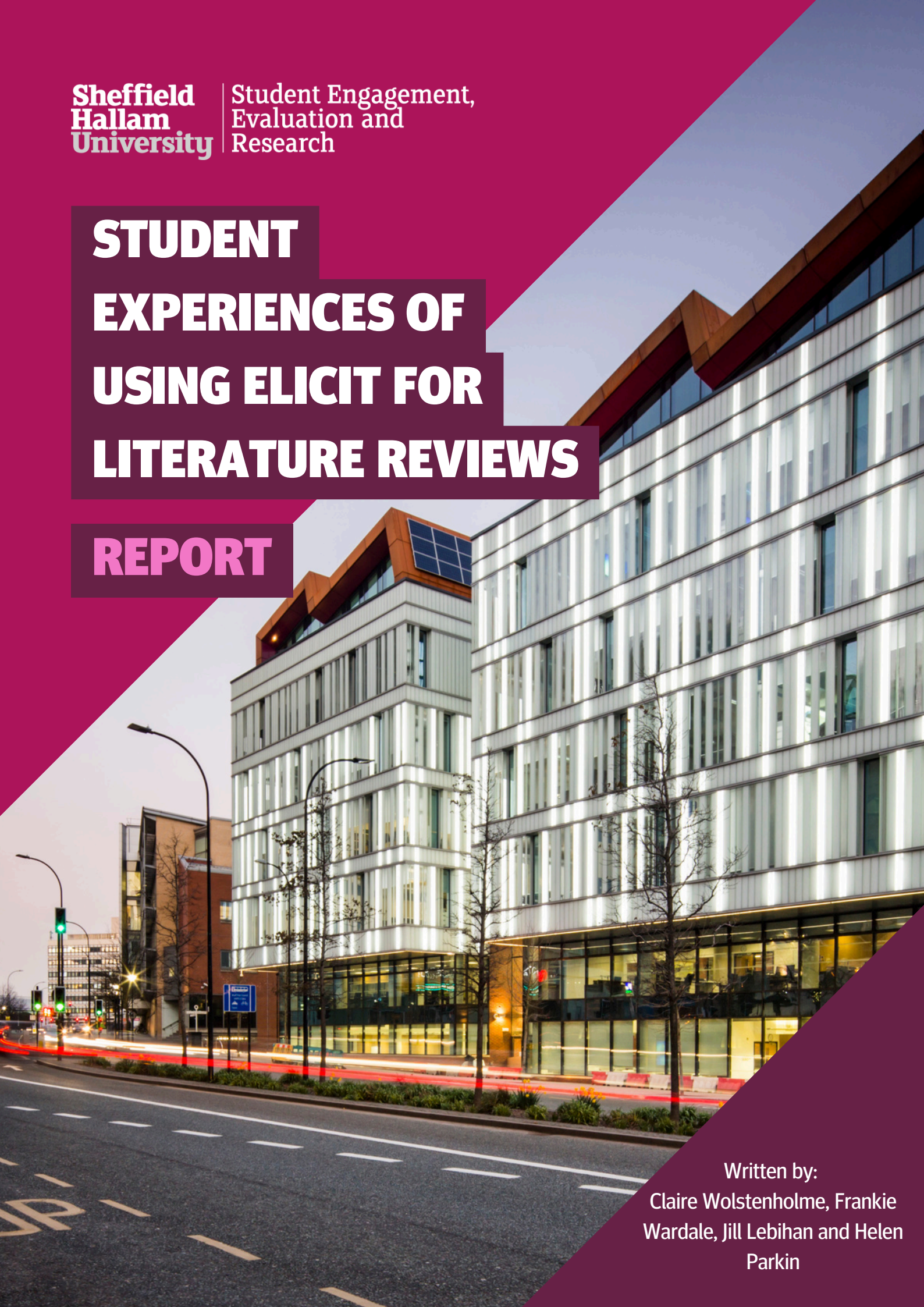
Sheffield Hallam University | Student Engagement, Evaluation and Research

STUDENT EXPERIENCES OF USING ELICIT FOR LITERATURE REVIEWS

REPORT



Written by:
Claire Wolstenholme, Frankie Wardale, Jill Lebihan and Helen Parkin

Student Experiences of Using Elicit for Literature Reviews

1. Who led on this project?

Frankie Wardale (Student Researcher), Claire Wolstenholme (Senior Lecturer, STEER), Jill Lebihan (Head of Student Engagement, STEER) and Helen Parkin (Senior Lecturer, STEER).

2. When did it take place?

October 2023 to August 2024.

3. What was the type of evidence?

Exploratory: Evidence of a specific topic that could be used to design an intervention.

4. Which stage(s) of the student lifecycle did it relate to (if any)?

Success (e.g. retention, attainment).

5. What question(s) was the project attempting to address?

- What are student perceptions of the usability of Elicit for writing literature reviews?
- To what extent does use of Elicit help or hinder the development of key research skills for students?

