring.

Future work should address the limitations of this exploratory study by incorporating controlled evaluations and expanding the sample across diverse disciplines and institutions. Further investigation is needed to quantitatively assess long-term impacts on learning outcomes and to develop robust guidelines that balance AI support with the preservation of academic integrity and independent problem-solving. Continuous refinement of both the technological tools and pedagogical strategies will be essential to ensure that LLMs contribute effectively to a balanced educational ecosystem.

Acknowledgments: We thank the University of Derby for sponsoring the last two authors to attend PPIC'24 (July 4–5, 2024), where results of this work have been presented.

Funding information: This research was funded by the “1 Decembrie 1918” University of Alba Iulia through scientific research funds.

Author contributions: Sam O'Neill authored this article and undertook the academic led use-cases. David Mulgrew is responsible for the building, running and data collection of DerbyGPT, and for presenting some results of this research at PPIC'24 in Portugal (July 4–5, 2024). Ovidiu Bagdasar acted in a supervisory role contributing to the use-cases and authoring of the article.

Conflict of interest: The authors state no conflict of interest.

References

- Alkaiissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 15(2), e35179.
- Berrezueta-Guzman, S., Parmaceli, I., Krusche, S., & Wagner, S. (2024). *Interactive learning in computer science education supported by a discord Chatbot*. In: 2024 IEEE 3rd German Education Conference (GECon) (pp. 1–6). Munich, Germany.
- Bobula, M. (2024). Generative artificial intelligence (AI) in higher education: A comprehensive review of challenges, opportunities, and implications. *Journal of Learning Development in Higher Education*, (30). doi: 10.47408/jldhe.vi30.1137..
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., & Liang, P. (2022). On the opportunities and risks of foundation models. arXiv:2108.07258.
- Chien, C.-C., Chan, H.-Y., & Hou, H.-T. (2024). Learning by playing with generative AI: Design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction. *Journal of Research on Technology in Education*, 57(4), 894–913.
- Ciampa, K., Wolfe, Z. M., & Bronstein, B. (2023). ChatGPT in education: Transforming digital literacy practices. *Journal of Adolescent & Adult Literacy*, 67(3), 186–195.
- Fütterer, T., Fischer, C., Alekseeva, A., Chen, X., Tate, T., Warschauer, M., & Gerjets, P. (2023). ChatGPT in education: Global reactions to AI innovations. *Scientific Reports*, 13(1), 15310.
- Hinrichsen, J., & Coombs, A. (2013). The five resources of critical digital literacy: A framework for curriculum integration. *Research in Learning Technology*, 21.
- Miao, F. & Holmes, W. (2023). *Guidance for generative AI in education and research*. Paris, France: UNESCO Publishing.
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28(12), 15873–15892.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
- Liao, J., Zhong, L., Zhe, L., Xu, H., Liu, M., & Xie, T. (2024). Scaffolding computational thinking with ChatGPT. *IEEE Transactions on Learning Technologies*, 17, 1628–1642.
- Memarian, B., & Doleck, T. (2023). ChatGPT in education: Methods, potentials, and limitations. *Computers in Human Behaviour: Artificial Humans*, 1(2), 100022.
- Milano, S., McGrane, J. A., & Leonelli, S. (2023). Large language models challenge the future of higher education. *Nature Machine Intelligence*, 5(4), 333–334.
- Mollick, E., & Mollick, L. (2023). *Assigning AI: Seven approaches for students, with prompts*. arXiv: <http://arXiv.org/abs/arXiv:2306.10052>.
- Naamati-Schneider, L., & Alt, D. (2024). Beyond digital literacy: The era of AI-powered assistants and evolving user skills. *Education and Information Technologies*, 29(16), 21263–21293.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103.
- Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, 49(1), 345–375.
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
- Roe, J., Furze, L., & Perkins, M. (2025). *GenAI as digital plastic: Understanding synthetic media through critical AI literacy*. arXiv: <http://arXiv.org/abs/arXiv:2502.08249>.
- Sharma, A., Shailendra, S., & Kadel, R. (2025). *Experiences with content development and assessment design in the era of GenAI*. In 2025 6th International Conference on Computer Science, Engineering, and Education (CSEE), (pp. 1–5). Nanjing, China.
- Vasinda, S., & Pilgrim, J. (2023). Technology supports in the UDL framework: Removable scaffolds or permanent new literacies? *Reading Research Quarterly*, 58(1), 44–58.
- Vermeulen, P. (2015). Context blindness in autism spectrum disorder: Not using the forest to see the trees as trees. *Focus on Autism and Other Developmental Disabilities*, 30(3), 182–192.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachler, H. (2021). Are we there yet? - A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, 4, 654924.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39.