# Evaluating an AI agent to assist student learning in a 200-level Genetics practical

Gillian MacKay (Genetics), Rebecca Bird, Janic Gorman, Stephen Scott, Ryan Walker, Chen Wang (Centre for Educational Design and Innovation), University of Otago

**Abstract**

*A Cogniti chat agent was developed to scaffold student learning in a 200-level Genetics laboratory. The agent was piloted with 131 students. We analysed student chat logs, time spent using the agent and a student survey. Early findings suggest that the chat agent provides a useful tool for the students in a busy lab environment, but that staff-student interaction is still necessary.*

## Introduction

Practical laboratory sessions in Genetics can be demanding, busy environments where students often need to apply critical thinking skills to interpret instructions and apply concepts. In a 200-level Genetics Drosophila laboratory, students generate their own data and then analyse it with the support of a series of leading questions. The teaching staff normally guide the students through these questions in a very structured manner. This activity can be a bottleneck in the lab and students often have to wait for teaching staff to become available to help them. We developed a GenAI chat agent, using Cogniti, to help students work through these questions in a similar manner to that provided by the teaching staff. The aims of using the agent were to help reduce any initial confusion, support students' independent problem solving, and utilise the limited one-on-one time provided by staff more effectively. This pilot aligns with the University of Otago's intent to "embrace the potential of Artificial Intelligence (AI) to enhance educational experiences" and in Genetics, where AI is supporting advances in the field (Lin et al., 2025). We developed a course-integrated AI agent in the AI-Supported, learner-as-collaborator context (Elliott et al., 2025, Ouyang and Jiao, 2021). Cogniti provides a safe, controlled, and equitable exposure to an AI agent. We specified the agent's role, personality and behaviour with 12 stepwise guidance prompts mapped to the lab questions and based on current teaching practice. The agent was piloted with 131 students in the lab, with ongoing access after the lab. Once the students had worked through the questions with the agent, they were asked by the agent to review their work with a member of the teaching team. The data available from this pilot includes in-lab observations, an anonymous student feedback survey and anonymous Cogniti student chat transcripts. Ethics approval was secured for the student survey.

## Results and Discussion

The agent was expected to support student learning through a Socratic conversation (Ouyang and Jiao, 2021). However, during the lab, it was observed that some students did not follow the prompts and took the agent off track by trying to drive the agent towards giving answers. Experienced demonstrators were better at guiding student engagement with the agent, and getting the student back on track when they didn't follow the logical development of the discussion with the agent and got the agent confused. Likert responses were generally very positive about the use of the GenAI agent. However, the students were split on the agent's ability to develop critical thinking, and most still wanted to interact with a human (Table 1). Sixteen participants also provided written feedback on the chat agent. Responses were analysed using an inductive coding approach. Seven students described the agent as a useful tool, with two

noting that it helped reduce barriers to asking questions in the classroom. Four reported negative experiences, including technical difficulties with the interface or the agent's response. Four students felt the agent supported their critical thinking during the laboratory exercises, while three felt it hindered their ability to think through problems independently. Three students expressed broader concerns about the use of generative AI in university teaching, including ethical and environmental considerations. Four students indicated that although the agent was helpful, they preferred interacting with laboratory demonstrators or suggested that the tool would be most effective when used alongside in-person support, this is similar to Doğru and Faulconer (2025). Students spent an average of 38 min with the agent, 29% of the interaction was after the lab. The average number of queries was 25, with 30% being after the lab. Some students spent very little time engaging with the agent, but it was noted that many students worked in pairs, of their own accord, when engaging with the agent in the lab setting.

| Question | Positive | Neutral | Negative |
|---|---|---|---|
| How easy was it to use the chat agent? | 92% | 8% | 0% |
| How accurate did you find the chat agent? | 97% | 3% | 0% |
| How useful did you find the chat agent? | 82% | 8% | 10% |
| I prefer to talk to a chat agent vs a human | 24% | 16% | 60% |
| The chat agent encouraged me to think critically | 55% | 19% | 26% |

***Table 1: Likert responses, grouped into positive, neutral and negative, for 38 participants.***

Results suggest that while chat agents may provide useful scaffolding for some students, they are most effective when integrated with teaching staff rather than replacing teaching staff. The aim was to support students to interact with staff at a higher cognitive level. Future improvements will focus on refining prompts and enhancing demonstrator training.

1. Is the response to the agent consistent across student cohorts?
2. Should the agent be expanded to support other aspects of the lab?
3. Do we have to provide alternatives for students who have ethical objections to AI?

**References**


Doğru, M. S., & Faulconer, E. K. (2025). ChatGPT as a virtual laboratory teaching assistant in undergraduate biology. *Research in Science Education*, 56, 379–399. https://doi.org/10.1007/s11165-025-10271-z

Elliott, S., Richter, M., Alozie, C., & Raidis, D. (2026). Student perceptions of personalised GenAI agents on student learning, success, engagement, and wellbeing: A pilot study of Cogniti AI. Technology, Knowledge and Learning. https://doi.org/10.1007/s10758-025-09942-7

Lin, K. H., Kao, T. H., Wang, L. C., et al. (2025). Benchmarking large language models GPT-4o, Llama 3.1, and Qwen 2.5 for cancer genetic variant classification. NPJ Precision Oncology, 9, 141. https://doi.org/10.1038/s41698-025-00935-4

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. Computers and Education: Artificial Intelligence, 2, 100020. https://doi.org/10.1016/j.caeai.2021.100020