6. What need(s) or issue(s) was this research project addressing?

Artificial Intelligence (AI) tools are becoming increasingly available for university students to use in their academic work. Whilst these tools are relatively new, it is important to investigate their potential effects on students' skills development, with a view to shaping policy around their use. One such AI tool is Elicit: a language prediction model that searches databases of papers and journal articles to retrieve relevant academic work from a research question or topic. It is designed to reduce time and effort given to literature searches and reviews through automation. Elicit can output a summary of abstracts and a summary of the top 4-6 papers on the chosen subject. Similar to conventional tools, Elicit shows how many times a paper has been cited, but differs by providing the context of each citation. This allows users to view the influence that the paper has had on subsequent research. Elicit also provides links to the papers it finds, allowing the user to validate its claims.

A report in February 2024 by HEPI showed that more than half of UK undergraduate students surveyed had used AI tools in their academic work and 73% thought they

would likely use AI post university (Freeman, 2024). Literature, although not focussed on Elicit AI specifically, has suggested that student's overreliance on AI tools in their academic work, can lead to a decrease in critical thinking and cognitive reasoning skills (Zhai et al, 2024). Research undertaken to obtain students views of AI use has shown a mixed picture, with positive attitudes towards its use in academic study, alongside concerns over reliability, misconduct, and skills development (Chan & Hu, 2023). A plethora of research has been undertaken around AI use in university students, however a large proportion is based outside of the UK and no research to date has looked at student's use of Elicit AI specifically.

7. What was the aim(s) of the project?

This study aimed to understand student perceptions on the use of Elicit in their academic learning and assessment. Specifically, we wanted to understand if students felt a tool such as Elicit may help or hinder development of key research skills for students. Additionally, the project aimed to uncover the perceptions of usability of Elicit, and the perceived barriers and facilitators to its usage.

8. What did the research project involve?

Participants were asked to trial Elicit to help inform their literature review of an upcoming assignment task. They could spend as little or as much time using Elicit as they chose, and were advised about how to reference their use of Elicit in their assignment if applicable. Participants then took part in a focus group to discuss their views of Elicit, including its utility as a literature review tool and their perceptions of the potential impacts on students of utilising such tools.

9. What was the methodology used?

Ethical approval was granted by Sheffield Hallam University in January 2024. The method of data collection was a focus group in March 2024, which was scheduled to take place shortly after the assignment hand in dates. This was to ensure recall of use of Elicit whilst not interfering with assignment completion. One participant was unable to attend at the arranged time, and therefore an online interview took place with this participant separately in April.

Information and an invitation to take part in the research was sent to all registered student researchers (SRs) (spanning students across all levels and courses) who worked in the directorate of Student Experience Teaching and Learning (SETL) at Sheffield Hallam University. Students were paid for 2 hours of their time (as SRs) as an incentive to take part. The aim was to recruit 8 participants on a first come, first served basis. In total, 6 participants agreed to take part across a variety of courses, with levels

of study 4, 6 and 7 represented. Participant demographic details are not displayed to protect identities of participants.

The focus group lasted 50 minutes and the interview just under half an hour. Recordings were transcribed initially using Microsoft Word transcription, before a member of the team listened to the recording, corrected any mistakes made by the software, formatted the transcript appropriately and removed all sensitive data. Analysis was a two-stage approach. Firstly, the research team read through the focus group transcript making notes, before coming together for an analysis meeting. Data was analysed using a thematic analysis approach of the focus group and interview transcripts.

10. What learning, findings and / or recommendations can be shared?

Findings are presented below in brief by the research questions.

10.1: RQ1. What are student perceptions of the usability of Elicit for writing literature reviews?

Only one participant was aware of Elicit before this research**, most participants had not previously heard of Elicit** but were aware of other large language models (LLMs) such as ChatGPT. The predominant literature review tools used by participants were Google and Google Scholar, as well as databases on the Library Gateway. One student had utilised other AI tools (Jenni AI and Paper pal) to search literature.

Participants **overall viewed Elicit as a useful tool to provide a basic overview of a topic area**, but agreed that it had some significant drawbacks. One such issue was the summary of research papers provided by Elicit which was seen as insufficient:

> *It's trying to get like a basic idea of the whole thing in just one sentence, which is not ideal.* (Focus group participant)

This may however have been related to participants experience and understanding of Elicit capabilities and features that some users were not aware of. For example, the participant who took part in an interview felt that the summaries were much more useful:

> *To give you a heads up is actually wonderful - when you click on it you can actually have a preview of the main paper, so you can read in-depth through the site- I like it. ...Once I get a paper from Elicit, I go onto the gateway and look for it, and then I see if people have cited the paper.* (Interviewee)

As the interviewee was the only participant that had used AI tools beyond ChatGPT, experience and familiarity with different AI tools could play a more significant role in how effectively someone uses and perceives AI instead.

Elicit was also said by participants to **be less useful for subjects that were either fairly niche, or where topic areas intersected.**

Participants felt that **Elicit was more reliable and trustworthy than other LLMs** for example ChatGPT, which is known for hallucinating references, but they acknowledged that this was a fairly low bar for comparison. However, participants in the focus group felt that **Elicit was not able to give information on the quality and reliability of the source material** it had found, which was seen as a weakness.

There was some **disagreement as to the extent that using Elicit sped up the process** of finding sources of literature, with some participants feeling that it did not save them a significant amount of time:

It speeds up, you know what the source is... But it's not a part that's particularly arduous. (Focus group participant)

I feel like it was helpful for understanding, but it wasn't quicker to find the references. (Focus group participant)

Others felt using Elicit was helping them get the information needed more quickly:

Quicker than I can imagine, I will say that it's a lifesaver.... Within seconds I know this is the paper for me, and then I am in it. (Interviewee)

It's quite useful to get like a basic understanding of what people have done in the past because of the summary it gives out...I think it's definitely faster (Focus group participant)

When asked about their future use of Elicit, answers were polarised between the interviewee who felt they 'definitely' would use Elicit further, to a focus group participant who felt the opposite way:

I don't think I'd ever use it; I don't think I'd use it to replace what I usually use (Focus group participant)

The majority of participants stated that Elicit was fairly limited in its usefulness and usability for support with literature review writing.

10.2: RQ2. To what extent does use of Elicit help or hinder the development of key research skills for students?

The extent to which using Elicit might **help or hinder a student to develop research skills,** was felt by most to be dependent on the user's knowledge, experience and reliance on such tools:

Elicit will give you papers, but it's not doing critical analysis for you – if you don't do the critical analysis, you are not going to get the score for it. (Interviewee)

There were concerns that students with less experience in academic research may **become reliant on such tools,** and therefore not develop the necessary research skills:

I feel like it would make a lot of students reliant on it if it was from the start of university, because I feel like you'd kind of get used to using it. You'd use it all the time, and so you take it away, they wouldn't know what to do with it if it wasn't there. (Focus group participant

Not Elicit, it's ChatGPT - it has a way of bringing someone into a particular space and then you are so dependent on it, you don't think outside the box. (Interviewee)

The focus group discussion led to debate about potential **inappropriate use of AI tools** more generally, with some individuals considering AI use a form of cheating or laziness on students' part:

It was quite dangerous because I've heard of people getting high marks beforehand using stuff such as ChatGPT. (Focus group participant)

Others felt that where AI tools were used, work could become a demonstration of students' ability to use the tools rather than their ability to conduct research and/or their academic abilities:

It becomes, 'how good are you at using Elicit or ChatGPT, rather than how good a researcher are you' and you know getting a degree based on the fact that you are good at talking to this type of thing. (Focus group participant)

It just feels unfair because if there are ten people not using it and I'm the one using it and then I get like a significantly higher grade it is more of a reflection of what Elicit can do rather than what I can do. (Focus group participant)

Participants felt it was important for **universities to keep up to date with AI tools** to ensure that students were not able to use these in inappropriate ways:

It's better for like universities to adapt from now and get to know what like the dangerous part of AI is and how to prevent you know bad things. (Focus group participant)

However, there was also an acknowledgement that AI was a new technology likely to grow and therefore **students needed to better understand and utilise it**, with one participant likening it to satellite navigation:

It's like if you're a driver and you're still using this physical map… So, I think if we get to a point where everyone's OK with it and [there's] clear guidance on it, I think it could help you. (Focus group participant)

Participants felt they might be **more comfortable and inclined to use AI tools if there was a clear supportive stance from the university** and policies relating to what is and is not acceptable.

10.3: Conclusion

Findings indicate that there is uncertainty and concern from some students in regard to AI tools such as Elicit, both in relation to their usefulness, and the potential implication on skills acquisition. Students also had reservations about the potential for unjust and unethical use of AI tools, calling for clearer guidance around acceptable usage.

This research was limited in that it was small scale, and participants perhaps did not have the time or support needed to fully understand the functions of Elicit, and therefore they may not have developed a full awareness of its capabilities whilst attempting to use it. However, this allowed the research to highlight the perceptions of Elicit in a natural context as participants' existing knowledge, experience and understanding of AI affected their views.

11. References

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1). https://doi.org/10.1186/s41239-023-00411-8

Freeman, J. (2024). Provide or punish? Students' views on generative AI in higher education. Higher Education Policy Institute. https://www.hepi.ac.uk/wp-content/uploads/2024/01/HEPI-Policy-Note-51.pdf

Microsoft CoPilot. Accessed 19.11.2024. Used for support with writing conclusion and abstract.

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments, 11*(1). https://doi.org/10.1186/s40561-024-00316-7

Sheffield Hallam University

Knowledge Applied